

2011

An agent-based simulation model for business reopenings in New Orleans post Hurricane Katrina

Helbert Arenas

Louisiana State University and Agricultural and Mechanical College, harena1@tigers.lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations



Part of the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Arenas, Helbert, "An agent-based simulation model for business reopenings in New Orleans post Hurricane Katrina" (2011). *LSU Doctoral Dissertations*. 3675.

https://digitalcommons.lsu.edu/gradschool_dissertations/3675

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

**AN AGENT-BASED SIMULATION MODEL FOR BUSINESS REOPENINGS IN
NEW ORLEANS POST HURRICANE KATRINA**

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Geography and Anthropology

by

Helbert Arenas

B.S., Universidad Nacional de San Agustin, Arequipa, Peru, 2000

M.S., International Institute for Geo-Information Science and Earth Observation, Enschede,
The Netherlands, 2004

December, 2011

Acknowledgments

I believe most of Ph.D. graduates would agree that a Ph.D. research is a long exhausting endeavor. Mine could not have been possible without the valuable help of several people to whom I will always be thankful.

First I would like to thank my family in Peru for supporting and encouraging me all these years.

A pesar de la distancia he sentido su calor y apoyo todos los días a lo largo de estos años. Muchas gracias Mamá, muchas gracias Zare, ojala algún día pueda pagarles todo lo mucho que han hecho por mi.

My special thanks to my major professor, Dr. Nina Lam, without her none of this would have been possible. Along these years she has shown me the work dedication of a true scientist. Thanks Dr. Lam for all the things you have taught me and especially for your endless patience.

My special thanks to the members of my committee: Dr. Leitner, Dr. Liu, Mr. Braud and Dr. Breitenbeck. Thanks for your time, your valuable feedback and suggestions.

I would like to thank my family in Baton Rouge, Mauricio and Belen, my siblings in the exile, thanks guys. You really make my time in LSU and Baton Rouge an enjoyable experience. Finally I would like to thank my coworkers and fellow students, your support, advice and insights are really appreciated.

This research was supported by a Doctoral Dissertation Research Improvement (DDRI) Award (BCS-1003609) and research grant (BCS-0729259) from the National Science Foundation. Partial support from a research grant from the National Oceanic and Atmospheric Administration (NOAA) Gulf of Mexico Sea Grant Program and a research grant from the

Inter American Institute for Global Change Research (IAI) (award number: CRN-2050) is acknowledged.

Table of Contents

Acknowledgments.....	ii
List of Tables	vi
List of Figures	vii
Abstract	ix
Chapter 1. Introduction	1
1.1 Problem Statement	1
1.2 Objectives	2
1.3 Significance	4
1.4 Dissertation Outline	5
Chapter 2. Business Community in New Orleans after Hurricane Katrina.....	6
2.1 New Orleans and Its Environment	6
2.2 Hurricane Katrina	10
2.3 Business Decisions after a Disaster	11
2.4 Previous Research on Business Return Post Hurricane Katrina in New Orleans	14
Chapter 3. Estimation of the Temporal Dynamics of Business Reopening Predictors.....	22
3.1 Introduction	22
3.2 Generalized Linear Models	22
3.3 Spatio-temporal Dependence	27
3.4 Analysis of the Third Telephone Survey	34
3.4.1 Data	34
3.4.2 Methods	35
3.4.3 Results	37
3.4.3.1 Binary Logistic Regression Model Results	37
3.4.3.2 Probit Regression Model Results	43
3.5 Discussion	45
3.6 Conclusions	46
Chapter 4. Agent-based Models and Simulation in Geo-information Science.....	48
4.1 Models and Simulation	48
4.2 Alternative Modeling Approaches	50
4.2.1 Mathematical Models and System Dynamics	50
4.2.2 Microsimulation	51
4.2.3 Queuing Models	53
4.2.4 Cellular Automata	54
4.3 Agent-based Models	55
4.3.1 Model Verification and Validation	57
4.3.1.1 Verification	57
4.3.1.2 Validation	58

4.3.2 Trends in Agent-based Models	59
4.3.3 Applications of Agent-based Models in Geography	61
4.3.3.1 Abstract Models	61
4.3.3.2 Regionally or Locally Specific Models	61
4.3.3.3 Highly Detailed Urban Realistic Simulation	64
4.4 Conclusions	65
 Chapter 5. Development of an Agent-based Model for Simulating Business Reopening in New Orleans Post Hurricane Katrina	68
5.1 Introduction	68
5.2 Study Area and Data Requirements	69
5.3 Model Conceptual Design	70
5.4 Model Operational Design	71
5.5 Model Calibration	75
5.6 Model Validation	77
5.7 Disaster Scenarios	84
5.7.1 Scenario 1	85
5.7.2 Scenario 2	85
5.7.3 Scenario 3	90
 Chapter 6. Conclusions	93
6.1 Summary	93
6.2 Contributions	96
6.3 Future Research	97
 References	99
 Vita	107

List of Tables

Table 2-1 : Survey statistics of the three surveys (source Lam et al. 2009).....	16
Table 2-2: Business classification according to their NAICS code (source: Lam et al. 2009)	18
Table 2-3 : Business opening ratio by type at the three time periods (source: Lam et al. 2009) ...	18
Table 2-4 : Average ranks of barriers and prospect for all businesses surveyed.....	20
Table 3-1: Neighbor influence estimation for business 1 using $p = 2$	33
Table 3-2 : Reopening counts by time period	33
Table 3-3 : Descriptive statistics of the variables for each analysis.....	38
Table 3-4 : Binary logistic regression model results (significant coefficient * sig<0.05)	39
Table 3-5:Probit regression models , coefficients (B) and marginal effects (Meff) (* Sig<0.05) .	44
Table 5-1: Equations for the reopening probability estimation for each period.....	76
Table 5-2: Probabilities of opening using the logistic functions.....	78
Table 5-3:Overall accuracy for each threshold (NA: not available, *:complete misclassification of one class).....	78

List of Figures

Figure 2-1 : Orleans Parish Location Map.....	7
Figure 2-2 : Orleans Parish DEM using NED from USGS.....	7
Figure 2-3 : Hurricane Katrina approaching New Orleans	12
Figure 2-4 : Levee breaches in New Orleans after Hurricane Katrina	12
Figure 2-5 : Flood depth and kernel density maps of businesses reopened by the three time periods	17
Figure 3-1: Probabilities estimation using a probit function.....	26
Figure 3-2 : Business density for firms included in Survey 3 A) Before Katrina, B) Right after Katrina, C) By September 30th 2005, D) By December 30th 2005, E) By March 31st 2006, F) By June 30th 2006, G) By September30th 2006, H) By October 2007.....	30
Figure 3-3: Estimation of neighbor influence on business 1	33
Figure 5-1: Calibration and Validation points.....	70
Figure 5-2: A)Model Flow chart, B)UML (Unified Modeling Language) design diagram of the business agent class	73
Figure 5-3: New Orleans neighborhoods.	80
Figure 5-4: Model evaluation by December 2005, A) Business density according to first survey (December 2005), B)Business density based on model results by December 2005, C) Differences between model and survey data.....	81
Figure 5-5: Model Evaluation by June 2006 A) Business density according to first survey (June 2006), B)Business density based on model results by June 2006, C) Differences between model and survey data.	82
Figure 5-6: Percentage of addresses actively receiving mail at the census tract level ((total addresses-(vacant-no stat))/total addresses).	84
Figure 5-7: Business density for scenario 1:A)By August 30 th 2005, B)By September 30 th 2005, C) By December 31 st 2005	86
Figure 5-8: Business density for scenario 1 (cont.): A) By March 31 st 2006, B) By June 30th 2006, C) By September 30th 2006.	87
Figure 5-9: Businesses density for scenario 2 and density difference between scenario 1 and scenario 2 : A) Immediately after Katrina, B) By December 2005, C)By June 2006, D)By September 2006.	89

Figure 5-10: Selected census block groups for test areas91

Figure 5-11: Business density for Scenario 3, and density differences between scenario 1 and scenario 3: A)By August 30th 2005, B) By September 30th 2005 , B) By December 31st 2005, D) By September 2006.92

Abstract

The empirical study of business responses to disasters is relatively scarce, despite that they are a fundamental part of the cities, providing services, jobs and, taxes that are essential for urban sustainability. This dissertation develops an agent-based simulation model to represent and understand the businesses reopening process in a dynamic environment in New Orleans after Hurricane Katrina. The objectives are two-fold: 1) To identify the main reopening predictors involved and estimate their relative importance through time, using an empirical data set collected from another study; 2) To represent the business reopening process through a computer simulation model, using the parameters derived from the first objective.

The results show that businesses located in flooded areas had lower reopening probabilities, however the effect was significant only in the first nine months after the disaster. Larger businesses had better reopening probabilities than smaller ones, although this variable stopped being significant after six months. Variables associated with higher social vulnerability, such as percent non-white population and percent population under 18, had a negative effect on the business reopening probabilities at different points of time. The influence of neighboring firms using 1-km buffer was found significantly positive only immediately after the disaster; it became significantly negative one year after the disaster.

The simulation model developed proved to mimic the reopening process at a suitable level. The model was used to simulate two scenarios: 1) First, the flood depth was reduced by 1 meter as a way to represent the implementation of measures designed to increase the buildings and infrastructure resistance to floods. The simulation results indicate that there are specific areas that would obtain greater benefit from these measures, however ten months

after the disaster the effect of the measures tends to diminish. 2) Second, the spatial effects of aids were simulated by making a limited number of businesses in specific locations totally resilient to the disaster. The results indicate that the beneficial effect is influenced by variables such as business density and socio-economic conditions of the area. The positive effect is perceivable until four months after the disaster, after this point it diminishes.

Chapter 1 : Introduction

1.1 Problem Statement

Most cities have proven through the history to be extremely resilient to natural disasters (Vale and Campanella 2005). However policies designed to make the post disaster recovery sustainable still need further study. Cities are complex environments in which multiple variables and components influence each other. Understanding the complex environmental system dynamics, the cause-and-effect relationship, and the interactions between human and natural components is critical to the design of policies aimed to ensuring sustainable development of cities and regions, especially when they are faced with the threats of both long-term climate change and short-term large-scale disturbance (Lam et al. 2009).

This research studies the effects of Hurricane Katrina in the business community of New Orleans. Hurricanes are the costliest disasters in the United States and in recent years their costs have increased dramatically (Pielke and Landsea 1998). The total cost of hurricanes in the 1990s was higher than the 1970s and 1980s combined. The higher cost is explained by the increased vulnerability. The population density in coastal areas has seen a steadily increase in the last decades (Rappaport and Sachs 2003). More people and infrastructure are located in areas threaten by hurricanes than before, making the disaster costlier when they occur.

Until recently most of the research in the field of post-disaster recovery has been focused on households or whole communities (Webb, et al. 2002). Businesses are a fundamental part of the cities. They provide services, jobs and taxes essential for urban sustainability; however, the empirical study of business responses to disasters is scarce (Lam et al.2007; Lam et al. 2009b).

Any long-term recovery policy must address the economic aspect of community recovery. In order to design better recovery policies, it is necessary to identify the significant recovery predictors and evaluate how they affect the return of businesses after a disaster. Most of the previous literature that focuses on business recovery indicates that businesses' responses to disasters are not random. There are statistically significant recovery predictors associated with a positive or negative outcome. The recovery process is a dynamic one. However, few studies were able to develop empirical models that used data collected almost immediately after the disaster. Most of the previous studies were based on surveys carried out several months, if not years, after the event, losing important temporal information. Moreover, as far as we know, there has not been any study that uses a dynamic simulation approach to examine the business recovery process. To understand the business recovery process, it is important to study the dynamics of the process using empirically derived parameters.

As in most human systems the business recovery process is an open one. There are many uncertainties that can not be fully addressed. The purpose of this research is to understand decision making under uncertainties by businesses and provide decision makers with information and tools to design better recovery policies. This study identifies relevant recovery predictors and provides tools to test "what if" scenarios to test recovery policies.

1.2 Objectives

This research develops an agent-based simulation model to represent and understand the businesses reopening process in a dynamic environment in New Orleans after Hurricane Katrina. The business post-disaster recovery is a complex phenomenon. Businesses are heterogeneous in nature, size, economic activity, and others. Based on their locations, a disaster also affects them in different forms. Agent-based models are particularly well suited

for modeling business recovery because they represent each business as an independent component with its own characteristics, allowing the representation of complex interactions between businesses themselves as well as with their geographic environments. A model is a simplified representation of reality that helps researchers to understand a real world process. An agent-based simulation model will allow researchers to conduct experiments that otherwise would be too expensive or difficult to conduct in the real world (Gilbert 2008). The reopening simulation model will not only provide understanding of the process, but also help planners design and test recovery policies so that resources can be better allocated.

The objectives of this dissertation research are two-fold. The first is to understand the business reopening process, in order to estimate rules and parameters to be used in the simulation model, making it as realistic as possible. We will identify the main reopening predictors involved and estimate their relative importance through time. We will use the business survey and related GIS data collected by previous studies in this study area (Lam et al. 2007, Lam et al. 2009). The study variables will include characteristics of the business, socio-economic characteristics of the residents in the business vicinity, and the influence of neighboring establishments on the reopening decision process. The second objective is to represent the business reopening process through a computer simulation model. The simulation model will assume each business as an independent agent. The business agents will estimate at discrete points of time its reopening probabilities based on its internal characteristics as well as the influence of neighboring business agents. The influence of neighboring business agents is determined using a neighborhood function, such as an inverse distance-weighting function, that will assign more weights to closest neighbors. The simulation result will be visualized using an interface programmed in JAVA.

Once the simulation model is calibrated and validated, the effects of a business opening in one location on the reopening probabilities of neighboring businesses, and how such effects propagate through time, can be estimated and visualized by the simulation model. The scenarios generated by the empirically based model to be developed in this research will help increase our understanding of how city recovers in a complex environment and what planning or aid strategies would be the best to help speed up recovery.

1.3 Significance

This research contributes to geography and spatial sciences with a special focus on post-disaster economic recovery and dynamic modeling. This research is innovative in three aspects: First, there is no previous research on the simulation of post-disaster business reopening process. The use of simulation will help understand the complex recovery process and how the reopening of one business in a given location at a given time will affect the business reopening probabilities in other locations in later time periods. Second, the use of empirical data for model calibration and validation for an agent-based model is also innovative. A validated simulation model that is supported by real data will allow researchers to design and test recovery policies so that resources can be better allocated. Third, this study will estimate how the status of neighboring firms influence the business reopening decision process.

Although the simulation model is based on post-Katrina New Orleans data, it can also be applied to other locations affected by disasters of the same nature. The simulation and visualization program will be made available to the open source community to allow their use, modification and improvement.

A better understanding of how human systems respond to disaster events will allow policy makers to improve disaster preparedness plans. The reopening probabilities and the model results will be useful to planners for evaluating alternative recovery aid strategies. The simulation model will be a valuable tool not only for planning purposes, but also for hypothesis testing purposes.

1.4 Dissertation Outline

Chapter 2 is a literature review of the New Orleans landscape, hazards, disasters and previous relevant research on the impact of Hurricane Katrina on the business community in New Orleans. Chapter 3 describes the procedure employed to estimate the businesses reopening probabilities. Chapter 4 describes agent-based models, the approach employed in this research. Chapter 5 describes in detail the agent-based simulation model developed. Chapter 6 contains the conclusions and suggestions for further research.

Chapter 2 : Business Community in New Orleans after Hurricane Katrina

2.1 New Orleans and Its Environment

New Orleans is located next to the Mississippi river close to its mouth. It is one of America's leading cargo ports, and is the only deepwater port in the country served by six railroads of category one. In addition to the railroads, it has access to the highway system through the Clarence Henry Truckway. Its railroad and highway connections give the port a privileged access to the whole country. In 2004 a study conducted by Martin Associates indicated that the port activity originated 160,498 jobs, \$8 billion in earnings, \$17 billion in spending and \$800 million in taxes only in the state of Louisiana (Port of New Orleans 2011).

The city of New Orleans extends between the Mississippi river and Lake Pontchartrain (See Figure 2-1). The older parts of the city are located next to the river. This is explained by the economic importance of the river as well as by the city's unique topography. The lands adjacent to the Mississippi river are higher than all other natural features in the delta area. The French settlers named these areas as *levées*, which means "raised up". The elevation of the levees is due to the river sediments deposits over thousands of years. The natural levees have an elevation of 8 to 15 feet above sea level. On top of the natural levees, artificial ones have been built since 1719. After the Civil War, the levee system became the responsibility of the federal government, which increased the elevation of the levees in between 15 to 25 feet (See Figure 2-2) (Campanella 2006).

Until the beginning of the 20th century, the city of New Orleans was limited to the immediate surroundings of the river levees. The area suitable for urban development extended less than two miles from the river shore close to the current location of Clairborne Avenue (Kelman 2003). The rest of the land, between this point and the Lake Pontchartrain shore was

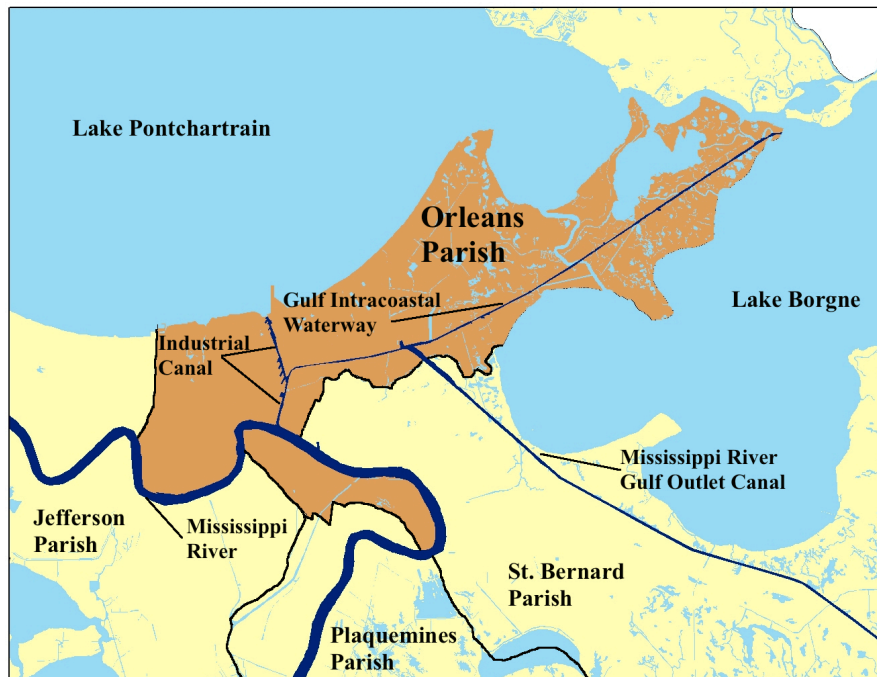


Figure 2-1 : Orleans Parish location map (map created by the author).

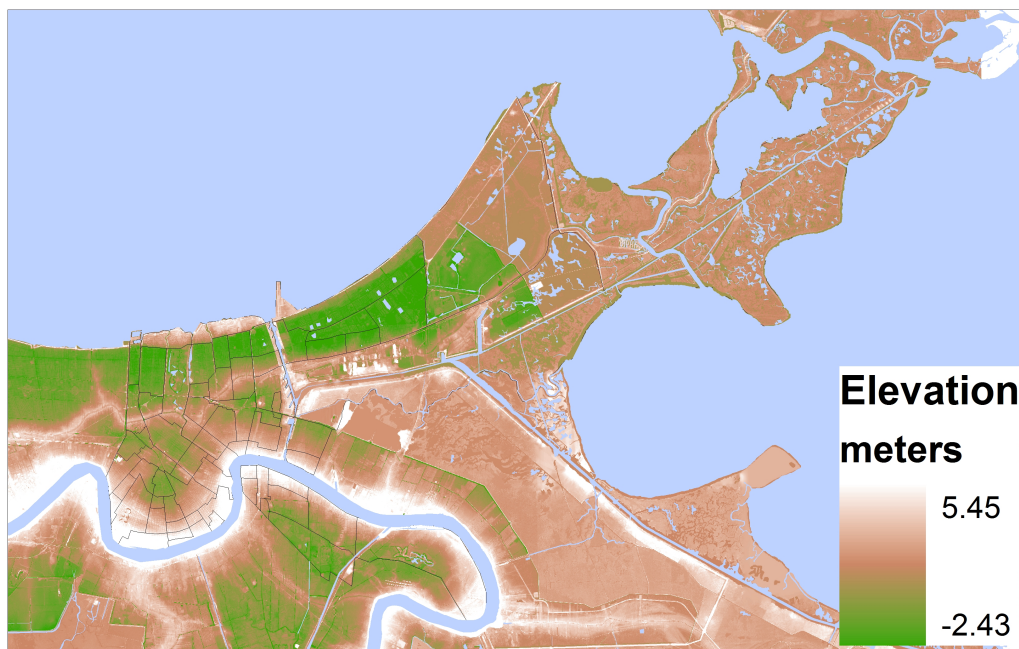


Figure 2-2 : Orleans Parish Digital Elevation Model using National Elevation Dataset from United States Geological Survey (map created by the author).

a marsh crossed by some bayou ridges, with an average elevation at or below sea level. Around 1895 sanitary engineers developed a system consisting of miles of canals and electric pumps. The goal of the system was to remove the water from the city. In 1914 the system was composed by almost seventy miles of canals and seven pumping stations. The system transforms the marshes next to the lake into dry land. The new reclaimed areas began to be rapidly urbanized around 1910. In 1890 the city imposed taxes on 132 million dollars in property, by 1914 the amount of property was almost twice the amount registered in 1890 (Kelman 2003).

Today second to the river levees, the Lake Pontchartrain shores are the areas with the highest elevation in New Orleans (See Figure 2-2). But it was not always like this. The newly developed land required some protection from hurricane surge coming from the lake that could flood the city. A system of levees was built around 1900 in this area, although this system had serious faults. A better defense mechanism was created based on a plan designed by W.H. Bell 50 years earlier. The idea was to create a levee half a mile into the lake and pump sediments from the bottom of the lake into the space created between the original shore and the new levee. This newly created area would protect the city from the lake. The project was completed in 1934. The new land was elevated 4 to 6 feet above the lake level, and more than 10 feet above the adjacent lowlands (Campanella 2006).

Because of its location and particular topography, New Orleans is threatened by floods because of hurricane surges and river high-water. Throughout its three centuries of history New Orleans has suffered 27 floods due to either of those causes. After each of these events the city has improved its defense mechanisms, rebuilding and raising its levee system (Kates et al. 2006).

Before Hurricane Katrina, in the 20th century, the city sustained severe damage in multiple occasions. On September 29th of 1915 a category 4 hurricane landed near Grand Isle, Louisiana. Due to the hurricane surge more than 25,000 buildings near the Lake Pontchartrain shore were flooded (Kates et al. 2006). Parts of the Mid-city district were flooded because of a failure in the pumping system due to damages in the power infrastructure.

The year of 1927 was especially difficult for the states in the Mississippi basin. Heavy rains starting in the summer of 1926 increased the river to unusual levels. Floods affected Arkansas, Illinois, Kentucky, Louisiana, Mississippi, Missouri, Tennessee, Texas, Oklahoma and Kansas. On April 15th of 1927 a power failure caused the pumping system ceased to work in New Orleans, and it could not be in a worse moment. That day New Orleans sustained 14 inches of rain. Because of the particular topography of the city and in the absence of the pumping system, water had nowhere to go and flooded the low elevation parts of the Big Easy (Kelman 2003). By the end of April, the river was full, it had broken levees at Mound Landing, Mississippi and the population in New Orleans was in the brink of panic. A controversial plan was proposed in an attempt to reduce the pressure in the river levees. The plan consisted in creating a breach, with explosives, downstream from New Orleans, allowing the water to flow to the sea alleviating the pressure. The plan was carried out on April 29th saving New Orleans from flooding. However the breach flooded St. Bernard and Plaquemines parishes effectively destroying them (Kates et al. 2006, Kelman 2003).

The night of September 9th of 1965, a category 3 Hurricane named Betsy landed in the southeastern Louisiana coast. The hurricane surge puts pressure in the levee system of the city eventually leading to the failure of the Industrial Canal and the Intracoastal waterway

levees which caused the flooding of the low areas of New Orleans, west of Industrial Canal and north of Gentilly Ridge (Goudeau and Conner 1968). Betsy destroyed more than 27,000 houses, causing the displacement of more than 300,000 residents. After Hurricane Betsy, the levee system was upgraded by the U.S. Army Corps of Engineers only to be tested again in 2005 with Hurricane Katrina (Kates et al. 2006) .

2.2 Hurricane Katrina

On August of 2005 the tropical depression Katrina formed on the Atlantic, on August 23rd it became Tropical Storm Katrina. It moved erratically westward over the Bahamas and became a category 1 hurricane, 15 miles northeast of Fort Lauderdale on August 25th. It made its first landfall the same day causing substantial damage and the lost of 14 lives. During the night of August 25th it moved westward over the Florida peninsula, reaching the Gulf of Mexico the next day. An upper anticyclone over the Gulf and warm sea surface temperature caused the intensification of Hurricane Katrina, which became a category 5 hurricane on August 28th. On August 29th Hurricane Katrina made its second landfall in southeast Louisiana as a category 3 hurricane (See Figure 2-3) (Graumann et al. 2005). The hurricane surge caused by Katrina increased the pressure on the levee system that protect navigation and drainage canals in New Orleans. The levee system eventually failed in several points causing torrents of salt water into adjacent neighborhoods which flooded up to 12 feet in some cases (See Figure 2-4). The death toll in New Orleans alone was over 1,600.

Four-fifths of New Orleans urban area was flooded. Over 107,000 housing units were damaged by flood, while more than 27,000 units sustained damages caused by wind. By September extensive areas of Orleans, St. Bernard, and Plaquemines parishes were without population or any economic activity (Campanella 2008). By New Year 140,000 out of

450,000 people returned to New Orleans. Services such as electricity, gas and potable water returned to unflooded areas though mid Autumn, however flooded areas had to wait several months, until 2006 for their services to return (Campanella 2006).

The idea of shrinking the city was proposed by some urban planners, however the idea was rejected by social activists who considered it as ethnic and class cleansing. Eventually the authorities gave a green light for the return to all areas in New Orleans, initiating a process that is still on course (Campanella 2008; Lam et al. 2009).

The current limited understanding of decision making after catastrophic events make this topic a relevant one for research. However, the nature of the events and their aftermath make them hard to study. The opportunities for systematic study and empirical observation are limited (Lam et al. 2009). Furthermore, until recently most of the research in the field of post-disaster recovery has been focused on households or whole communities and not in businesses (Pielke 1997; Webb, et al. 2002).

2.3 Business Decisions after a Disaster

Research on business recovery after natural disasters is still scarce (Dahlhamer and Tierney 1998, Waugh and Smith 2006). Some of this research uses aggregated units of analysis which obscure the conclusions of the research. Businesses are very heterogeneous, therefore it is necessary to conduct this kind of research at the micro-scale level (Zhang et al. 2009). Previous research at the microscopic level has identified certain business characteristics that could predict to a certain degree the recovery probabilities of a firm after a disaster. These are: the direct disaster impact at the business location, the business size, the business economic activity and lifelines interruption (Alesch et al. 2001; Dahlhamer and Tierney 1998; Webb, et al. 2002).



Figure 2-3 : Hurricane Katrina approaching New Orleans (source: CIMSS 2011)



Figure 2-4: Levee breaches in New Orleans after Hurricane Katrina (source: MSNBC 2011)

A significant predictor is the measure of the impact of the disaster at the business location. In the case of businesses after the Northridge earthquake, a division of the affected area according to the earthquake shaking intensity was identified as a significant recovery predictor. Earthquake shaking intensity is not only associated with the damage on the business itself but also with the damage the disaster has caused on neighboring businesses and residents (Dahlhamer and Tierney 1998). Businesses face more recovery challenges when their customer base is concentrated in areas deeply affected by the disaster. After a disaster customers might have to relocate due to residential damage or spend their income in repairs, restricting other expenditures on local businesses (Alesch et al. 2001).

Since 1993, the Disaster Research Center at the University of Delaware has conducted a number of post-disaster business surveys at different locations in the United States. The results of the surveys indicate that larger organizations are more prepared than smaller ones (Webb et al. 2000). After a disaster it is more difficult for smaller firms to recover. In the case of larger firms it is possible to have employees specifically dedicated to make pre-disaster preparations which is not the case for smaller firms. In the event of a disaster these preparations work and larger firms rebound faster (Yoshida and Deyle 2005; Kroll et al. 1991). Research conducted after the Northridge earthquake by Dahlhamer and Tierney (1998) concluded that larger organizations have better survival chances than smaller ones. A larger organization is more likely to have its infrastructure spread over a larger geographic area, in a disaster event, it is more likely that at least part of its infrastructure has not been damaged, allowing the organization to return to business faster. Smaller businesses tend to depend on local neighborhood customers, making them more susceptible to the negative effects of population displacements (Alesch et al. 2001; Dahlhamer and Tierney 1998).

Previous research suggests that the business economic activity has a strong influence on the type of preparedness firms have in case of disasters (Dahlhamer and D'Souza 1997). Studies conducted after the Northridge earthquake and in Memphis/Shelby county found that firms in the finance, insurance and real state (f.i.r.e) sectors were better prepared for disasters than businesses involved in other activities. This might be a result of a stricter regulation. The same research suggests that businesses in wholesale, retail and services tended to be less prepared (Webb et al. 2002).

2.4 Previous Research on Business Return Post Hurricane Katrina in New Orleans

In order to study the factors that affect business decision after a disaster, Lam et al. (2009) conducted a research on New Orleans business community after Hurricane Katrina. The purpose of the research was to identify the spatial and temporal dynamics of factors considered important by businesses for their decision making. The information obtained would enable researchers to model and predict the reopenings over time and space.

The research conducted by Lam et al. (2009) used the August 2005 Louisiana Department of Labor Micro File for Economic Development. This file contains 45 variables for each business in New Orleans. Among the most relevant are: NAICS code (North American Industry Classification System), physical address, telephone, contact person, longitude, latitude, zip code, census tract, census block, parish location, number of employees and aggregated wages. Using this information Lam et al. (2009) conducted three telephone surveys on December 2005, June 2006 and October 2007, with the purpose of studying the business return process on the short, intermediate and long term. The telephone surveys were conducted by the Louisiana State University Public Policy Research Laboratory. The

questionnaires were similar in order to compare them, although the third survey included more questions.

In order to analyze the telephone surveys, Lam et al. (2009) made the following assumptions: 1) Businesses that did not answer the telephone calls, after five or more attempts, and businesses with disconnected telephones were considered closed. 2) Businesses that did answer the telephone calls but refused to participate in the surveys were considered open. Under these assumptions less than 26% of the businesses reopened in the first four months after Katrina, by the time of the second survey the reopening rate increased to 39%, the rate increased to 66% by the time of the third survey (See Table 2-1).

Figure 2-5B depicts the business density in Orleans Parish before Hurricane Katrina. Figures 2-5C, 2-5D and 2-5E depict the business density based on the first, second and third survey respectively.

Lam et al. (2009) assumed the “disconnected” status of the telephones as indicating a short-term decision to not reopen, while the “no answer” was assumed as a sign of uncertainty, a “wait and see” approach to the crisis. Using these assumptions the telephone surveys indicate in the first four months, 15% of the businesses decided not to reopen anytime soon (disconnected), while 59% of the businesses took a “wait and see” approach indicating a high degree of uncertainty by this time. The second survey was performed 10 months after the disaster. It indicates that 28% of the businesses had their telephones disconnected, indicating a not to reopen decision, while businesses in the “no answer” category diminished to 32%. The results indicate a reduction in the number of firms with a “wait and see” attitude. The third telephone survey, taken two years after the disaster, indicates that only 5% of the businesses had their telephones disconnected, whereas 29% of the businesses were in the “no answer”

Table 2-1 : Survey statistics of the three surveys (source Lam et al. 2009)

	Dec-05	Jun-06	Oct-07
Total Sample	9132	9139	6155
Total attempted	8574	8808	5837
Others	359	439	2294
Revised (Attempted-Others)	8215	8369	3543
Completed survey	975(12%)	1418(17%)	1232(35%)
Assumed open	1173(14%)	1867(22%)	1101(31%)
Disconnected	1259(15%)	2376(28%)	170(5%)
No answer	4808(59%)	2708(32%)	1040(29%)

Note: The percentage values were computed using “Revised” (Total attempted minus Others) as denominator. (Note: “Others” include “no eligible respondent”, “incorrect business”, and “not a business.”; whereas “Assumed open” include “hard or soft refusal”, “busy”, “call back”, “fax”, “already taken survey”, “partially completed”, and “mail back”.)

category, indicating a reduction of uncertainty.

Lam et al. (2009) classified the businesses in seven groups based on their economic activity as stated in the firms’ NAICS code (North American Industry Classification System). The purpose of the classification was to create groups conformed by businesses of similar economic function in order to analyze the reopening rates for each group (see Tables 2-2 and 2-3).

The reopening rates for each business group are depicted in table 2-3. The results indicate that the highest reopening rate was 33% by the time of the first survey to Group 7 (professional, scientific and technical services). The lowest reopening rate in the first survey was 17%, which is Group 4 (educational, health care, social assistance, and public administration). By the time of the second survey Group 7 was still the one with the highest return rate (48%) while Group 4 was still the lowest return rate (32%). By the time of the third survey, the reopening rate differences between groups diminished. Two years after the

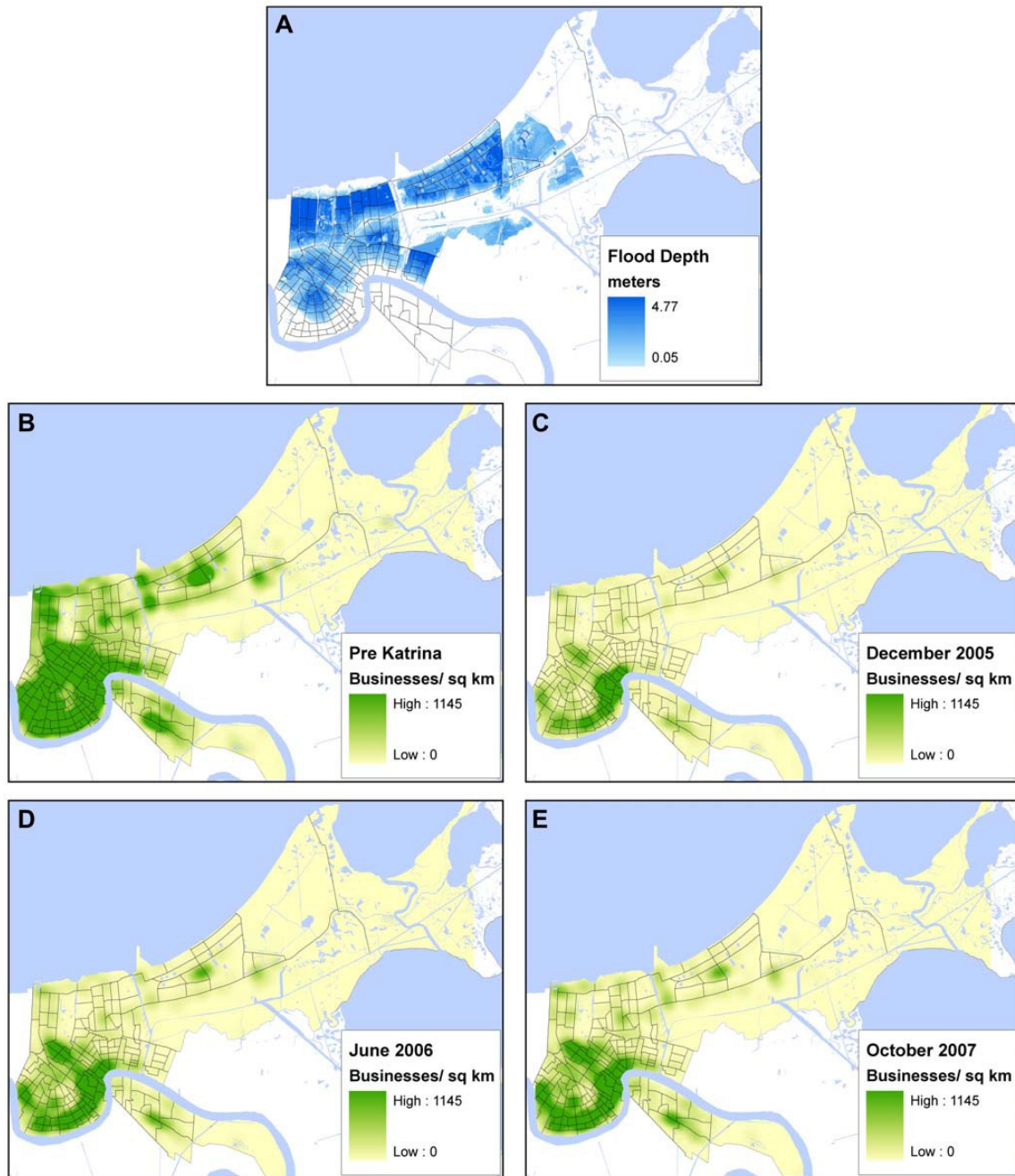


Figure 2-5: Flood depth and kernel density maps of businesses reopened by the three time periods (source: Lam et al. 2009).

Table 2-2: Business classification according to their NAICS code (source Lam et al. 2009)

Group	NAICS code	Description
1	11,21,22,23,31,32,33	Mining, utilities, construction, manufacturing, agriculture, forestry, fishing, and hunting
2	42,44,45	Wholesale and retail
3	51,52,53	Information, finance, insurance, and real estate
4	61,62,92	Educational, health care, social assistance, and public administration
5	71,72	Arts, entertainment, recreation, accommodation, and food services
6	48,49,55,56,81	Management of companies, waste management, transportation, warehousing, and other services
7	54	Professional, scientific, and technical services

Table 2-3 : Business opening ratio by type at the three time periods (source Lam et al. 2009)

Group	Dec-05			Jun-06			Oct-07		
	Total	Open	% Open	Total	Open	% Open	Total	Open	% Open
1	601	128	21	627	234	37	279	178	64
2	1533	368	24	1589	593	37	681	464	68
3	880	233	26	882	361	41	367	237	65
4	975	166	17	1007	322	32	414	252	61
5	1034	266	26	1072	397	37	481	305	63
6	1750	390	22	1746	601	34	681	429	63
7	1401	464	33	1423	676	48	598	397	66
Sum/Ave	8174	2015	25	8346	3184	38	3501	2262	65

Note: The “Total” column is the sum of “Completed survey”, “Disconnect”, “No answer”, and “Assumed open” in Table 1. The “% open” figures were derived by assuming “Disconnect” and “No answer” as businesses closed. Note that within the “Completed survey” category, a small portion of businesses remained closed even though they participated in the survey.

disaster the average reopening rate for all groups was close to 65%.

The businesses surveyed by Lam et al. (2009) were asked to rate a list of problems considered barriers to businesses reopenings, using a scale from 1 (less important) to 5 (most important). In the first survey, the most important barrier was the levee protection with an average score of 3.19, the second most important barrier was the lack of customers with an

average score of 2.89. In the second survey, the levee protection was still the main barrier with an average score of 3.20. The second and third most important barriers were utilities and communications with scores of 3.15 and 3.18. In the third telephone survey crime, was considered the most important barrier with an average score of 2.94 while the second most important barrier was levee protection with an average score of 2.87 (See Table 2-4).

The results indicate that immediately after the disaster, levees were considered the main concern for business owners. By the time of the second survey levees were still important, however other problems became almost equally relevant: utilities and communications. By the time of the third survey day to day concerns, such as crime, became more relevant.

The results of Lam et al. (2009) coincide with previous research. Firms related to scientific and technical services had a faster return rate than firms in other economic activities, while firms in the fields of health care, social assistance and public administration showed a lower return rate in the first survey. By the time of the third survey the differences between the reopening rates of the different business groups diminished. The average businesses reopening rate by October 2007 was 65% a figure very close to the jobs return rate (67%) estimated by the Brookings Institute during the second quarter of 2007 (The Brookings Institute 2008). The high scores given to utilities and communications coincide with previous research by Dahlhamer and Tierney (1998) regarding the importance of these elements for business recovery.

The multiple surveys allowed Lam et al. (2009) to study the attitude change among business owners through time after the disaster. It is interesting that 59% of the businesses by

Table 2-4: Average ranks of barriers and prospect for all businesses surveyed (source: Lam et al. 2009)

Problem	Dec-05	Jun-06	Oct-07
Damage	2.64	3.02	2.41
Insurance	2.66	2.64	2.58
Employee	2.73	2.97	2.69
Customer	2.89	2.76	2.68
Crime	-	2.41	2.94
Levee	3.19	3.20	2.87
Utilities	2.58	3.15	2.33
Communication	2.72	3.18	1.99
Environmental	2.23	2.42	1.82
Governmental	2.66	2.47	2.34
Financing	2.47	2.31	2.41
Prospect	2.54	2.31	2.30
Recovery progress	-	2.41	2.68

Note: Except for the “recovery progress” variable, scores range from 1 to 5 and the higher the score, the more important the problem.

the time of the first survey took a wait and see approach. These firms kept their telephone service, but waited to reopen until certain conditions were fulfilled. The surveys indicate that the main concerns by this point were the status of critical infrastructure: levees, utilities and communications. An emergency plan that considers a quick restoration of water, power, roads, public transportation, and telecommunications most likely would eliminate major concerns among business owners. Ten months after the disaster, infrastructure was still among the main concerns among business owners; indicating that these problems were not adequately solved by this time.

In summary the results of the surveys indicate that the business return is a dynamic process. The attitudes of business owners changed through time, indicating an uncertainty reduction and change about the main concerns for their business decisions.

In this study we are interested in predicting the reopening status of firms in New Orleans. A common approach in statistics is to try to predict the value of one variable based on the combinatory effect of a set of possible predictive variables. The reopening status of the businesses can only have two possible values, open/closed. The study of binary dependent variables requires special manipulation and statistical models such as logit and probit regression models that will be described in the next chapter.

Chapter 3 : Estimation of the Temporal Dynamics of Business Reopening Predictors

3.1 Introduction

It is often the case in scientific studies that researchers want to predict the value of a variable based on a combination of a set of predictors. A common approach is the use of linear models, which are a basic statistical tool. The purpose of linear models is to try to predict the value of a continuous variable of interest based on a linear combination of a set of explanatory variables.

A linear model can be represented as:

$$E(y) = \mu = X\beta + \varepsilon \quad (1)$$

In this model we are trying to predict the values of a dependent variable y whose expected values are represented by $E(y)$ or μ . A vector of size $n \times 1$ containing the set of explanatory variables is represented by X while β is the associated set of parameters for the explanatory variables and ε is the model disturbance.

The continuous nature of the dependent variable represents a shortcoming for the use of these models in social sciences. In many cases the dependent variables in social sciences are categorical, ordinal, or binary that represent attitudes, behaviors, characteristics or decisions, which cannot be predicted by common linear models due to their non-continuous nature. To handle these type of analysis statisticians have designed the generalized probabilistic models (Liao 1994).

3.2 Generalized Linear Models

A generalized linear model uses a function to link the linear combination of explanatory variables with the dependent one. The goal of the generalized probabilistic

models is to establish the probabilities of an event, how likely a decision is to be made, or the belonging to a certain group, etc, using a set of continuous or categorical predictors.

To create a generalized model we introduce the variable η , which is the result of the linear combination of $X\beta$. However the relation between η and μ is not necessarily linear and can be defined by a linking function. A linear model results when the linking function is identity.

$$\eta = X\beta + \varepsilon \quad (2)$$

$$\mu = g(\eta) \quad (3)$$

There are two common approaches to study dependent variables with a binary distribution, binary logistic (logit) and probit regression models. Binary logistic regression models started being used in the early 1980's when they become available in statistic software. The main reason for their use was the limitations of ordinary least squares regression when dealing with binary dependent variables (Peng and So 2002). The coefficients in this type of models are estimated using the maximum likelihood. Binary logistic models follow the form:

$$Prob(y = 1) = \frac{1}{(1 + e^{-\eta})} \quad (4)$$

A binary regression model is meaningful if the explanatory variables used in the model manage to explain the dependent variable better than the intercept only. The most common measure for this goal is the likelihood ratio statistic, which approximately follows the chi-square distribution. Only if the model has a significant χ^2 it is possible to interpret its parameters. However because of the existence of the link function the interpretation task is not straightforward. In general we can interpret the model coefficients by analyzing the sign of the parameters and their statistical significance. If a variable coefficient is significant, a

positive sign would indicate that an increase on its value would increase the likelihood of the dependent variable, assuming that all other independent variables are kept constant. The logistic coefficients are not standardized. It is not possible to make comparisons between variables unless they have the same measurement units. There is not a truly equivalent to the standardized betas that one can expect from an ordinary least squares regression.

Another way to interpret the results is through the analysis of the marginal effects. A marginal effect is the estimation of how much the probabilities of the event would change if there is a change in one unit value of an independent variable (Liao 1994, Aldrich and Nelson 1984). In a binary logistic regression model it is possible to estimate the marginal effects by transforming the variable coefficients into odds. An odd is the ratio between the probabilities of occurrence of an event against the probabilities of non occurrence of the event (See Formulae 5). Odds and probabilities are not the same, although it is possible to transform one into the other (see Formulae 6).

$$odds = \frac{prob_{(event)}}{1 - prob_{(event)}} \quad (5)$$

$$prob_{(event)} = \frac{odds}{1 + odds} \quad (6)$$

The results of a binary logistic regression model will indicate how much the odds of an event would change if there is an increase in one unit of any of the independent variables. An odds value of 1 would indicate no relationship. A value greater than 1 would indicate a positive relationship, while a value smaller than 1 would indicate an odds decreases with decreasing value of the independent variable (Spicer 2005, Norusis 2003, Peng and So 2002).

Probit regression models started being known in the field of econometrics in the early 60's (Liao 1994) . These models involve the modeling of the response variable with the normal cumulative distribution function. A probit model follows the form:

$$Prob(y = 1) = \Phi(\eta) \quad (7)$$

$$Prob(y = 1) = \Phi(X\beta + \varepsilon) \quad (8)$$

where Φ is the normal cumulative distribution function (CDF) of the standard normal distribution, η is a linear regression of the z score of the probabilities of the event. A z score is the measure in standard deviations from the mean. For example let's assume we are using a probit regression model to estimate the probabilities of an event, using a set of explanatory variables X . We have $\eta = X\beta + \varepsilon$, where η is the result of the linear combination from the explanatory variables. In order to transform η into probabilities we use it as if it was a z score in a standard normal probabilities curve. If the value of $\eta = -1.76$, then $\Phi(-1.76) = 0.0392$, which would indicate that 3.92% of the area of the probability curve is below this value, which is the probability of the event occurrence. In the case when $\eta = 0.58$, then $\Phi(0.58) = 0.719$, which would indicate that 71.9% of the area of the probability curve is below a z score of 0.58, then the probability of the event occurrence is 71.9% (See Figure 3-1).

When using probit models it is necessary to interpret the model coefficients with care because of the link function. A change in the value of the predictors does not have a direct effect on the event probabilities. The linear effect of the variation on the variables affects the probit (η), which needs to be processed by the function Φ in order to be translated as a change in the event probabilities.

Like in the case of the binary logistic regression models, variable coefficients that are found significant would affect the dependent variable according to their sign. Positive

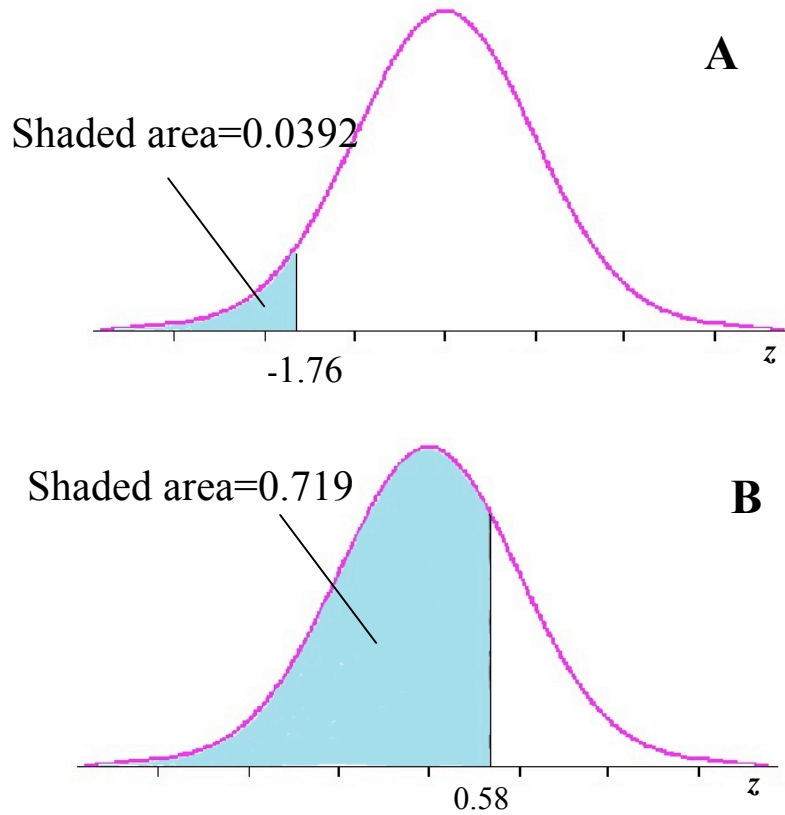


Figure 3-1: Probabilities estimation using a probit function.

coefficients would indicate that an increase in the variable value is associated with a higher likelihood of occurrence of the dependent variable, while a negative sign would indicate the opposite, assuming all the other independent variables are constant. It is also possible to analyze the probit model results using the marginal effects.

In most cases binary logistic and probit models will provide very similar results. The resulting probabilities from both models are almost identical. It has been reported that when probit and binary logistic models are applied to the same dataset, the difference between the coefficients is a constant factor believed to be approximately 1.8 (Sharif, Zaharim and Sopian 2009), where

$$\text{Probit coefficients} \times \text{constant} = \text{Binary logistic model coefficients}$$

although some researchers have proposed the factor to be 1.6 based on series of trial and error (Liao 1994).

3.3 Spatio-temporal Dependence

Classic statistical analysis considers each member of the sample as independent, not being affected by other members of the sample. In the case of businesses decisions, this seems unrealistic. The reopening decision of one business would certainly affect the decisions of its neighbors. A model that takes in account the influence of neighboring elements is introduced by LeSage and Pace (2009) and called the SAR model (Spatial Autoregressive Regression model).

$$y = \rho W y + X\beta + \varepsilon \quad (9)$$

where y is the status of a firm, W is a spatial weight matrix that represents the indirect influence of neighbors on y .

In LeSage et al. (2011) the authors studied the business reopening decisions after Hurricane Katrina in New Orleans. They introduced a modified version of Formula 9:

$$y^* = \rho W y^* + X\beta + \varepsilon \quad (10)$$

In this version, y^* represents the latent unobservable utility/profit associated with the closed/open status of the firms. LeSage et al. (2011) defined W as a spatial matrix of size $n \times n$ where n is the number of firms. W contains either $1/m$ or 0 , where m is a fixed number of selected closest neighbors for each business.

To estimate the reopening probabilities LeSage et al. (2011) used probit regression models. They used a sample comprised of 673 stores distributed in three major New Orleans commercial corridors, St. Claude Avenue, from Poland Avenue to Faubourg Treme, the entire Magazine Street, and South and North Carrollton avenues. The purpose of the study was to

explore variables that influenced the firm's reopening decision. The explanatory variables in the analysis were: flood depth, (logged) median income for the census block group in which the firm is located, the business size, the socioeconomic class of the clientele, and store ownership. LeSage et al. (2011) found that at the first three month horizon the significant reopening variables were flood depth, logged (median income), store size, proprietorship, and the spatial spillovers. The flood depth had a negative impact. Stores located in flooded areas had lower reopening chances. The logged median income of the resident population had a positive effect, stores located in wealthier areas had better reopening chances. Small stores had lower reopening probabilities than medium size ones. Sole proprietorships stores had higher reopening probabilities compared to regionally owned chains. In this time horizon the spatial spillovers had a positive effect. At the six month horizon the significant predictors were flood depth, customer status, proprietorship and the spatial spillovers. Flood depth still had a negative effect like in the previous time horizon. Stores with low status customers had lower reopening chances than stores with medium status clientele. As in the previous time horizon the sole proprietorship had better reopening chances than regionally owned chains. At the twelve month horizon the significant variables were flood depth, logged median income, the clientele status, and the spatial spillovers. Flood depth was still a significant variable negatively affecting the reopening chances, like in previous time horizons. The logged median income behaved in a manner similar to that of previous time horizons, having a positive effect. Stores with low status customers had lower reopening chances than stores with medium status customers. In all the time horizons the spatial spillovers had a positive reopening effect.

In this dissertation we estimate the reopening probabilities every three months. The purpose and methods of this research are similar to those of LeSage et al. (2011), but they differ in two aspects: the sample used for the analysis and the estimation of the spatial spillovers. While the samples used by LeSage et al. (2011) are located along three main avenues in Orleans Parish, the samples used in this research comprise 1358 businesses selected randomly from the total set of businesses included in the August 2005 Louisiana Department of Labor Micro File for Economic Development in the Greater New Orleans Area. The spatial distribution of the sampled businesses is not even, it follows the spatial trends of the actual business distribution with some areas having high business concentration and other areas having scarce businesses presence (See Figure 3-2).

Using the notation specified by Liao (1994), the method employed by LeSage and Pace (2009) would be represented as:

$$y^* = \Phi(\eta) \quad (11)$$

where:

$$\eta = \rho W y^* + X\beta + \varepsilon \quad (12)$$

Using the same notation, the probit model employed in this study would be represented as:

$$y_t^* = \Phi(\eta_t) \quad (13)$$

$$\eta_t = \rho_t W y_{(t-1)} + X\beta_t + \varepsilon \quad (14)$$

while a binary logistic model would have the same η_t and follow the form described in Equation 4.

Although both studies use probit regression models and some common variables, this study handles the spatial spillovers differently in several ways: First we consider that

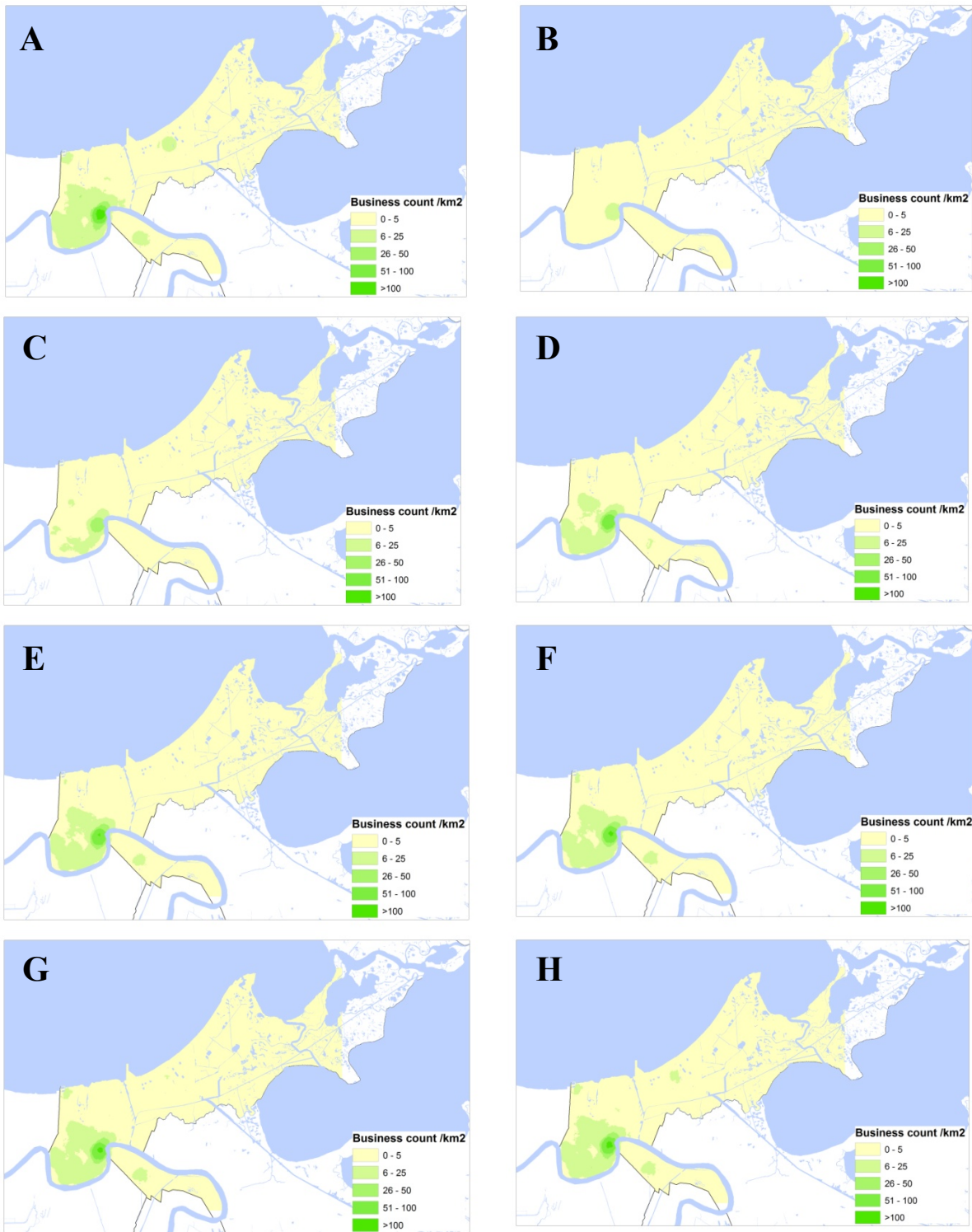


Figure 3-2 : Business density for firms included in Survey 3 A) Before Katrina, B) Right after Katrina, C) By September 30th 2005, D) By December 30th 2005, E) By March 31st 2006, F) By June 30th 2006, G) By September 30th 2006, H) By October 2007.

the spatial spillovers have a temporal dimension. LeSage et al. (2011) considers that a change in the value of the variables of one firm would instantly affect its neighbors, whereas we consider that a change of this nature would affect the firm neighbors in a later point of time. $Wy_{(t-1)}$ represents the status of the firm's neighbors in time $(t-1)$, which is used as the input for the estimation of the reopening probabilities for businesses in time t .

A second difference is that in our study, the spatial spillovers are based on the actual reopening status of neighboring firms (Wy) and not on the return probabilities (Wy^*). Our spatial spillovers do not consider the indirect effect that a change in a predictive variable would have on neighbor firms. A change on the spatial spillovers is only given by a change in the status of a neighbor firm from closed to open.

A third difference is in the approach used to calculate the spatial spillovers. In LeSage et al. (2011) the authors use a fixed number of the closest neighbors; assigning all of them the same weight. In this study we consider all the firms located within a given radius of influence as neighbors. Because of the heterogeneous spatial distribution, the number of neighbors changes depending on the location. The amount of influence assigned to each neighbor is not the same, it is based on an inverse distance weight function, the closer the neighbor the stronger the influence. The formula used to estimate the neighbor influence is:

$$W_y = \frac{\sum_{k=1}^N C_k v_k}{\sum_{k=1}^N C_k} \quad (15)$$

and

$$C_k = \frac{1}{d(y, y_k)^p} \quad (16)$$

where for each observation y we have N neighbors, v_k is the status of neighbor k (either 0 or 1), while C_k is the inverse distance function value between observation y and its neighbor y_k ,

power by p . Higher values of p would increase the importance of the closest neighbors (Longley et al. 2005, Dubin 2009).

A fourth difference is the removal of opened firms from the analysis. In this study the dependent variable is the status of the businesses every three months. We assumed that once a business decides to reopen it remains reopened for the rest of the study period. Estimating the reopening probabilities in time " $t+n$ " of a business that has already reopened in time " t " would only dilute the effects of the variables in time " $t+n$ ". In this study we remove from our sample businesses that have already reopened in a previous time period. However neighbors that have reopened in previous time points are still considered when estimating the spatial spillovers.

The approach used to define the spatial spillovers can be better understood with an example: Let's assume we have six businesses, as represented in Figure 3-3. The distances between the businesses are indicated next to the line that links them, for instance the distance between business 2 and business 1 equals 9, while the distance between businesses 1 and 5 equals 6. In this example we are going to estimate the influence of neighboring firms on business 1.

In the previous time step businesses 3 and 5 were open, while businesses 2,4 and 6 were still closed. Table 3-1 contains the required operations to estimate the neighbor influence using equations 15 and 16.

In this study we estimate the businesses reopening probabilities for several periods of time. Businesses that had reopened in a given point of time were removed from the sample in subsequent time periods. For example, firms that reopened between October and December 2005 were excluded from the probit analysis for the period of January to March 2006.

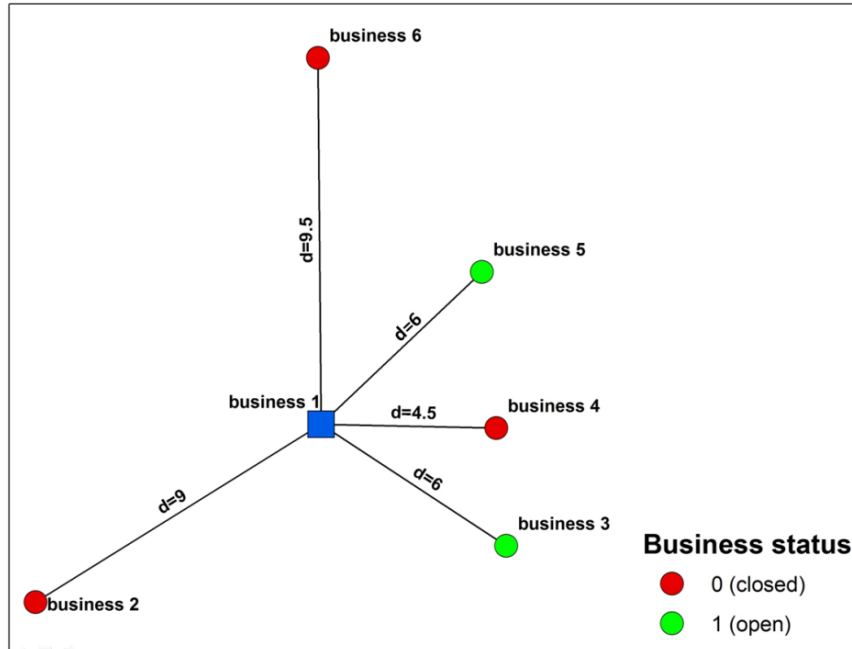


Figure 3-3: Estimation of neighbor influence on business 1

Table 3-1: Neighbor influence estimation for business 1 using $p=2$

	v_k	$d(y, y_k)$	C_k	$C_k v_k$
business 2	0	9	0.01	0.00
business 3	1	6	0.03	0.03
business 4	0	4.5	0.05	0.00
business 5	1	6	0.03	0.03
business 6	0	9.5	0.01	0.00
			0.13	0.06

Table 3-2 : Reopening counts by time period

Time Period	Open	Closed	Total
August 30th - September 30 th 2005	268	917	1185
October 1 st 2005 – December 31 st 2005	452	465	917
January 1 st 2006 - March 31 st 2006	129	336	465
April 1 st 2006 - June 30 th 2006	50	286	336
July 1 st 2006 – September 30 th 2006	55	231	286
October 1 st 2006 – October 2007	63	168	231

However, firms that reopened before January 2006 were still considered for the neighborhood influence estimation.

3.4 Analysis of the Third Telephone Survey

3.4.1 Data

The data for this research were obtained through three telephone surveys carried out in December 2005, June 2006, and October 2007, as described in Lam et al. (2009) and in Chapter 2.

Dependent Variable: For this study the dependent variable is the status of a set of businesses in Orleans Parish after Hurricane Katrina, open or closed. The information source is the third telephone survey carried out in October 2007. This survey asked the respondents:

Q1a. Did your business close following the 2005 Hurricanes? Y/N

Q1b. If yes, how long was it closed? _____ months.

Using the answers to this question it is possible to reconstruct the business return process at a monthly interval. In this research we assumed businesses that had their telephones disconnected were still closed. The number of open businesses is not cumulative. For instance, between October 1st and December 31st 2005 452 businesses reopened out of 917. In the period between January 1st and March 31st 2006, 129 firms out of the remaining 465 decided to reopen. The analysis for the first quarter of 2006 does not consider businesses that reopened before January 1st 2006. Including in the sample businesses that had already reopened in previous time periods may dilute the effects of the predictors and bias the results.

Independent variables: The independent variables considered for the analysis can be divided in three groups:

1) Business characteristics: (a) Business size: We used as a proxy the natural logarithm of the total wages of each business. (b) Disaster damage: To estimate the disaster

damage we estimated the flood depth at the business location as a proxy. The flood depth in meters data was obtained from the Katrina & Rita clearing house cooperative, managed by Louisiana State University (L.S.U. 2005).

2) Social vulnerability: The social vulnerability of the residents of an area has been previously studied. Our selection of variables was derived from the study by Cutter et al. (2004), which includes the following: population density (population by square kilometer), % non white population, % population with age under 18, % population with age over 65, median household income, % renter occupied houses, and % female population. These variables were obtained at the census block group level from the Census 2000. Our study area includes 485 census block groups.

3) Spatial spillovers: Classic statistical analysis considers each member of the sample as independent, not being affected by other members of the sample. In the case of businesses decisions, this seems unrealistic. The reopening decision of one business would certainly affect the decisions of its neighbors. In this research the spatial spillovers were estimated by assigning a binary value to the status of the firms to indicate open/close (1 or 0). Then we calculated the inverse distance weighting value of the neighbors reopening status within a distance threshold of the business of interest as a variable to indicate the spatial spillover effect. We considered only firms located within 1000 meters of the business of interest.

3.4.2 Methods

In this research we estimate the reopening probabilities for the businesses in Orleans Parish using both probit and binary logistic regression methods. This study compares the results of both approaches and identifies the significant predictors.

To perform any spatial analysis it is necessary to represent the businesses as spatial features. Our dataset includes the address of each business. To transform the address into coordinates we used the Geocode tool of ArcGIS. The final result is a set of points, projected in UTM coordinates (zone 15). Each point represents a business.

Data for the socio-economic variables of the residents in the vicinity of the businesses were obtained from the Census Bureau for Orleans Parish in Census 2000 at the census block group level. To associate the business locations with the census block groups we performed a spatial join in ArcGIS (point in polygon operation). The flood depth information obtained for Orleans Parish was associated with the businesses location using a point in raster operation in ArcGIS.

The spatial spillovers were calculated using a Java program specifically designed for that task by the author. The businesses were assigned a value based on its status, close=0, open=1. The Java program estimated the inverse distance weight value for each business based on the status value of its neighbors within 1000 meters.

Using the previously described predictors we estimated the reopening probabilities for the businesses included in the third survey, for the following periods: 1) August 30th - September 30th 2005, 2) October 1st 2005 – December 31st 2005, 3) January 1st 2006 - March 31st 2006, 4) April 1st 2006 - June 30th 2006, 5) July 1st 2006 – September 30th 2006, 6) October 1st 2006 – October 2007.

We used two statistical methods to calculate the reopening probabilities of the businesses: probit analysis, using the software SAS, and binary logistic regression, using the software SPSS. Both methods although different, provide similar results as it will be further discussed in the next sections.

3.4.3 Results

Table 3-3 shows the descriptive statistics of the variables used in this analysis. Previous research suggests that probit and binary logistic regression models have in most cases similar results (Liao 1994). The results of this research seem to confirm this conclusion. We created six models for each type of regression. The models predict the reopening probabilities of businesses that reopened in six time points as described in Table 3.2. The significant predictors identified in each type of regression models are the same. The significance of the variables is not the same through time. The coefficients for the binary logistic models and their marginal effects can be seen in Table 3-4. In the case of the probit analysis, the model coefficients and their marginal effects are depicted in Table 3-5.

3.4.3.1 Binary Logistic Regression Model Results

The results of the binary logistic regression models for each time period are depicted in Table 3-4. A R^2 measure indicates how much of the variability of the predicted variable is explained by the model. Binary logistic regression models do not provide a direct equivalent to the R^2 that results from an Ordinary Least Squares regression. Alternatives are pseudo R^2 like Nagelkerke's. However these measures usually have very low values and should not be interpreted in the same way as the R^2 resulting from a OLS regression.

An alternative measure for the model fitness is the percentage of correctly predicted values. The model constructed for the period between August 30th and September 30th 2005 can correctly predict the values of 77.4% of the sample. For the second study period between October 1st and December 31st 2005 the model can correctly predict the values of 62.9% of the sample. For the period between January 1st and March 31st the model can correctly predict 72.3% of the status of the firms. The model corresponding to the period between April 1st and

Table 3-3 : Descriptive statistics of the variables for each analysis

		Ln (wages)	Flood depth	Pop dens	% non white	%pop <18	%pop >65	Ln (mghi)	% renter	% female pop	ni
Aug-05 Sep-05 (n=1187)	Min	0.00	0.00	0.13	0.65	0.00	0.00	8.32	0.00	27.85	0.00
	Max	17.24	2.79	91.72	100.00	55.06	33.31	12.21	98.63	78.61	1.00
	Std dev	2.63	0.59	17.92	32.39	12.90	8.08	0.60	18.38	8.57	0.15
	Mean	9.76	0.40	25.47	46.78	16.67	13.23	10.14	53.84	48.01	0.14
Oct-05 Dec-05 (n=917)	Min	0.00	0.00	0.13	0.65	0.00	0.00	8.32	0.00	27.85	0.00
	Max	17.24	2.79	91.72	100.00	55.06	33.31	12.21	98.63	78.61	1.00
	Std dev	2.74	0.60	17.83	32.77	12.93	8.02	0.60	18.62	8.57	0.21
	Mean	9.64	0.43	25.61	47.32	17.14	13.26	10.14	53.33	48.15	0.32
Jan-06 Mar-06 (n=465)	Min	0.00	0.00	0.13	0.89	0.00	0.00	8.65	1.45	27.85	0.00
	Max	17.24	2.73	84.96	100.00	55.06	33.31	11.89	89.72	78.61	1.00
	Std dev	2.79	0.64	18.21	33.58	12.89	7.79	0.59	19.12	8.41	0.25
	Mean	9.45	0.57	25.62	54.49	19.61	12.80	10.10	53.48	49.06	0.61
Apr-06 Jun-06 (n=336)	Min	0.00	0.00	0.13	0.89	0.00	0.00	8.81	1.45	27.85	0.00
	Max	17.24	2.73	77.23	100.00	55.06	32.33	11.89	89.72	78.61	1.00
	Std dev	2.98	0.67	18.16	34.36	12.70	7.31	0.59	19.35	8.07	0.24
	Mean	9.38	0.63	25.81	55.78	20.83	12.40	10.12	52.41	49.39	0.70
Jul-06 Sep-06 (n=286)	Min	0.00	0.00	0.13	0.89	0.00	0.00	8.81	2.67	27.85	0.00
	Max	17.24	2.73	77.23	100.00	55.06	32.33	11.89	89.72	78.61	1.00
	Std dev	3.12	0.68	18.23	34.46	12.78	7.20	0.59	19.31	7.91	0.24
	Mean	9.28	0.65	26.30	55.41	20.84	12.40	10.14	52.50	49.65	0.73
Oct-06 Oct-07 (n=231)	Min	0.00	0.00	0.13	0.89	0.00	0.00	8.81	2.67	27.85	0.00
	Max	17.24	2.72	77.23	100.00	55.06	32.33	11.89	89.72	78.61	1.00
	Std dev	3.28	0.67	18.84	34.64	13.08	7.52	0.58	19.37	8.02	0.22
	Mean	9.21	0.61	26.08	54.13	20.26	12.28	10.12	53.72	48.76	0.80

June 30th can predict 85% of the status of the businesses. The model for the period between July 1st and September 30th can correctly predict the status of 80.1% of the firms in the sample. The model representing the period between October 1st 2006 and October 2007 can predict 73.2% of the sample status.

The results of the binary logistic regression provide us with the coefficients required to estimate the reopening probabilities as well as a measure of the marginal effect of a change of

Table 3-4 : Binary logistic regression model results (significant coefficient * sig<0.05)

Period:	Aug 30th	Sep 30th	Oct 1st	Dec 31st	Jan 1st	Mar 31st	Apr 1st	Jun 30th	Jul 1st	Sep 30th	Oct 1st	Oct 2007
Nagelkerke R^2	0.04		0.18		0.08		0.07		0.16		0.05	
% Correctly Classified	77.4		62.9		72.3		85.1		80.1		73.2	
	B	Exp(b)	B	Exp(b)	B	Exp(b)	B	Exp(b)	B	Exp(b)	B	Exp(b)
ln(wages)	* 0.08	1.09	* 0.09	1.09	0.03	1.03	0.08	1.08	0.08	1.08	0.03	1.03
flood depth	* -0.07	0.93	* -0.17	0.84	* -0.11	0.90	-0.08	0.92	-0.01	0.99	0.00	1.00
pop dens	0.00	1.00	0.00	1.00	0.00	1.00	-0.01	0.99	0.00	1.00	0.01	1.01
% non white	0.00	1.00	* -0.01	0.99	0.01	1.01	0.00	1.00	0.00	1.00	0.01	1.01
% pop <18	-0.01	0.99	0.00	1.00	* -0.04	0.96	0.02	1.02	-0.03	0.97	0.00	1.00
% pop >65	0.00	1.00	0.01	1.01	0.02	1.02	0.00	1.00	-0.01	0.99	-0.01	0.99
ln(mhhi)	0.27	1.30	-0.04	0.96	0.18	1.19	-0.61	0.54	0.16	1.17	0.35	1.42
% renters	0.01	1.01	0.00	1.00	0.01	1.01	-0.02	0.98	-0.02	0.98	0.00	1.00
% fem pop	0.01	1.01	0.01	1.01	0.02	1.02	-0.04	0.96	* 0.11	1.12	-0.05	0.95
ni	* 0.72	2.05	0.00	1.00	0.33	1.39	-0.26	0.77	* -1.46	0.23	-0.75	0.47
Constant	-5.61	0.00	-0.12	0.89	-4.26	0.01	7.02	1121.08	-6.86	0.00	-2.33	0.10

one unit in any of the variables. The marginal effect is denoted by $Exp(B)$, which is the odds ratio.

The odds is the ratio between the reopening probabilities and the probabilities of not reopening (see Equation 5). If we add one unit in any of the variables while the others remain constant, we will have new odds and new reopening probabilities where:

$$Exp(B) = \frac{\text{new odds}}{\text{previous odds}} \quad (17)$$

The results of the binary logistic regression provide us with an $Exp(B)$ estimation for each of the variables. This would indicate how much the odds value would be affected when there is a change of one unit in the variable. Odds are different from probabilities, although it is possible to calculate the reopening probabilities based on the odds using Equation 6. If the variable coefficient is positive, the odds ratio is greater than 1, an increase of value in this

variable would positively affect the probabilities. If the coefficient of the variable is negative, its odds ratio would be lower than 1, and an increase of it would decrease the probabilities.

The results of the binary logistic regression models indicate that business size is a significant positive predictor in the first and second periods (see Table 3-4). Larger businesses had better reopening probabilities than smaller ones in the first months after the disaster, until December 2005. After this point the size variable ceased being significant. In the first time period, between August 29th and September 31st 2005, an increase in 1 unit of the natural log of the wages would represent an odds ratio of 1.09.

In the second time period, between October 1st and December 31st of 2005 an increase in one unit in this variable would represent an odds ratio of 1.09, same as the first time period. To estimate the business size we used $\ln(\text{total wages})$ as proxy, therefore it is necessary to be cautious when translating a change of one unit of this variable into dollars.

There is no straight way to translate the odds ratio into a change of probabilities. We can calculate the odds for any reopening probability. By adding one unit in a variable and keeping constant of the rest, we are modifying the probabilities and therefore the odds. The odds ratio is the relationship between the new and the previous calculated odds. The effect of an increase in $\ln(\text{wages})$ on the odds ratio and the reopening probabilities can be better explained with an example. Consider a business Y with a reopening probability of 0.45. Using Equation 5 we can calculate its reopening odds: $[0.45/(1-0.45)]=0.818$. The results of the logistic regression indicate that for the first time period an increase in one unit of $\ln(\text{wages})$ would result in an odds ratio of 1.09 (see Table 3.4). We can calculate the resulting new odds using Equation 17 ($\text{new odds}=\text{odds ratio} \times \text{previous odds}$). In this case the new odds is 0.892. Using Equation 6 we can calculate the new reopening probabilities:

$0.47 = [0.892 / (1 + 0.892)]$. An increase of one unit in the value of $\ln(\text{wages})$ would affect the reopening probability of Y, increasing it by 0.021 (0.471-0.45).

The results of the analysis suggest that flood depth is a significant variable in the first three time periods, negatively affecting the reopening probabilities. The results indicate that businesses located in flooded areas had lower reopening probabilities than their peers established in not-flooded locations. In the first period an increase in one meter of flood depth would represent an odds ratio of 0.93. In the second time period, this variable had a negative effect on the reopening probabilities. However its importance changed through time. An increase in one meter of flood depth would represent an odds ratio of 0.84. In the third time period, between January 1st and March 31st of 2006, an increase in one meter of flood depth would represent an odds ratio of 0.9.

For example, in the case of the previously introduced business Y, it had a reopening probability of 0.45. Its reopening odds was 0.818. In the case of the first period, an increase in one meter of flood depth would represent an odds ratio of 0.93. Using Equation 17, the new reopening odds was 0.761. The calculated reopening probabilities for this odds value was 0.432. An increase in one meter of flood depth would result in a decrease in the reopening probability of 0.0178 (-1.79%).

The percentage of non white population is a significant negative predictor only in the second time period. An increase in one percent non white population would represent an odds ratio of 0.99.

Let's return to the previously introduced business Y. Let's say it has a reopening probability of 0.45 in the second time period, which would mean it has a reopening odds of 0.818. An increase in one unit of the % non white population would represent an odds ratio

of 0.99. Using Equation 17, we calculate the new reopening odds associated with the new values for the predictors. The new reopening odds is 0.810, then the reopening probability is 0.448. An increase in 1% of non white population in the vicinity of business Y in the second period would result in the decrease in its reopening probability of 0.0025 (-0.25%).

The percentage of population with age under 18 was a significant predictor only in the third time period (January – March 2006). The variable negatively affects the reopening probabilities. An increase in 1% of this variable would mean an odds ratio of 0.96.

Using the example of business Y with a reopening probability of 0.45, we have that in the third period, an increase in 1% of the percentage of population with age under 18 would result in a new odds ratio of 0.96. The new reopening odds value is 0.785, hence the new reopening probability for business would equal to 0.44. An increase in 1% of the population with age under 18 would mean a decrease of the reopening probability of 0.01 (-1.01%).

The percentage of female population was only significant in the fifth time period. An increase in one unit of this variable would mean an odds ratio of 1.12. Consider an example of business Y with a reopening probability of 0.45. The odds of business Y is 0.818. The odds ratio is 1.12, which indicates that an increase in 1% of the female population of the census block where the business is located would result in a new reopening odds of 0.916 (Using Equation 17: $0.916=0.818 \times 1.12$), which would indicate a new reopening probability of 0.478. An increase in 1% of the female population positively affects the reopening probability by increasing it by 0.028 or 2.82%.

The neighbor influence was significant only in the first and fifth time periods. In the first time period it affected the reopening probabilities in a negative form, while in the later time period the neighbor variable worked in the opposite way. In the first time period an

increase of 1 in the neighbor influence would represent an odds ratio of 2.05, where in the fifth time period the same increase would represent an odds ratio of 0.23.

The variable neighbor influence has a range between 0 and 1. The value 0 would indicate all the neighboring businesses are closed. Value 1 would mean the opposite. Using the example of business Y, it has a reopening probability of 0.45 and an odds value of 0.818. In the first period the odds ratio is 2.05. An increase in one unit of the neighbor influence meaning all the neighbors are open, would result in a new odds of 1.677 (using Equation 17), and a new reopening probability of 0.626 (according to Equation 6). An increase in one unit of the neighbor influence would represent an increase of 0.1765 in the reopening probability or 17.65%.

3.4.3.2 Probit Regression Model Results

The results of the probit and binary regression models are the same in terms of identifying the same significant predictors for the same time periods (See Table 3-5). Business size was identified as a significant positive predictor for the first two time periods. Larger businesses had better reopening probabilities than smaller ones. In the first time period an increase in one unit of the natural logarithm of the wages would increase the reopening probabilities by 0.014 (1.411%), while in the second time period it would increase the reopening probabilities by 0.0192 (1.92%).

Flood depth was a significant negative predictor in the three first time periods. Businesses located in flooded areas had lower reopening chances than their unflooded peers. In the first time period an increase of one meter in flood depth would represent a decrease by 0.0118 (-1.18%) in the reopening probabilities. In the second period it would

Table 3-5: Probit regression models, coefficients (B) and marginal effects (Meff) (* Sig<0.05)

	Aug	Sep	Oct	Dec	Jan	Mar	Apr	Jun	Jul	Sep	Oct	Oct
	30th	30th	1st	31st	1st	31st	1st	30th	1st	30th	1st	2006
	B	Meff	B	Meff	B	Meff	B	Meff	B	Meff	B	Meff
ln(wages)	*0.05	0.01	*0.05	0.02	0.02	0.01	0.05	0.01	0.04	0.01	0.02	0.01
flood depth	*-0.04	-0.01	*-0.10	-0.04	*-0.06	-0.02	-0.05	-0.01	0.00	0.00	0.00	0.00
pop dens	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.01	0.00
% non white	0.00	0.00	*-0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
% pop <18	-0.01	0.00	0.00	0.00	*-0.02	-0.01	0.01	0.00	-0.02	0.00	0.00	0.00
% pop >65	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	-0.01	0.00	-0.01	0.00
ln(mhhi)	0.15	0.04	-0.02	-0.01	0.12	0.04	-0.37	-0.08	0.10	0.02	0.19	0.06
% renters	0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	-0.01	0.00	0.00	0.00
% fem pop	0.00	0.00	0.00	0.00	0.01	0.00	-0.02	-0.01	*0.06	0.01	-0.03	-0.01
ni	*0.43	0.13	0.00	0.00	0.19	0.06	-0.17	-0.04	*-0.87	-0.20	-0.45	-0.14
Constant	-3.26		-0.13		-2.71		4.11		-4.08		-1.29	

represent a decrease of 0.038 (-3.8%). Finally in the third time period the same increase in flood depth would affect negatively the reopening probabilities by 0.021 (-2.1%).

The percentage of white population was identified as a significant predictor only in the second time period. An increase in one percent of the non white population would reduce the reopening probabilities by 0.003 (-0.3%).

The percentage of population under 18 was identified as a negative significant predictor only for the third time period. An increase in one percent of this segment of the population would reduce the reopening probabilities by 0.008 (-0.8%).

The percentage of female population was identified as a significant positive predictor only in the fifth time period. An increase in one unit of the percentage of female population would increase the reopening probabilities by 0.015 (1.5%).

The neighbor influence was identified as a significant predictor only in the first and the fifth time periods. In the first time period an increase in 1 unit of this variable would

increase the reopening probabilities by 0.13 (13%). In the fifth time period the same increase would represent a decrease in the reopening probabilities by 0.20 (-20.0%).

3.5 Discussion

Our results indicate that smaller businesses had lower reopening chances than larger ones. However this was only true until December 2005. After this point this variable stops being statistically significant. It is necessary to mention that we are using a linearized variable $\ln(\text{wages})$, therefore an increase of one unit of the linearized variable would mean differently in real world dollars, depending on the value of the original variable. For instance, consider a business Y, which has a total wage equivalent to \$10,000, then $\ln(\text{wages})=9.21$. An increase in one unit of the variable $\ln(\text{wages})$ would mean a new $\ln(\text{wages})$ equal 10.21, which would be equivalent to \$27,183. The difference is \$17,183. However, an increase in one unit of $\ln(\text{wages})$ from 10.21 to 11.21 would represent an increase of \$46,708. Policies designed to help small businesses should consider the temporal dimension. The help is most needed immediately after the disaster, especially in the case of small firms.

As it would be expected, businesses located in flooded areas had lower probabilities to reopen, although the temporal window for flood depth is bigger. By the end of March 2006 the flood depth stopped being a significant reopening predictor. Policy designed to help flooded firms should focus on this time frame.

A major disaster causes massive population displacements, and in many cases the population remains mobile for several months. Detailed and accurate return population estimates become difficult to obtain in this context (Plyer et al. 2010) . After a disaster, businesses are affected by the reduced purchasing power of their clients or by their displacement (Alesch et al. 2001). In the case of New Orleans, population displacement is hard to study at the micro level due to the lack of data. The 2006 American Community

Survey (ACS) conducted by the U.S. Census Bureau provides statistically reliable data comparable with the Census 2000, at the county level. The results of this survey were analyzed by Frey et al. (2007). The authors concluded that the return of young people and families with children to New Orleans was lower compared to other segments of the population, while the returning population was disproportionately composed of white, childless and better educated population.

In our analysis we used as independent variables socio-economic indicators describing the resident population at the census block group level, using Census 2000 data. Our results indicate that the percentage of non white population and the percentage of population with age under 18 were negative predictors by December 2005 and March 2006. Frey et al. (2007) identified the population with low return as having similar characteristics as our negative reopening predictors. Census block groups that had more non white population and more population with age under 18 were the ones with lower population return, hence impacting negatively the businesses reopening, which could be due to a reduced clientele for businesses that rely on local customers. Previous research in other geographic areas agrees with this finding (Alesch et al. 2001; Dahlhamer and Tierney 1998). Policies designed to improve the business return should also take into consideration the resident return. Businesses are connected with the residents as they are both customers and labors.

3.6 Conclusions

The New Orleans business community is heterogeneous, composed of elements with different characteristics that were affected by the flood caused by Hurricane Katrina in an uneven manner. The reopening process is dynamic and not linear. The reopening rate is not constant through time. The variables with significant influence in the reopening decision are not the same at different points in time.

Both the binary logistic regression and the probit regression analyses identified the same significant reopening predictors, which were: $\ln(\text{wages})$, flood depth, % non white population, % population with age under 18, % of female population, and the neighbor influence. The neighbor influence was identified as a significant predictor, which indicates that the status of neighboring firms affects the reopening decision. There is a level of interaction between the businesses; they influence each other.

Both regression procedures indicate that the significance of the reopening predictors fluctuates through time. Most of the significant variables appeared in the first three periods until March 31st 2006.

A comparison of binary logistic regression and probit analysis indicates that both results are very similar. Although it is much easier to estimate the reopening probabilities on the fly using a logistic function than a probit one.

The business community in New Orleans is heterogeneous in multiple aspects. Important characteristics that influence the business reopening decisions could have a wide range of values for any specific firm. The recovery process post Hurricane Katrina is a dynamic phenomenon. The relevance of certain factors is not the same through time. Science has developed several simulation approaches to mimic dynamic processes. However most of these approaches consider the average values of the system only, the detailed differences among components are not modeled. In the next chapter we will describe an agent-based modeling approach that considers the various characteristics of the business community at a detailed level.

Chapter 4 : Agent-based Models and Simulation in Geo-information Science

4.1 Models and Simulation

A model is a simplified representation of the real world, where significant, relevant elements of it are depicted while others are ignored. Simulation is one of the many purposes for developing models. A simulation model is an attempt to represent a dynamic system of the real world. Possible uses of a simulation can be to: 1) perform experiments to obtain better understanding of how the system works, 2) reproduce the dynamics of the system and predict future outcomes, 3) develop capabilities to substitute human elements (Gilbert and Troitzsch 2005).

The use of computer simulation models is relatively new in science. The first models started being used in the 1960s, they became common in the 1990s, mainly because of greater computer processing availability. Computer simulation models are constructed as computer programs with inputs and outputs. Due to their nature, they require the modeler to be precise in defining the model parameters. Computer models in social sciences allow their creators to perform experiments. A common approach to experimentation is to compare the outcomes resulting from applying certain treatment with the outcomes of not applying it. Applying this approach in social sciences is difficult. Computer models when well validated allow social scientists to perform experiments without affecting humans (Gilbert 2008). Another advantage of using simulation models is that they allow scientists and researchers to experiment with ideas, especially when the phenomenon is too complex to be handled by simple equations.

Computer models used to represent how cities function can be static or dynamic (Batty 2006). Cities are complex systems with heterogeneous components that interact between them

(Torrens 2000). Critics of traditional urban models claim that they are centralized, deal poorly with dynamic phenomena, do not use detailed information, are not flexible, and lack realism. Currently, detailed information as well as increased computer power have allowed developers to model urban systems even at the individual level. New models in urban geography are heavily influenced by geographic information science, artificial intelligence, complexity studies, and simulation in social and natural sciences (Torrens 2001). Models of this kind use as their basic unit elements called automata. An automaton is a piece of software capable of receiving information from outside itself, process it, and proceed logically based on a set of instructions programmed within itself (Torrens 2001; Gilbert 2008). The main advantage of simulation models of this kind over traditional static approaches is a better representation of the dynamics of heterogeneous discrete elements that compose an urban system. The most common uses of automata in modeling spatial phenomena are: cellular automata and multi agent systems (Benenson and Torrens 2004; Torrens and Benenson 2005; Dietzel and Clarke 2004).

A business community is a complex system with diverse interactions, and is composed of heterogeneous members that influence each other. Agents have been proposed as a way to model systems of this type by previous researchers (An et al. 2005; Batty 2007; Miller and Page 2007; Epstein 2006). According to Benenson et al. (2005) and North and Macal (2007), object oriented programming is the most natural way of implementing large-scale automata based models. Using this approach it is possible to create an abstract class “agent” with a given number of attributes and methods and then construct specific members of this class, each with different attribute values. An important part of the model is the storage and display of automata states. Saving the states of the model at different run time points allow

researchers to examine the process outside the simulation for calibration and validation purposes.

4.2 Alternative Modeling Approaches

There is a significant amount of literature regarding dynamic models in social sciences. These models have different characteristics which affect their suitability under different circumstances. The most common of these modes are mathematical models, microsimulation, queuing models, and cellular automata (Gilbert and Troitzsch 2005).

4.2.1 Mathematical Models and System Dynamics

A common approach to represent dynamic systems has been through the use of mathematical equations. A classic example is the Malthusian growth model, presented in “An Essay on the Principle of Population” by Reverend Thomas Robert Malthus (1798), one of the most influential books on population. The Malthusian model is:

$$P(t) = P_0 e^{rt} \quad (1)$$

where P_0 is the initial population, r is the growth rate and t is time. This model considers the population as a whole and does not allow the study of elements or groups within the population.

Another classic example, although more complicated, is the one presented by Lotka and Volterra in 1911. Their model used differential equations to represent the continual change in an evolving system. This model incorporated the interactions between two or more species as well as age, structure and other characteristics of individuals (Webb 2008) (Volterra 1926). The basic Lotka-Volterra model can be represented by a pair of equations:

$$\frac{dx}{dt} = x(\alpha - \beta y) \quad (2)$$

$$\frac{dy}{dt} = -y(\gamma - \delta x) \quad (3)$$

Where y is the number of predators, x is the number of preys, and $\frac{dx}{dt}$ and $\frac{dy}{dt}$ represent the growth rate of the two populations through time.

System dynamics evolved from models based on difference and differential equations. Using this technique a system is modeled as a whole without considering the heterogeneity of the parts that compose it. An important difference between system dynamics and differential equations is that the first one uses discrete time steps as an approximation of continuous time. A second difference is the possible use of non continuous functions in system dynamics (Gilbert and Troitzsch 2005). The first language designed for system dynamic models was DYNAMO, although there are other well known tools, with additional features such as graphic interfaces such as STELLA, PowerSim and VenSim (Ventana Systems inc 2010). Using a system dynamics approach, Fiddaman (2002) developed an interesting model to test policies regarding climate change. This model includes elements of economy, energy, population, climate, and CO_2 emissions (Fiddaman 2002). The model was used to test tax and emissions permits policies following the Kyoto Protocol. The results of the model indicate that all policies have a positive result, however, tax policies outperform emission permits.

4.2.2 Microsimulation

A major limitation of differential equations and system dynamics models is that they do not take into account the different individuals, groups, and subclasses that comprise a population. In differential equations or system dynamics models the characteristics of the individuals are averaged and it is impossible to distinguish particular elements of the whole. The microsimulation modeling approach to this problem is to represent individuals with their particular characteristics and different transition probabilities. Models of this type are

stochastic while system dynamic models are deterministic. Microsimulation models consist of at least two levels: the level of individuals and the aggregated level. Microsimulation models are more data demanding than differential equations or system dynamics models, it is not uncommon that models of this type contain non aggregated data for several thousand individuals. Because of their high storage and processing requirements, there were few examples in the 70's and 80's. However, with the increase of computer power availability, they are becoming more popular in recent years (Gilbert and Troitzsch 2005).

An interesting example of microsimulation was developed by Liang (2009). In this research two models were integrated to simulate the evacuation of the population of New Orleans due to hurricanes. The first model represents the residents of each of the blocks in New Orleans as an agent in an agent-based model. At a coarser level a microsimulation model developed in VisSim simulates the flow of the residents through the highway system of the area (Liang 2009). The calibration of the model was done by running the model with different sets of values and identifying the ones that better match real traffic count data under normal conditions. The model was successfully validated after being compared to Hurricane Katrina evacuation traffic count data from August 27th 12.00 am to August 29 12:00 am in two locations: Fluker on I-55 and Slidell on I-10. The unit of analysis of this study was the census block. The model compared three evacuation strategies, 1) a simultaneous evacuation in which all census blocks of New Orleans started evacuating at the same time, 2) a staged evacuation based on social vulnerability; the most vulnerable census blocks based on social variables evacuate first, and 3) a staged spatial vulnerability based evacuation, where census blocks more vulnerable based on spatial factors evacuate first. The results indicate that both staged evacuations reduced the road network clearance time. The most effective in terms of

time was identified as the staged spatial vulnerability evacuation strategy, saving 3.6 hours compared to the simultaneous evacuation (Liang 2009).

4.2.3 Queuing Models

These are stochastic models, in which the temporal dimension is not continuous nor is it represented as regular time steps. In queuing models the temporal dimension is represented by events scheduled in an agenda. The result of a given event can generate new events that are inserted into the agenda, modifying it. In a queuing model there are four types of objects: the agenda, servers, customers, and queues (Gilbert and Troitzsch 2005).

A seminal work in the use of queuing models in geography is Hypercube Queuing Model, a computer program developed by Richard Larson under a grant from the National Science Foundation in 1974 (Larson 1975). The model is designed to be used by police departments, fire departments, and emergency medical services. The purpose of the model is to help planners to find the best mobile unit and service area configuration. The model allows the simulation of the increment/reduction of units and the redrawing the unit's service area. The output of the model consists in performance measures in terms of unit's workload and travel times. Using the model the planner can select the configuration with the best response time of mobile units, and the most suitable partition of the service area. The model requires the study area to be divided in reporting area, an expected number of emergencies per hour for each reporting area, and the average duration of travel for the units. The model is best suited for situations where at least 10% of the mobile units are busy at any point of time. It specifically takes into account that mobile units might have been dispatched to other emergencies. Under those circumstances the model assigns the emergency to other unit available in the district or calls from a unit of a neighboring district (Chaiken 1975) .

4.2.4 Cellular Automata

Cellular automata models represent a dynamic system using a regular lattice of identical cells. Each cell has a finite set of possible states determined by a set of transition rules (O'Sullivan and Torrens 2000). According to Gilbert and Troitzsch (2005), the main characteristics of a cellular automata model are: (1) It has a set of identical cells arranged in a regular grid of 1, 2, or 3 dimensions. (2) It has a set of finite possible states for the cells. (3) The model advances in discrete steps, in each step the cell can change its state. (4) The state change rules consider the neighbors' states to determine the cell's state. Because of this characteristic cellular automata models are better for modeling local interactions.

Because of the spatial representation of cellular automata models, they have become very attractive for geographers. Their regular space representation facilitates the use of data obtained from remote sensing imagery as data input. Due to their nature, modelers are able to examine dynamic processes at different time scales. An interesting example of the use of cellular automata in geography was presented by Dietzel and Clarke (2004). In their research they evaluate the capabilities of SLEUTH to predict the land use change in the area of San Joaquin in California. SLEUTH is an acronym for the spatial data layers required in the model: Slope, Land use, Exclusion, Urban extent over time, Transportation, and Hill-shaded backdrop (Clarke 2008). SLEUTH software was originally developed by Clarke in 1992 under the sponsorship of the United States Geological Survey's Urban Dynamics program. Its development continued under the NSF funded Urban Research Initiative (NASA 2007). The model has two main components, the urban growth model and a land use dynamics component called Deltatron (Dietzel and Clarke 2004). The model requires as input data raster files containing information regarding: slope, hillshade (for visualization purposes),

land use at two different time periods, an exclusion layer (indicating urban growth constraints), four different temporal layers with urban extent, and finally a transportation layer. Using the input data the model is able to estimate the value for five parameters, each with a value between 0 and 100: 1) Diffusion (how much urban growth is likely to disperse locally), 2) Breed coefficient: it defines the likelihood of new detached settlements, 3) Spread coefficient: it determines how likely is the diffusion from existing urban areas, 4) Slope resistant factor: it determines how likely is the development on steep slopes, 5) Road gravity factor: it determines the likelihood of new settlements along roads. The model evaluates values of 1, 25, 50, 75, and 100 for each of the parameters. From the 3125 possible combinations the model selects the one that gives more accurate results when compared to the reality. The model is tested in the area of San Joaquin in California with data from 1988 and 1996. The results indicate that SLEUTH can accurately replicate 93%, 77%, and 72% using 5, 10, and 15 land use classes, respectively (Dietzel and Clarke 2004).

4.3 Agent-based Models

Traditional modeling techniques for dynamic processes in Geography are based on differential or difference equations. When applied to geography these techniques usually divide the study area into administrative units, the available information of individuals is then aggregated at these unit levels (Benenson et al. 2002). The processes at the unit level are represented by equations that assume that all the individuals included in an areal unit have the same behavior. Inferences about individuals using the aggregated data lead to what is known as ecological fallacy. This is the assumption that members of a group have the average characteristics of the group as a whole (Longley et al. 2005). A circumvent solution to the ecological fallacy is the representation of heterogeneous individuals. In order to create models

of these characteristics, in recent years tools such as automata have been introduced in the field of Geoinformation Science. An automaton is a piece of software capable of autonomous actions based on its perceptions of the surrounding environment. Using automata it is possible to model at a very high detail level complex systems by representing its individual parts each with an automaton (Torrens and Benenson 2005).

An agent is a kind of automaton; its exact definition is still an open debate in the fields of computer science and artificial intelligence. Most of the researchers in the field would agree that an agent is a piece of software capable of autonomous actions in order to reach a goal (O'Sullivan 2008). Agents should also be able to perceive their environment and be able to react to it according to a set of rules of behavior. The rules can be complex, deterministic or stochastic, fixed or adaptive (Billari et al. 2006). While the environment can be synthetic like in a computational model or the real world. An agent-based model is a computational method in which agents represent real world actors interacting within a synthetic representation of reality (O'Sullivan 2008). Agent-based models have become attractive to social scientists because they allow them to explicitly describe relations between individual, heterogeneous entities as well as relations between entities and their environment (Gilbert 2008).

A model is a simplified representation of the real world where the scientist specifies how it is believed the reality operates. Models are in many cases used to perform experiments. An experiment consists of applying certain treatment to an isolated system and comparing the result to a system where no treatment has been applied. Doing this in social sciences and in the real world is difficult. A major attractiveness of agent-based models is that when well validated, a model of this kind can be used to perform social experiments without affecting real humans (Gilbert 2008).

According to Gilbert and Troitzsch (2005), agents typically have certain properties (although depending on the model, the degree of implementation of each of them varies): 1) Autonomy: An agent is able to perform actions without the intervention of a human user or a central control. 2) Social ability: Agents are able to communicate with other agents. 3) Reactivity: An agent should be able to perceive its environment and take actions based on that information. 4) Proactivity: Agents should be able not only to react to their environment but to take action in order to reach a certain goal.

4.3.1 Model Verification and Validation

Both users and developers of a model are concerned about the accuracy of a model, i.e., how accurate a model is in depicting reality. The evaluation of the model accuracy consists of two processes: verification and validation (Sargent 1998).

4.3.1.1 Verification

When coding a computer program it is common at the initial stages to find bugs. Bugs are segments of code that do something different from the modeler intended. In the verification process the developers of the model look for possible bugs in the computer program, ensuring that the program works as planned (Sargent 1998). There are some techniques suggested in order to reduce the possible coding bugs (Gilbert 2008):

- Code elegantly: The code should be written in a careful way, avoiding shortcuts, and using meaningful variable names.
- Include lots of output and diagnostics: The final output of the model is not sufficient to identify bugs. It is necessary to generate intermediate outputs of the different processes that compose the model, in this manner it is possible to identify units of code with problems.

- Observe the simulation step by step:
- Add assertions: Most of the variables have a designed range of values, the code should alert the developer when one of the variable has a value outside this range.
- Add a debugging switch: The assertions and diagnostics will make the model run slower. By using a debugging switch it is possible to control how much of the diagnostics or assertions code is executed in the model, allowing the user to set different testing levels.
- Add comments and keep them up to date: Computer programs allow the programmers to write text that is not executed by the model. This text is used to describe the functionality of the code.
- Use corner testing: This technique consists of running the model using extreme values for the parameters. The analysis of the outcomes would indicate possible bugs.

4.3.1.2 Validation

In this process the model is tested to determine if it has an acceptable level of accuracy for the intended purpose of use. There are two types of validation: conceptual and operational. A conceptual validation tests if the model is based on valid theories and assumptions. A second part of a conceptual validation is the evaluation of the model's logic and structure. An operational validation tests the accuracy of the outputs of the model by comparing them to the reality. This comparison should take into account the intended purpose of use of the model. An operational validation can be subjective or objective. A subjective validation can be done by examining the graphic outputs of the model or by using a variable-sensitivity analysis. An objective validation uses statistical techniques to compare the outputs of the model with reality (Sargent 1998). Verification and validation processes are required for any model

intended to have a practical use. However in the field of agent-based modeling, in many cases due to the scarcity of data, or the complexity of the system, it is not unusual to see examples where verification and validation have not been fully implemented, as it will be explained in a later section.

4.3.2 Trends in Agent-based Models

Heath et al. (2009) surveyed published works on agent-based models. To be included in the survey the work had to be published in a peer-reviewed outlet during the period between January 1998 and July 2008. Using these criteria the authors compiled a sample of 279 works from 92 publication outlets. The survey indicates that there is an increasing trend in the amount of research using agent-based model. The authors found 52 studies in 2007 and only 4 in 1999. In the survey the authors found 34 studies for the year 2008, however they only considered works published until July 20th 2008. The four main fields of the models were economics (29%), social science (24%), biology (14%) and military (13%).

Heath et al. (2009) indicate that there is not a predominant tool or computer language among the ones used to develop the models. Although the results indicate that the most popular tools are the ones that are of public domain.

Agent-based models can be programmed using conventional programming languages or using libraries or simulation frameworks. The most popular modeling frameworks as described by Gilbert (2008) are: Swarm, Repast, Mason, NetLogo. All of them are open source except for NetLogo, which also has better documentation. The documentation of the other software libraries is limited. Their user base is increasing except for Swarm that is becoming less popular. The development languages for all of them is Java, except for

NetLogo which has a proprietary language. Only Repast and Mason have limited support for geographic information systems.

The disadvantage of using libraries is that because of their complexity it could take months for a user to take full advantage of their capabilities (Gilbert 2008). According to Heath et al. (2009), the two most popular computer languages used to develop agent-based simulation models are C++ and Java. Both of them are object oriented languages and have some syntax similarities. Java's development was strongly influenced by C++, and most of its syntax derives from it, although it is simpler and has fewer low level capabilities. Java was designed to be portable and architecture neutral. It runs on a java virtual machine (JVM), which makes it slightly slower than C++ which interacts directly with the operating system (Horstmann 1999). Both languages can handle spatial data and perform spatial operations using external libraries. Among the most known open source libraries are FDO, GDAL/OGR, GEOS, GeoTools, MetaCRS, OSSIM and PostGIS (OSGeo 2011). Developing an agent-based model using a conventional programming language has the disadvantage of having to start from scratch. However the advantage is more flexibility (Gilbert 2008).

As mentioned above, two types of validation are needed: conceptual and operational (Heath et al. 2009). The conceptual validation certifies that the model is based on sound theories and assumptions, while the operational validation compares the output of the model with the outputs of the real system. Heath et al. (2009) considers two types of operational validation: statistical and non-statistical. The statistical validation relies on formal statistical hypothesis tests to check the validity and significance of the simulation outputs when compared to the outputs of the real system. The non-statistical validation relies on qualitative assessments of the outputs, for example an expert opinion.

The survey of the published works gives interesting results regarding validation. The results indicate that 29% of the models were not conceptually nor operationally validated, 17% had only conceptual validation, 19% had only operational validation and 35% had both conceptual and operational validation. For the models that were both operationally and conceptually validated, 95.0% used only non-statistical validation, 0.5% used only statistical validation, while 4.5% used a combination of both validation techniques.

4.3.3 Applications of Agent-based Models in Geography

4.3.3.1 Abstract Models

In this type of models the researchers do not attempt to represent the real world. The purpose of the model is to explore the implications of assumptions about agent behavior. In geography a classic example is Schelling's segregation model (O'Sullivan 2008). In this model, agents of two different classes are initially distributed in a grid space randomly. Agents are designed to prefer the vicinity of neighbors of their same class. In the model the agents move looking for a suitable combination of neighbors. The findings indicated that even with relatively tolerant agents, highly segregated patterns would emerge (Schelling 1971).

4.3.3.2 Regionally or Locally Specific Models

Most of the examples of this type are in the fields of Land use/land cover change (LUCC), pedestrian flow, and recreational behavior simulation for park management support (O'Sullivan 2008).

An interesting example in the field of LUCC is the one developed by Schreinemachers (2006). In this research the author uses an agent-based model to analyze the effects of soil fertility decline, population growth, and market structures on poverty and productivity. The study area is located in southeastern Uganda, it has an area of 12 km^2 , containing 520

households approximately. The data for this model was collected through household surveys and field observations. Each farm is represented as an agent. The decision model for the agents takes into account three factors: investment, production, and consumption. The decision rules were created using mathematical programming economic models (Schreinemachers and Berger 2006). The study concluded that trends in land and labor productivity were more strongly associated with population dynamics than with soil fertility decline. Another conclusion of this research is that improved access to innovations plus access to credit improves food security.

A well known study that uses agent-based models to represent pedestrian movement in real scenarios is by Batty et al. (2003). In this research the authors model the route of the Notting Hill Carnival. This is an annual 2-day event that gathers over 1 million people in an area of 3.5 km^2 in west central London. The peak time of the carnival is in the second day between 4pm and 6pm when there are around 260,000 visitors in the area. Because of the large crowd it is considered a major public safety concern. The purpose of the simulation is to test different parade routes to identify the safer one. In this model the authors create three types of agents, walkers (visitors to the parade), paraders (moving vehicles) and bands (fixed sound systems). The streets are considered a fourth type of agents and are used to direct the flow of the crowds. The data used to model the agent behavior was collected in the Carnival 2001. It came from video footage from police helicopters, data on entry and exit volumes at the subway station from surveys by London Underground Ltd., and bus volumes in entry points. The model has limitations due to software constrains. The maximum number of agents that can be represented is 16,000. However the authors were working to expand this limit and represent full size populations. The resulting model is suitable for simulating crowd control,

using the street agents to regulate the density of the crowd in specific areas of the carnival. An interesting characteristic of this model is that not only it allows researchers to simulate a real process but also intervene in it. Batty et al. (2003) concluded that the safest scenarios were the ones with longer parade routes. In these scenarios the public is spread in a larger area reducing the crowd density. For each of these scenarios a new traffic exclusion zone needs to be determined, as well as public entry points. The conclusions of this study are accepted by the Carnival Review Group and should be implemented in the future.

An interesting work in the field of recreational areas management support models is by Anwar et al. (2007). In this research the goal is to represent the interactions between whale-watching boats and whales in the Saguenay St. Lawrence Marine Park. The study area covers 1200 km^2 and includes several ports. The model considers two types of agents, the boat agents and the whale agents. The synthetic environment of the model is represented in raster format and contains information regarding the features in the area: land, water, reefs, ports and bathymetric information. The agent behavior for the boats was constructed based on GPS point samples collected from 341 excursions, scientific reports, and personal observation. In the model the authors consider two strategies for the whale-watching boats, cooperative and non-cooperative. The cooperative strategy stipulates that the boat agents share information regarding the whale location.

The results of the model indicate that a cooperative strategy for the boats largely increases the time of whale observation, however it also increases the risk of collision between boats and whales as well as increases the disturbance to marine mammals environment. The authors of this research plan to improve the whale behavior rules to better

represent the real movements of the marine mammals, taking into consideration of age, and species among other variables. For this model the researchers used Repast and Java.

4.3.3.3 Highly Detailed Urban Realistic Simulations

Models of this type attempt to recreate urban environments at a very detailed scale. There are previous examples of urban models using a cellular automata approach. However such approach has its share of criticism. Cellular automata represent the space with a regular grid. When representing urban environments the limitations of this form of representation became evident. It is difficult to represent network elements and spatial relationships of elements of different sizes. Benenson et al. (2005) propose the development of urban models using discrete objects that directly represent real world entities with fixed positions. They also propose the use of multi agent-based systems on top of the fixed entities to represent mobile elements of the model. The result of this approach is an environment called OBEUS (Object-Based Environment for Urban Simulation). Some of the ideas in OBEUS were first reported in Benenson et al. (2002). In the 2002 study, the authors represented the dynamics of ethnic residential distribution in the area of Yaffo in Tel Aviv between 1955 and 1995. In this model each householder is represented as an agent. The attributes for each agent are income, origin and religious affiliation. The synthetic environment for the agents is composed of entities representing the streets and buildings in the Yaffo area. The attributes for the streets are width, type, and street name. The attributes of the buildings are type (dwelling, business, public, etc), architecture style and number of floors. The householders decision rules regarding their residential location preferences are based on the surrounding environment and the neighborhood householders. For the parameter estimation, the results from a previous study carried out by one of the authors were used (Omer 1999). The results of the model were

successfully compared to Israeli's Population Census of 1995. The results of the model indicate that the most significant model variables to represent the population dynamics were the relation between householders and their neighbors and the relation between householders and the architectural style of the dwelling building.

4.4 Conclusions

Agent-based models and cellular automata have some similarities although they have two differences. The first is the space representation. Cellular automata represent space using a grid of 1, 2, or 3 dimensions, while agents can also be represented as 0 dimensional features. The second difference is that cellular automata have fixed locations while agents can change their location, allowing the representation of dynamic spatial relationships.

In this study we represent the businesses reopening process in New Orleans after Hurricane Katrina using primarily an agent-based modeling approach. The location of each business is determined by its address. There are businesses that share the same address, therefore in a 2 dimensional environment, those businesses are represented by 0 dimensional features with the same location.

Although the location of the businesses is fixed, their characteristics and spatial representation make them more suitable to be represented as agents. The businesses reopening process in New Orleans after Hurricane Katrina is dynamic. The components of the businesses community are heterogeneous. Each business is autonomous; it decides by itself when to reopen based on internal variables and its environment. There is communication between businesses. Each business perceives the reopening status of its neighbors. The final reopening decision takes into account this influence. Although the location of the businesses

do not change, the status of the businesses, open or closed, change through time, making the neighbor influence a dynamic process.

Although agent-based modeling has been applied with many examples in geography, the number of these models calibrated and validated using real world data is scarce. Furthermore the number of models validated using some statistical procedures is even fewer (Heath et al. 2009). One of the goals of this study is to develop a model for designing and testing of post-disaster recovery policies. A model of this nature needs to be correctly validated in order to be useful, using objective and subjective approaches. The subjective approaches could be in the form of expert opinion, whereas the objective approach requires a statistical measurement.

The businesses reopening process in New Orleans is located in an urban area that requires to be represented at a detailed scale. The simulation model requires that each business be represented independently as a discrete object. A suitable approach for a model of this type was proposed by Benenson et al. (2005). Although OBEUS is no longer in development (Benenson 2009), it is possible to implement some of its ideas using a computer programming environment.

A preliminary survey suggested that Repast or Mason could be the best tools for model development, because of their GIS capabilities. However the limited documentation of both simulation frameworks makes them difficult to use. An alternative development strategy would involve the use of a conventional computer programming language. The two most popular languages Java and C++ have similar capabilities. Both can access open source libraries with spatial capabilities like FDO, GDAL/OGR, GEOS, GeoTools, MetaCRS, OSSIM or PostGIS (OSGeo 2011). An important difference between Java and C++ is that the

former was designed to be conceptually simpler (Horstmann 1999), which makes it easier to use.

Chapter 5 : Development of an Agent-based Model for Simulating Business Reopening in New Orleans Post Hurricane Katrina

5.1 Introduction

Hurricane Katrina, with an estimated cost in damages around \$80 billion, is considered the most costly natural disaster in U.S. history (Stringfield 2010). It caused large population displacements, disrupting the economic life of the largest city in Louisiana. The post-disaster recovery is still an ongoing process several months after the disaster. A city affected by a disaster needs its businesses for a sustainable recovery. Studies on the post-disaster business recovery have been previously described in Chapter 2 and Chapter 3. A business post-disaster recovery process is dynamic and it is composed of heterogeneous parts that interact among them. These characteristics make it ideal to represent as an agent-based model. The use of this type of models in the field of post-disaster recovery is scarce (Oh et al. 2006). To our best knowledge there is no other attempt to model the recovery process of a business community after a disaster using an agent-based approach. A validated model of this kind would allow the simulation of diverse scenarios which could help decision makers to develop better policies.

The third survey data analysis, described in Chapter 3, comprised 1,185 businesses in total (see Table 3.2). Based on these results we propose the development of a model able to predict the reopening status for all the businesses in Orleans Parish. The simulation will include all the Orleans Parish businesses contained in the August 2005 Louisiana Department of Labor Micro File for Economic Development in the greater New Orleans Area. The total number of businesses is 10,169. The information in this dataset was used in Chapter 3 to estimate the reopening probabilities using binary logistic and probit regression models. The

sample used in the third survey (October 2007) is different from the first and second surveys (December 2005 and June 2006), allowing the use of the third survey results for calibration purposes and the first and second ones for validation procedures (See Figure 5-1).

5.2 Study Area and Data Requirements

The study area is Orleans Parish. Information regarding the business reopening status was obtained through three telephone surveys carried out in December 2005, June 2006 and October 2007. The surveys comprised all the businesses in Orleans Parish included in the August 2005 Louisiana Department of Labor Micro File for Economic Development in the greater New Orleans Area (the list includes businesses in Orleans, Jefferson and St. Tammany parishes) (Lam et al. 2009). Using this data source it was possible to estimate the business size using the natural logarithm of the total wages of the firm. The file also included the address of the business. Using this information and combining it with the flood estimation data for Orleans Parish after Hurricane Katrina (L.S.U. 2005), it was also possible to estimate the flood depth at the business location. One of the goals of this research is to study the relationship between the business reopening process and socio-economic characteristics of the residents in the study area. Information for the socio-economic variables was obtained from the Census 2000 at the block group level.

The socio-economic variables included in the simulation model were selected based on previous research by Cutter (2004). They were obtained at the census block group level, from the Census 2000. The selected socio-economic variables considered for the model were: population density, % non white population, % population with age under 18, % population with age over 65, median household income, % renter occupied houses, and % female population. Previous research carried out by Cutter (2004) identified these variables as

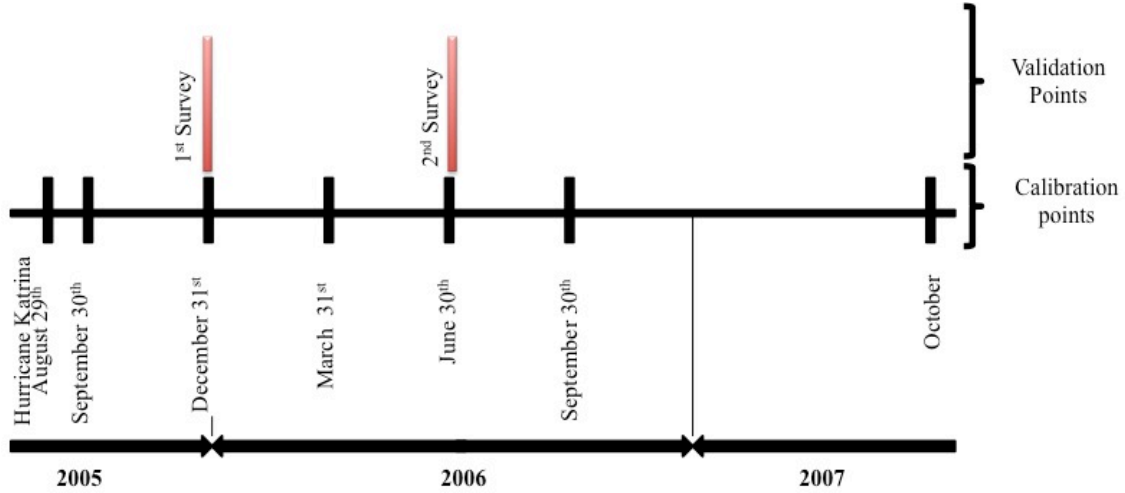


Figure 5-1: Calibration and Validation points

indicators of segments of society particularly vulnerable to disasters. After Hurricane Katrina, severe displacements of vulnerable population were recorded (Gutmann and Field 2010).

5.3 Model Conceptual Design

In this simulation model each business is considered as an agent. Each agent will estimate at discrete points in time its own reopening probability and decide to reopen or not, based on its internal characteristics as well as the input from its neighbors. In Chapter 3 we estimated the reopening probabilities using probit and binary logistic regression models. The simulation model uses the logistic regression functions instead of probit models to estimate the reopening probabilities. This model was chosen because it is easier to calculate on the fly.

The probability estimation functions have the following form:

$$Prob_t (y = 1) = \frac{1}{(1 + e^{-\eta_t})}$$

$$\eta_t = \rho_t W y_{(t-1)} + X \beta_t + \varepsilon$$

If $Prob_t > Threshold \Rightarrow Business\ status = Open$

where $Prob_t(y = 1)$ is the reopening probability of a business, ρ_t is the spatial dependence coefficient, $Wy_{(t-1)}$ is the spatial influence from neighboring firms, estimated using an inverse distance-weighting function, X is a set of variables describing business characteristics and socio-economic characteristics of the residents in the firm proximity, and β_t is the set of coefficients for the selected variables at a given point in time. Using this information it is possible to determine the reopening probability for each business for each period.

5.4 Model Operational Design

The businesses reopening process in New Orleans requires a very detailed or micro representation of an urban environment, where each business should be treated as an individual object. The approach proposed by Benenson et al. (2005) fulfills these requirements. OBEUS, the software introduced by Benenson et al. (2005) is not available and it is no longer in development (Benenson 2009). To implement the ideas proposed in OBEUS, we surveyed the most common approaches used to implement agent-based models. Most of the previous studies use either a simulation framework or a conventional computer programming language. From the most popular simulation frameworks, only Repast and Mason provide support for GIS. A survey of these frameworks suggests that their documentations are limited. An alternative is the development of the model using a conventional computer programming language and external libraries for spatial analysis. The two most popular programming languages for agent-based modeling, according to Heath et al. (2009), are Java and C++. In this study we decided to use Java because of its simpler concepts. There are many libraries that enable Java to handle geographic information and perform spatial analysis. The most well known open source libraries are FDO, GDAL/OGR, GEOS, GeoTools, MetaCRS, OSSIM, and PostGIS (OSGeo 2011). For personal preferences

it was decided to use GeoTools and PostGIS for the analysis of raster and vector data. The selected tools are of public domain, with each of them having a large base of users and good documentation.

The model uses information stored in raster as well as vector format. The vector information includes businesses' locations as well as census block groups boundaries for the study area. The raster information includes the flood depth estimation for Orleans Parish (L.S.U. 2005).

The vector information is stored in a database in a PostgreSQL system. This is an enterprise open source object-relational database management system (ORDBMS). PostgreSQL is capable of handling spatial information with PostGIS, a special add-on specifically designed for this purpose. The geometric features and methods supported by PostGIS are defined by the OpenGIS "Simple Features for SQL" specification (Ramsey 2011; Worsley and Drake 2002).

A flow diagram of the simulation model is depicted in Figure 5.2A. The model has an initialization phase in which it performs tasks that are needed only once. After this phase the model iterates until the number of time steps required is fulfilled.

The initialization phase has 5 steps: (1) First, the model reads an XML file that contains the coefficients required for the equations used to estimate the business reopening probabilities (See Figure 5-2). (2) Second, the model iterates along a user defined business list and instantiates the abstract business class with the attributes of each business. (3) Third, the model estimates the flood depth at the business location. The information regarding the business flood depth is obtained through a point in raster operation. For this operation the model uses Geotools, a free, open source Java geospatial toolkit released under GNU Lesser

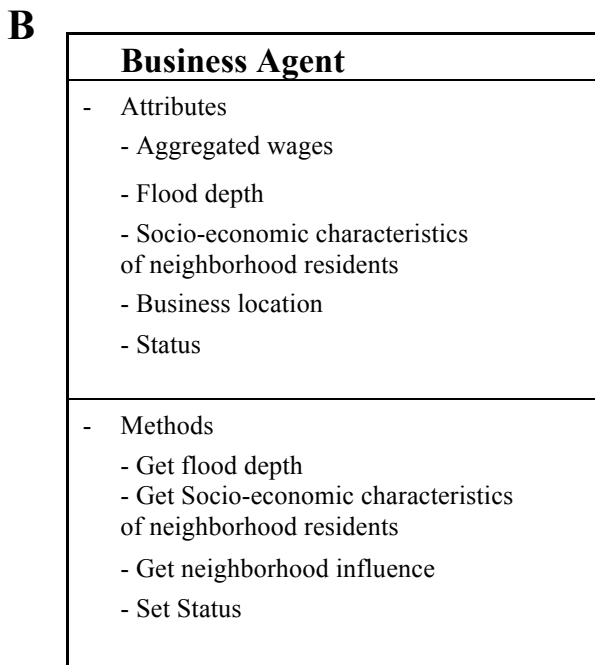
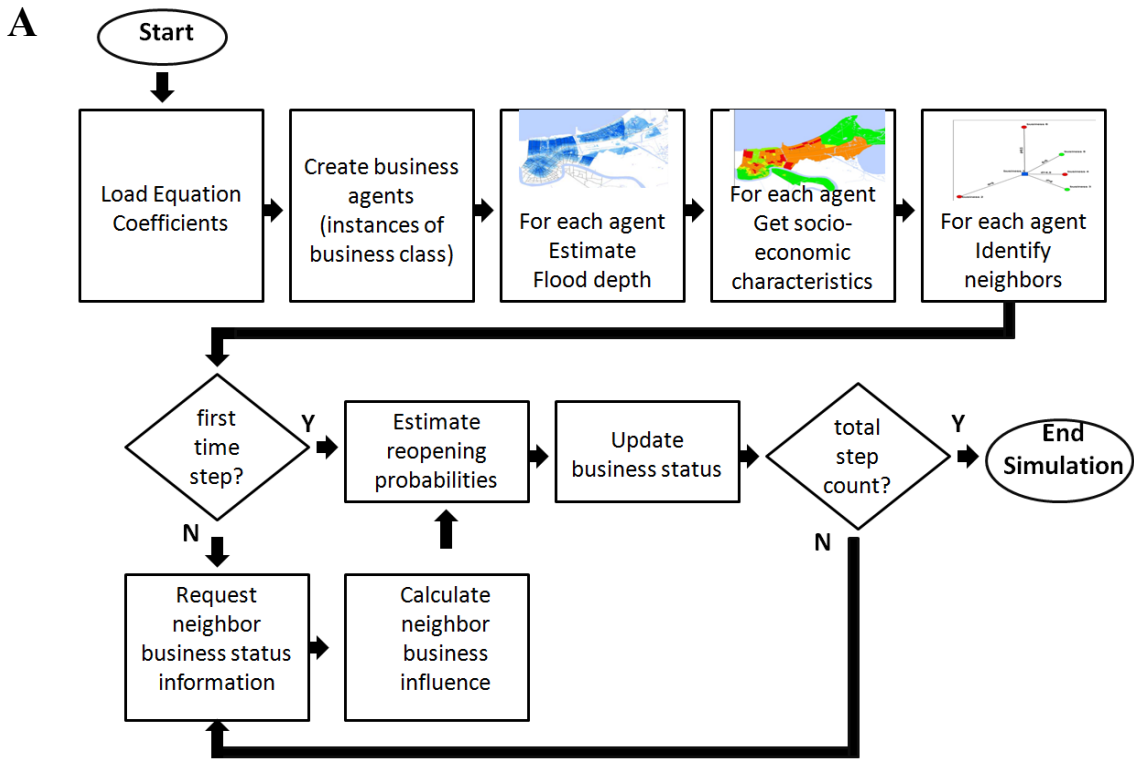


Figure 5-2: A) Model Flow chart, B)UML (Unified Modeling Language) design diagram of the business agent class

General Public License by the Open Source Geospatial Foundation (OSGeo 2011). (4) Fourth, the model obtains the socio-economic characteristics of the census block group where the business is located. This operation is a point in polygon operation. The spatial operation is performed using PostGIS commands in the form of an extension of standard SQL. (5) Fifth, the model identifies for each business its neighboring firms within 1000 meters. To estimate the influence of neighboring firms it is necessary to establish the distances between firms. One alternative is to do this task every time the model runs. A second alternative is to estimate the spatial relationships between all the businesses in the dataset and store these relationships in a database table. Every time it is necessary to determine the neighbor firms for a given business, we can simply query the relationships table without having to perform any distance calculation. The second alternative was selected because of its higher efficiency.

Once all the variables have been obtained, the model calculates the reopening probabilities for each business using the logistic regression coefficients loaded from an XML file in the initialization phase. At the first time step, it is not necessary to estimate the neighbor influence ($\rho = 0$), in subsequent steps the neighbors are identified using the relationships table. Once the model has identified the neighboring businesses for each firm, it queries about their status (open/close), and using this information plus the distances, the model calculates the neighbor influence using an inverse distance-weighting function (see Equations 15 and 16).

The reopening probabilities are compared against a threshold, which determines if the business opens or remains closed. The statuses of the businesses are then updated in the database. The simulation model is iterative, once the process reaches the last time step (6) the

simulation stops. The results of the model run for each time step are stored in a table in a PostgreSQL database and can be retrieved for later examination.

Following an object-oriented approach we designed an abstract business agent class. In the simulation model each business agent will be implemented as an instance of this class. The class bundles together attributes and methods common to all businesses. Figure 5.2B depicts a Unified Modeling Language (UML) design diagram of the business agent class.

5.5 Model Calibration

An analysis of the results of the third telephone survey carried out in October 2007 allowed us to construct logistic functions to estimate the reopening probabilities for the businesses in the sample (See table 5-1). The reopening probabilities are continuous between 0 and 1. The status of the business is a binary variable, with values of 0 for closed and 1 for open business. To determine if a business is open or closed we compare the probabilities against a threshold. If the probabilities are higher than the threshold we assume the business status is open (1) otherwise the business is closed (0).

To estimate the value of the threshold we estimated the reopening probabilities for the businesses that reopened in each time period. The reopening probabilities were calculated using the logistic regression coefficients derived in Chapter 3 (see Table 3-4) using SPSS. We observed that the reopening probabilities for the businesses that indeed reopened for the selected time periods vary through time. A summary of the statistics for the reopening probabilities for these businesses is depicted in table 5-2.

The average reopening probabilities for businesses that did not close after Hurricane Katrina was 0.17. Firms that reopened in the period between August and September 2005 had an average reopening probabilities of 0.24. While firms that reopened in the period between September and December 2005 had an average reopening probability of 0.54. As we can see

Table 5-1: Equations for the reopening probability estimation for each period.

Did not close after Katrina	$n_{t0=}$	$0.01(\ln \text{ wages}) - 0.043(\text{flood depth}) - 0.016(\text{pop dens}) - 0.002(\% \text{ non white}) - 0.006(\% \text{pop} < 18) - 0.006(\% \text{pop} > 65) + 0.243(\ln \text{ mhhi}) + 0.005(\% \text{renters}) + 0.031(\% \text{female pop}) - 5.336$
August 30 th – September 30 th 2005	$n_{t1=}$	$0.084(\ln \text{ wages}) - 0.72(\text{flood depth}) - 0.001(\text{pop dens}) + 0.003(\% \text{ non white}) - 0.012(\% \text{pop} < 18) - 0.002(\% \text{pop} > 65) + 0.266(\ln \text{ mhhi}) + 0.009(\% \text{renters}) + 0.008(\% \text{female pop}) + 0.719(\text{ni}) - 5.609$
October 1 st – December 31 st 2005	$n_{t2=}$	$0.087(\ln \text{ wages}) - 0.714(\text{flood depth}) + 0.002(\text{pop dens}) - 0.012(\% \text{ non white}) + 0.004(\% \text{pop} < 18) + 0.007(\% \text{pop} > 65) - 0.041(\ln \text{ mhhi}) - 0.001(\% \text{renters}) + 0.008(\% \text{female pop}) - 0.004(\text{ni}) - 0.118$
January 1 st – March 31 st 2006	$n_{t3=}$	$0.029(\ln \text{ wages}) - 0.105(\text{flood depth}) + 0.002(\text{pop dens}) + 0.010(\% \text{ non white}) - 0.037(\% \text{pop} < 18) + 0.015(\% \text{pop} > 65) + 0.176(\ln \text{ mhhi}) + 0.008(\% \text{renters}) + 0.016(\% \text{female pop}) + 0.326(\text{ni}) - 4.264$
April 1 st – June 30 th 2006	$n_{t4=}$	$0.081(\ln \text{ wages}) - 0.083(\text{flood depth}) - 0.011(\text{pop dens}) + 0.00(\% \text{ non white}) + 0.023(\% \text{pop} < 18) - 0.002(\% \text{pop} > 65) - 0.614(\ln \text{ mhhi}) - 0.020(\% \text{renters}) - 0.044(\% \text{female pop}) - 0.257(\text{ni}) + 7.022$
July 1 st – September 30 th 2006	$n_{t5=}$	$0.077(\ln \text{ wages}) - 0.009(\text{flood depth}) + 0.004(\text{pop dens}) - 0.001(\% \text{ non white}) - 0.030(\% \text{pop} < 18) - 0.014(\% \text{pop} > 65) + 0.159(\ln \text{ mhhi}) - 0.018(\% \text{renters}) + 0.114(\% \text{female pop}) - 1.456(\text{ni}) - 6.864$
October 1 st 2006 – October 1 st 2007	$n_{t6=}$	$0.029(\ln \text{ wages}) - 0.002(\text{flood depth}) + 0.014(\text{pop dens}) + 0.01(\% \text{ non white}) - 0.001(\% \text{pop} < 18) - 0.009(\% \text{pop} > 65) + 0.351(\ln \text{ mhhi}) - 0.005(\% \text{renters}) - 0.050(\% \text{female pop}) - 0.751(\text{ni}) - 2.333$

the reopening probabilities for the firms varies from period to period. In order to increase the similarity between the simulation model results and reality it was decided to have a different threshold for each time period.

We selected a list of possible thresholds and proceeded to compare the reopening probabilities against them. To evaluate which threshold value performed better, we estimated the overall accuracy of the statuses of the businesses resulting from comparing their probabilities with the different thresholds at each time period using the following equation.

$$\text{overall accuracy} = \frac{(\text{correctly classified open}) + (\text{correctly classified close})}{(\text{total number of businesses})}$$

Initially the comparison was done using thresholds values: 0.2, 0.25, 0.3, 0.35, 0.4, 0.45 and 0.5. The thresholds, which resulted in the highest overall accuracy for each time period were selected as the parameters to be used in the simulation model. In certain cases the tests using certain thresholds would result in complete misclassification of open or closed businesses. In those cases the threshold values were not considered and the overall accuracies were not estimated. The selected thresholds are highlighted in Table 5-3.

5.6 Model Validation

The main purpose of the validation is to prove that the simulation model accurately represents the real world system. A validated model can then be used in “what if” computational experiments. A common approach for model validation is to test the model against a set of real world cases and see if the model is capable of reproducing them (North and Macal 2007). In this research the actual status of the businesses, as recorded by the first and second telephone surveys, carried out in December 2005 and June 2006 respectively are compared against the corresponding partial results from the simulation model. The overall accuracy assessment indicates that the model can correctly predict almost 64% and 64.7% of the status of the businesses by December 2005 and June 2006, respectively.

To analyze the similarities and differences between the simulation model outputs and

Table 5-2: Probabilities of opening using the logistic functions.

	Did not close After Katrina	August Sep. 2005	Oct.- Dec. 2005	Jan- March 2006	April- June 2006	July- Sep. 2006	more than one year after the disaster
Min	0.06	0.07	0.05	0.09	0.05	0.08	0.10
Max	0.27	0.45	0.76	0.50	0.34	0.66	0.50
Average	0.17	0.24	0.54	0.32	0.18	0.27	0.30
Count	173	268	452	129	50	55	63

Table 5-3: Overall accuracy for each threshold (NA: not available, *:complete misclassification of one class)

Threshold values	Did not close after Katrina	Aug - Sept 2005	Oct.- Dec. 2005	Jan.- Jun 2006	April- June 2006	July- Sept. 2006	more than one year after the disaster
0.15	55.15	NA	NA	NA	NA	NA	NA
0.2	80.19	65.73	52.89	46.23	74.1	65.73	42.42
0.25	84.97	NA	NA	NA	NA	NA	NA
0.3	*	76.92	56.05	62.15	82.44	76.92	60.6
0.35	*	NA	59.11	68.39	84.22	NA	NA
0.4	*	79.72	62.92	72.47	84.82	79.72	71.86
0.45	*	80.16	62.38	72.9	*	80.07	73.16
0.5	*	80.77	62.92	72.26	*	80.77	73.16
0.55	*	80.25	62.35	NA	*	81.47	*

the survey results we estimated the business density of open businesses according to both datasets. In order to facilitate the description of the results we decided to use the names of the neighborhoods in the study area (See Figure 5-3).

Figures 5.4A and 5.4B depict the business status by December 2005, according to the telephone survey and to the simulation model respectively. The business status by June 2006 is represented in Figures 5.5A and 5.5B. In the first case, the map represents the actual

business status based on the telephone survey, while in the second it represents the simulation model results by the same point of time.

The similarities between the two datasets are obvious. However, for a more detailed analysis we estimated the differences between the kernel density surfaces corresponding to actual values and the simulation results:

$$[Difference] = [Density\ from\ actual\ values] - [Density\ from\ simulation\ results]$$

Ideally the values of the difference should be 0. Areas where the model has overestimated the reopening probabilities of firms have negative values, while areas in which the model has underestimated the probabilities have positive values.

Figure 5-4C depicts the differences between the actual values and the simulation results by December 2005. In this case the model has underestimated the reopening probabilities of firms located in the Central Business District. The model has also overestimated the reopening probabilities in the boundaries between West End, Lakeview and Lakeshore Lake Vista.

Figure 5-5C depicts the differences between the actual values and the simulation results by June 2006. In this case the main model underestimation areas are located in Gert Town, Tulane/Gravier and in the Central Business District, while secondary smaller clusters is located in Freret, East Riverside and Irish Channel. The main clusters of overestimated probabilities are located in the French Quarter, between Mid City and City Park and between LakeView and West End and Lakeshore Vista.

In general the differences between the predicted and the actual business status are located in specific areas, while for the rest of the study area the differences are minimum or none. The most noticeable underestimations occur in the Central Business District, while the



Figure 5-3: New Orleans neighborhoods (digitized from a PDF map from the Greater New Orleans Community Data Center website)

overestimation occurs in Mid City, Lake View and West End and Lakeshore Lake Vista.

The most obvious difference between the Central Business District and other areas of New Orleans is the business density. This area has a higher concentration of firms, and its density is comparable only to the French Quarter, its neighbor. The second obvious difference is the type of buildings in the area. Many of the businesses that operate in the Central Business District are located in high rise buildings. In the event of a flood the only businesses directly affected would be the ones operating at the ground level. The others would be mostly unaffected and could return to operate as soon as the minimum business requirements are provided.

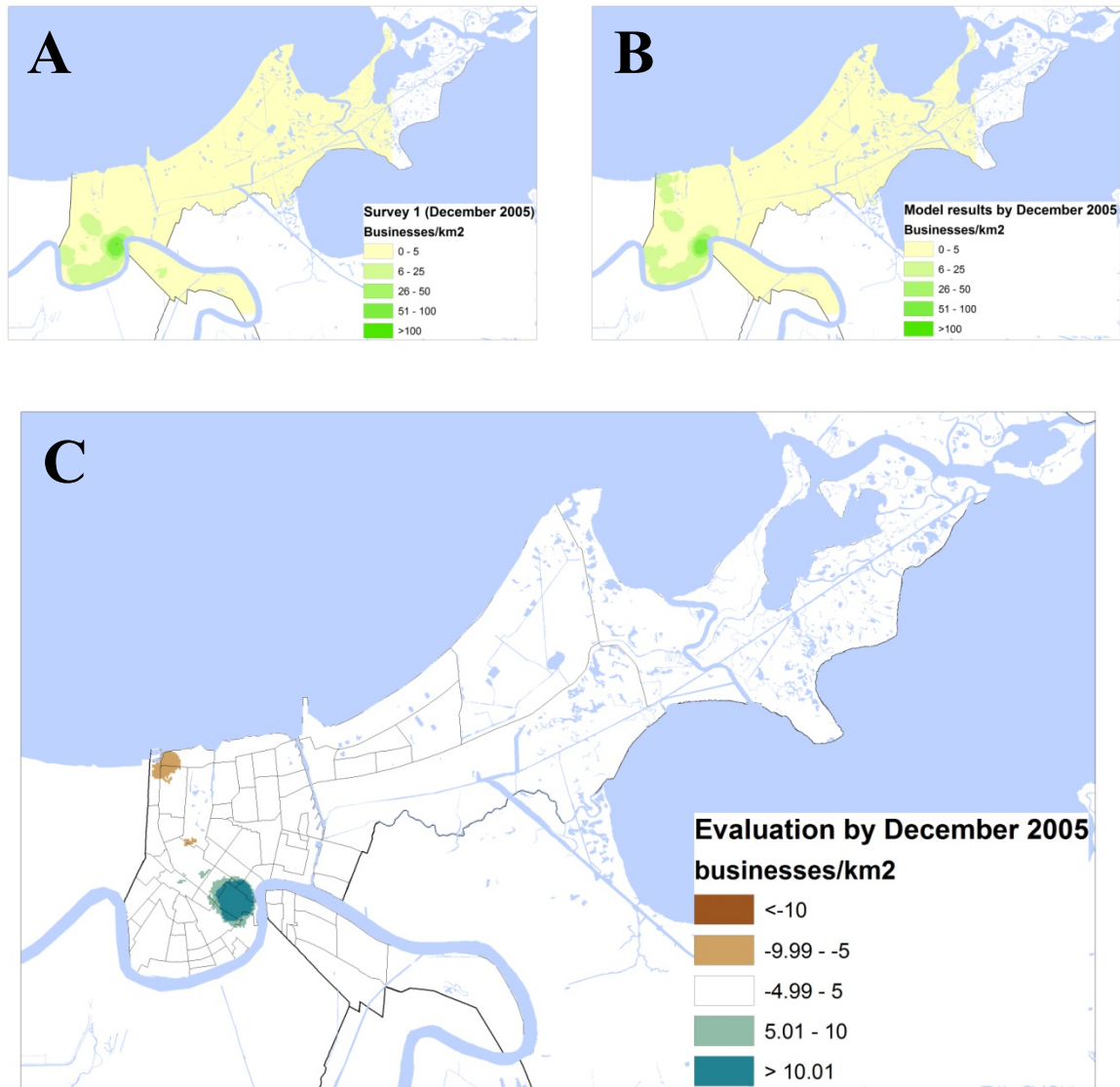


Figure 5-4: Model evaluation by December 2005, A) Business density according to first survey (December 2005), B) Business density based on model results by December 2005, C) Differences between model and survey data (negative values indicate overestimated predicted results, positive values indicate underestimated model results).

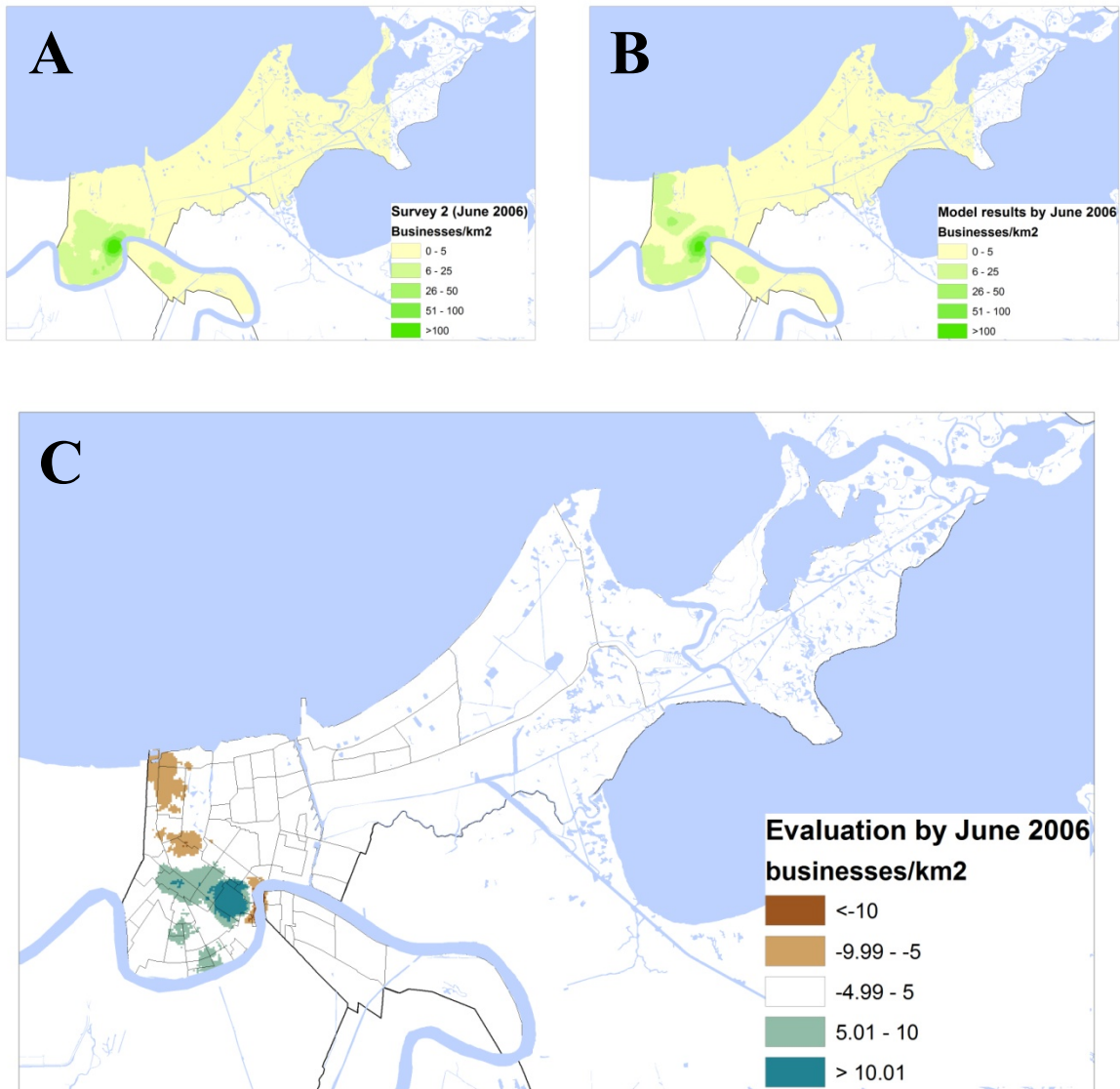


Figure 5-5: Model Evaluation by June 2006 A) Business density according to first survey (June 2006), B) Business density based on model results by June 2006, C) Differences between model and survey data (negative values indicate overestimated predicted results, positive values indicate underestimated model results).

The overestimation of the reopening probabilities mostly occurred in the Mid City, Lake View, and West End, neighborhoods and in a less degree in the French Quarter. Lake View and West End were severely flooded. That the residents return rate could affect the businesses reopening decision. There is not an exact detailed count on the return of evacuees to New Orleans in the months after Katrina. Postal delivery counts have been proposed as a proxy to estimate the population return after the disaster (Plyer et al. 2010; Plyer and Hodges 2008) . The United States Postal Service submits to the Department of Housing and Urban Development quarterly reports at the census tract level with counts of the number of addresses actively receiving mail. However due to the havoc created by Katrina, the USPS was not able to deliver accurate counts in first several months after the disaster (Plyer and Hodges 2008) . The quarterly report corresponding to June 2006 indicates the percentage of active addresses at that time and is mapped in Figure 5-6.

A traditional interpretation of the variable household income is that high income households are less vulnerable to disasters (Cutter and Emrich 2006). Lake View and West End are neighborhoods populated by high income households. The average household income in these neighborhoods is approximately \$51,020 (there is not a perfect match between the neighborhoods and the census block groups boundaries, therefore this is an approximation), while the mean household income for Orleans Parish is \$25,590. However the population return for these neighborhoods at the time of the survey was very low (See Figure 5-6). Elliot (2006) reported that the evacuee return likelihood is higher among low income homeowners, mostly because they can not afford to relocate for long periods. High income households on the other hand have the option of staying out from their homes for longer time. This fact is reflected on the low population return for Lake View, and West End, which might have

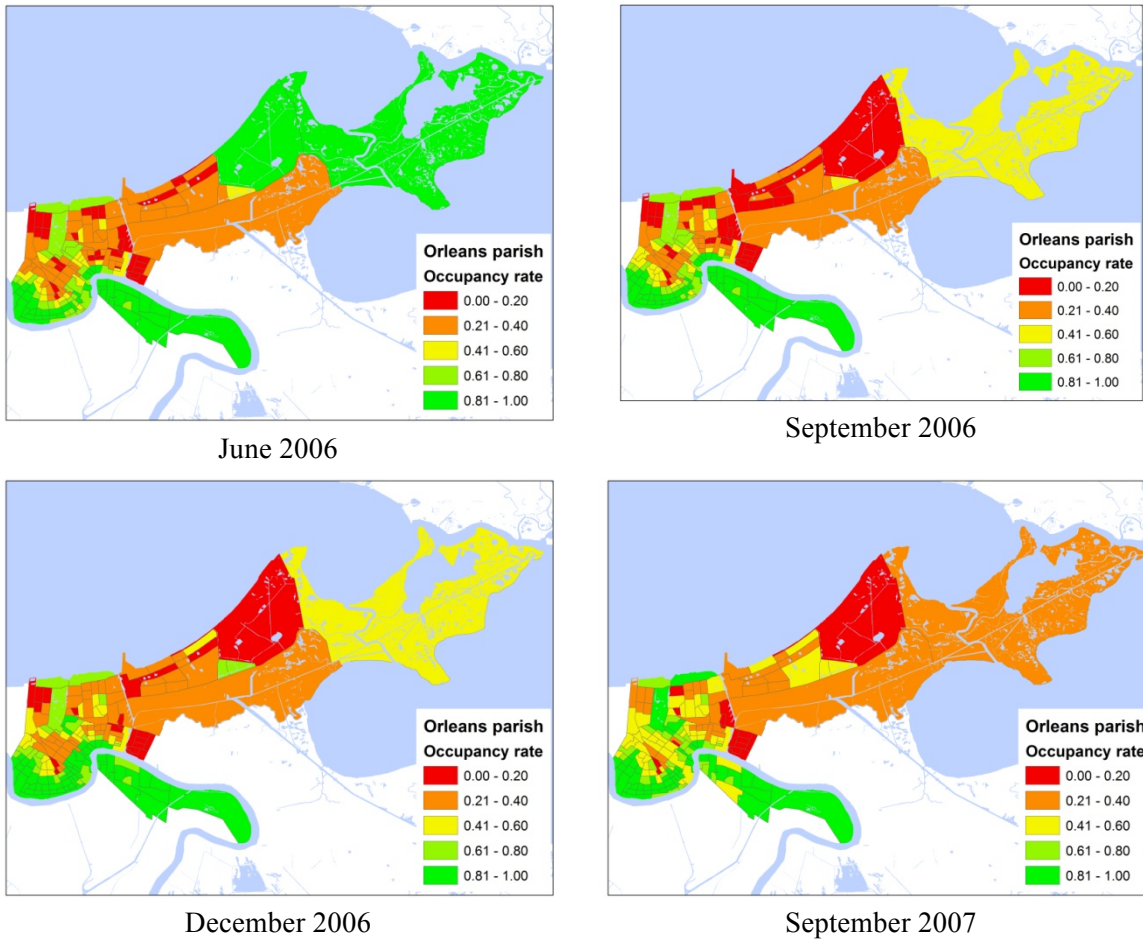


Figure 5-6: Percentage of addresses actively receiving mail at the census tract level ((total addresses-(vacant-no stat))/total addresses).

negatively affected the businesses reopening.

Overall, the simulation model can correctly predict the reopening status of between 64% and 65% of the businesses included in the evaluation. The evaluation of the results indicates that the model is suitable to be used for simulating alternative scenarios.

5.7 Disaster Scenarios

We created three scenarios to analyze the reopening process with the simulation model. In the first scenario we extrapolated the results from the sample used to estimate the binary logistic functions to all the business population in Orleans Parish. In the second

scenario we evaluated the reopening process with flood conditions different from the actual one. And finally in the third scenario we tested the effects of post-disaster policies that could encourage certain businesses to remain open after a severe flood.

We estimated the coefficients of the binary logistic functions using all the businesses included in the third survey (October 2007), which represent a subset of all the businesses listed in the August 2005 Louisiana Department of Labor Micro File for Economic Development in the Greater New Orleans Area.

5.7.1 Scenario 1

In our first scenario we extrapolated the results from the sample to the whole business population. The evolution of the density of the reopened businesses as predicted by Scenario 1 can be seen in Figures 5-7 and 5-8. The reopening trends follow the findings of Chapter 3, indicating an increase of the recovery rate until December 2005 (See Figures 5-7A, 5-7B, and 5-7C). After this point the recovery rate decreases. The number of returning businesses still increases but at a much lower pace. After June 2006 the business return rate is minimal, indicating a much stable phase of the process (See figure 5-8B and 5-8C).

5.7.2 Scenario 2

For the second scenario we considered a fictitiously less severe flood for Orleans Parish. The goal is to study how the businesses would react to a disaster of different magnitude. The scenario can be understood as an evaluation of the return process if certain measures designed to increase the business resilience are implemented. The original flood depth data were obtained from the Louisiana State University GIS Information Clearinghouse. This information is in raster format with a spatial resolution of 25 meters. For the less severe flood scenario, we created a new raster with the same spatial resolution, although with flood

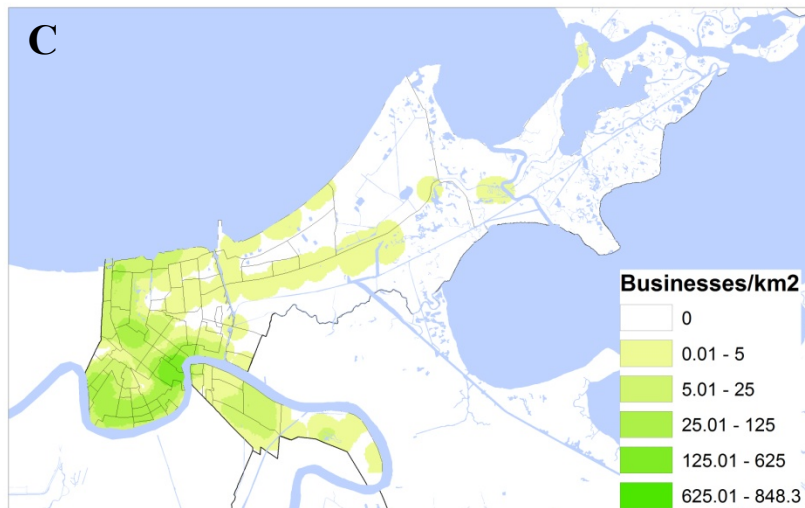
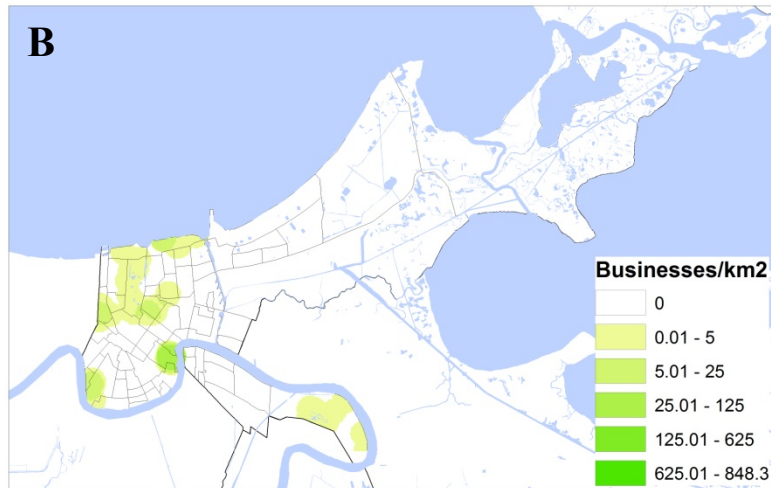
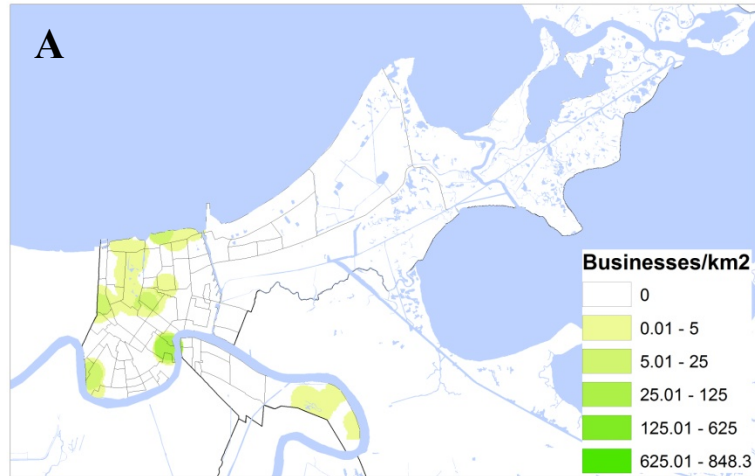


Figure 5-7: Business density for Scenario 1: A)By August 30th 2005, B)By September 30th 2005, C) By December 31st 2005

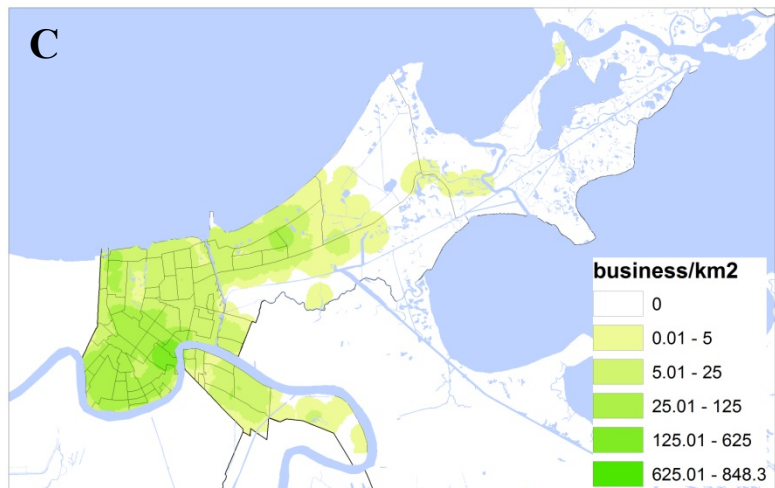
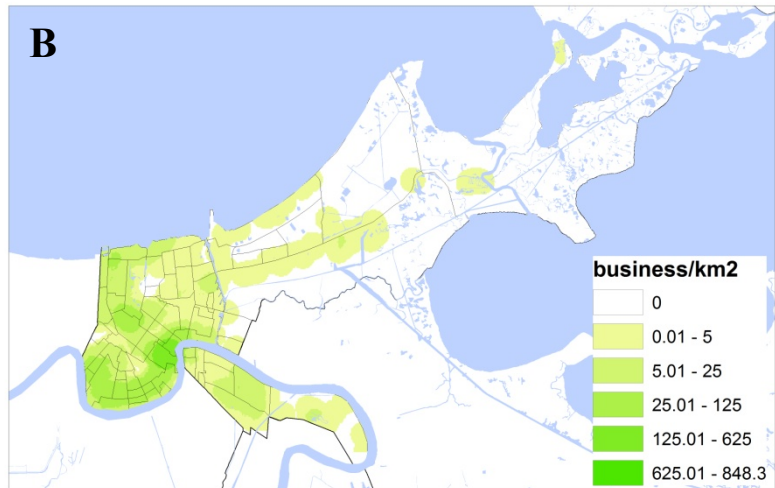
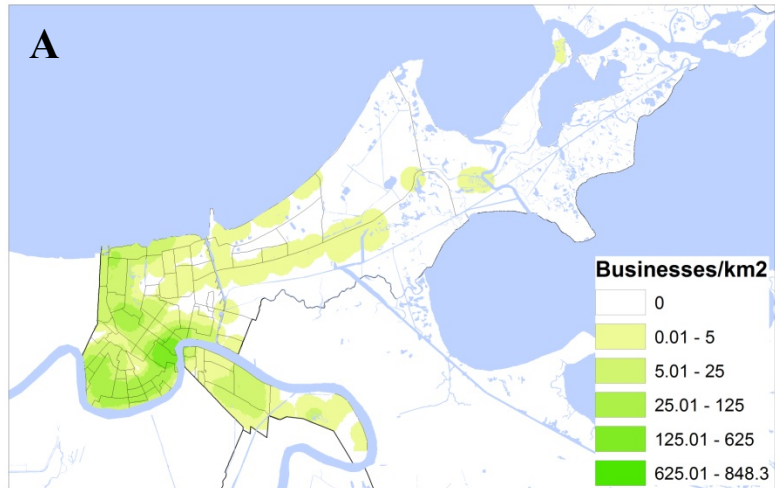


Figure 5-8: Business density for Scenario 1 (cont.): A) By March 31st 2006, B) By June 30th 2006, C) By September 30th 2006.

depth values lower by 1 meter. To accomplish this task we used Erdas Imagine.

The results for Scenario 2 indicate that less flood generates better reopening probabilities as it would be expected. To visualize the differences between Scenario 2 and Scenario 1, we calculated the differences in the reopened business density at different points of time.

$$\text{Differences} = \text{Business density for scenario 2} - \text{Business density for scenario 1}$$

The differences between Scenarios 2 and 1 are very small from immediately after the disaster until the end of September 2005 (See Figure 5-9A). A reduced flood damage would increase reopening in: Fairgrounds, Leonidas, Lakewood, Lakeshore Lake Vista and Lake Terrace. Although these reopenings are small in number in the period between October 2005 and December 2005, less flood damage has a stronger effect. The most significant differences between scenarios 2 and 1 occurred in the Central Business District and in Mid City (See Figure 5-9B). The differences between both scenarios remain more or less constant until June 2006. In the period between June and September 2006 the differences between both scenarios began to fade as can be seen in Figure 5-9D. The major differences occurred in Central Business District and Mid City and in a lower degree in Tulane/Gravier.

The positive effects of reducing the flood depth can also be observed in Treme/Lafitte, Bayou St. John, Fairgrounds and part of the Seventh Ward, although in a lesser degree. All these neighborhoods were actually affected by the flood. In the less severe flood scenario, businesses located in these areas show in general higher reopening probabilities compared to the first scenario especially starting in the last quarter of 2005. After June 2006 Scenarios 1 and 2 show more similar reopening results, indicating that the effect of flood damage decreases.

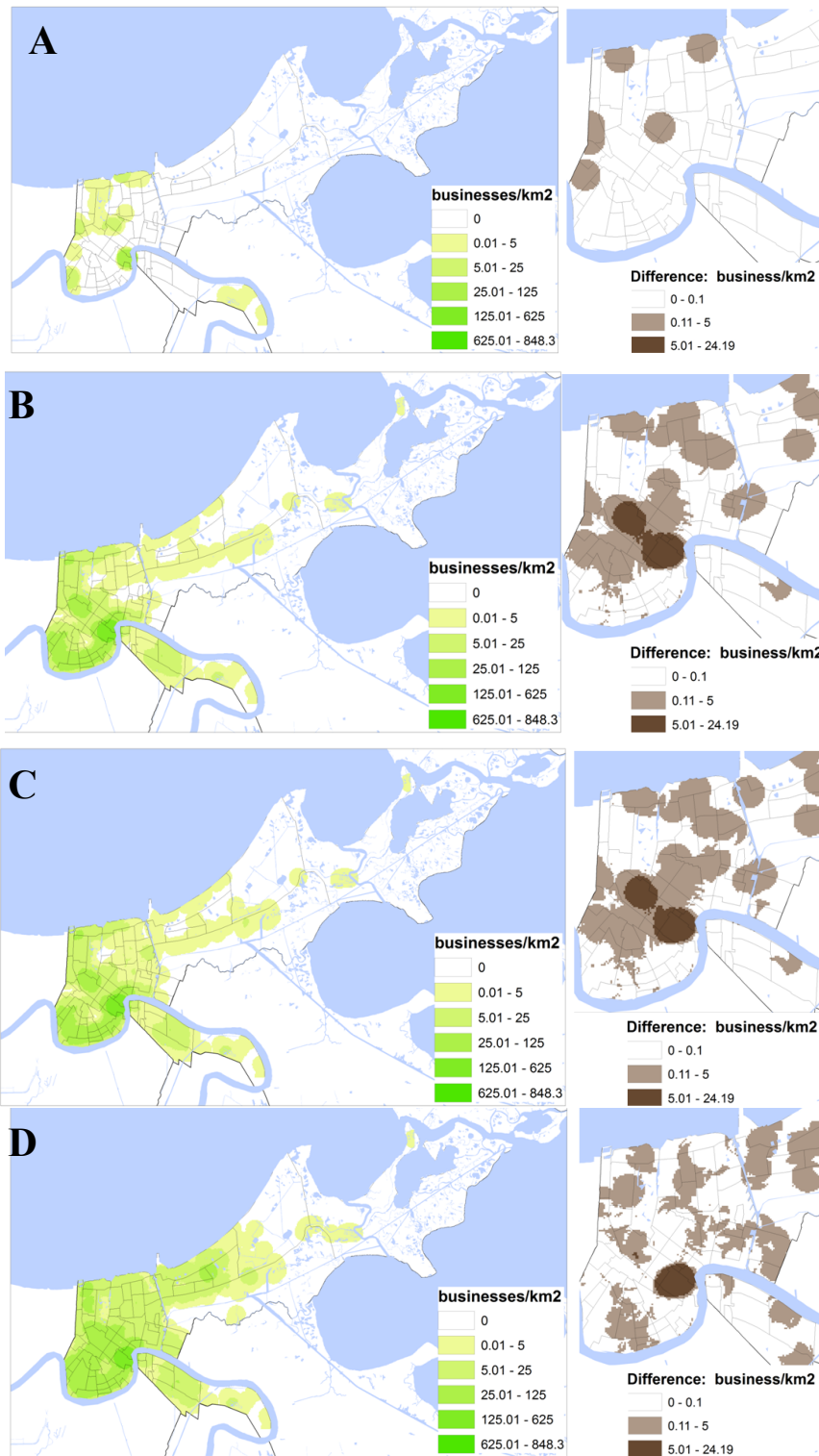


Figure 5-9: Businesses density for scenario 2 and density difference between scenario 1 and scenario 2 : A) Immediately after Katrina, B) By December 2005, C) By June 2006, D) By September 2006.

5.7.3 Scenario 3

For the third scenario we simulated that certain businesses would remain open after a disaster, due to a given policy designed for this purpose. The goal of Scenario 3 is to observe the effect of reopened businesses on their neighbors. In Scenario 3 we encouraged 50% of the businesses to remain reopen after the disaster in selected locations, and observed the effect this would have on the rest of the firms. The selected locations suffered a significant amount of flood and did not rebound immediately after the disaster.

The first test area is located in Mid City, it is composed of census block groups 220710050001 and 220710054003 (See Figure 5-10). Its boundaries are: Canal St, N. Hennessey St, Orleans Av., and N. Rendon St. This area had on average 1.03 meters of flood, 44.7% of its population was non white, while the median household income was \$35,627.

The second test area is located in West Lake Forest. It comprises census block groups 220710017361 and 220710017362 (See Figure 5-8). It is bordered by the I-10 highway, Read Blvd., Dwyer Rd. and Crowder Blvd. On average it suffered 1.4 meters of flood. Most of its population was non white (97.9%) with an average household income of \$22,057.

We encouraged 50% of the businesses located in our selected areas to remain open after the disaster so that we can see the effect of adding more reopened businesses on neighboring firms. The selection of the businesses that remain open was random.

The simulation results for Scenario 3 are depicted in Figure 5-11. Similar to the procedure employed in Scenario 2, we estimated the differences between the results of Scenario 3 with the results obtained from Scenario 1.

$$\text{Differences} = \text{Business density for Scenario 3} - \text{Business density for Scenario 1}$$

The results show that although the neighbor influence was a statistical significant

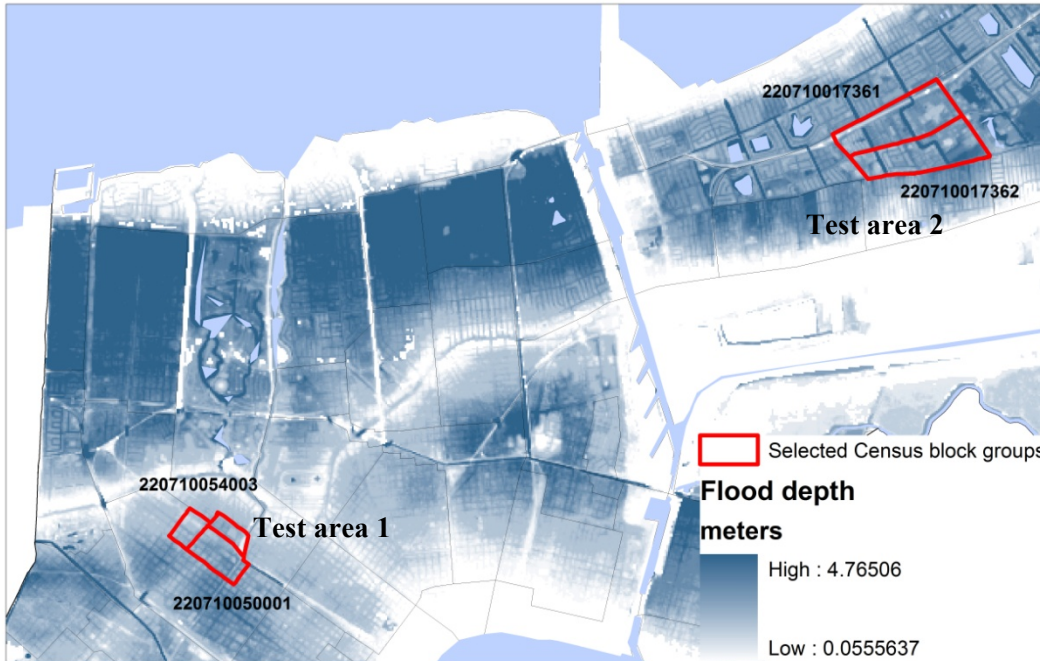


Figure 5-10: Selected census block groups for test areas

predictor for the period between August 30th and September 30th, it is hard to distinguish any impact on the businesses located in the proximities to test area-1 in this period. The major difference was in the next period between October and December 2005. Starting in this period we can observe an increase in the reopening rates to the east of test area-1. This effect occurs until the period June-September 2006. In this period Scenario 3 starts reaching similar reopening rates to Scenario 1.

In test area-2 it is not possible to see any effect on the neighboring firms outside the selected block groups. The density of the businesses in the proximity to this test area is lower compared to Mid City, hence the effect that the neighbor influence might have is reduced. The analysis of the third survey indicates that this area began to see some business return, only after June 2006. The results suggest that the conditions of the vicinity of test area 2 nullify any benefit resulting from forcing some business to reopen.

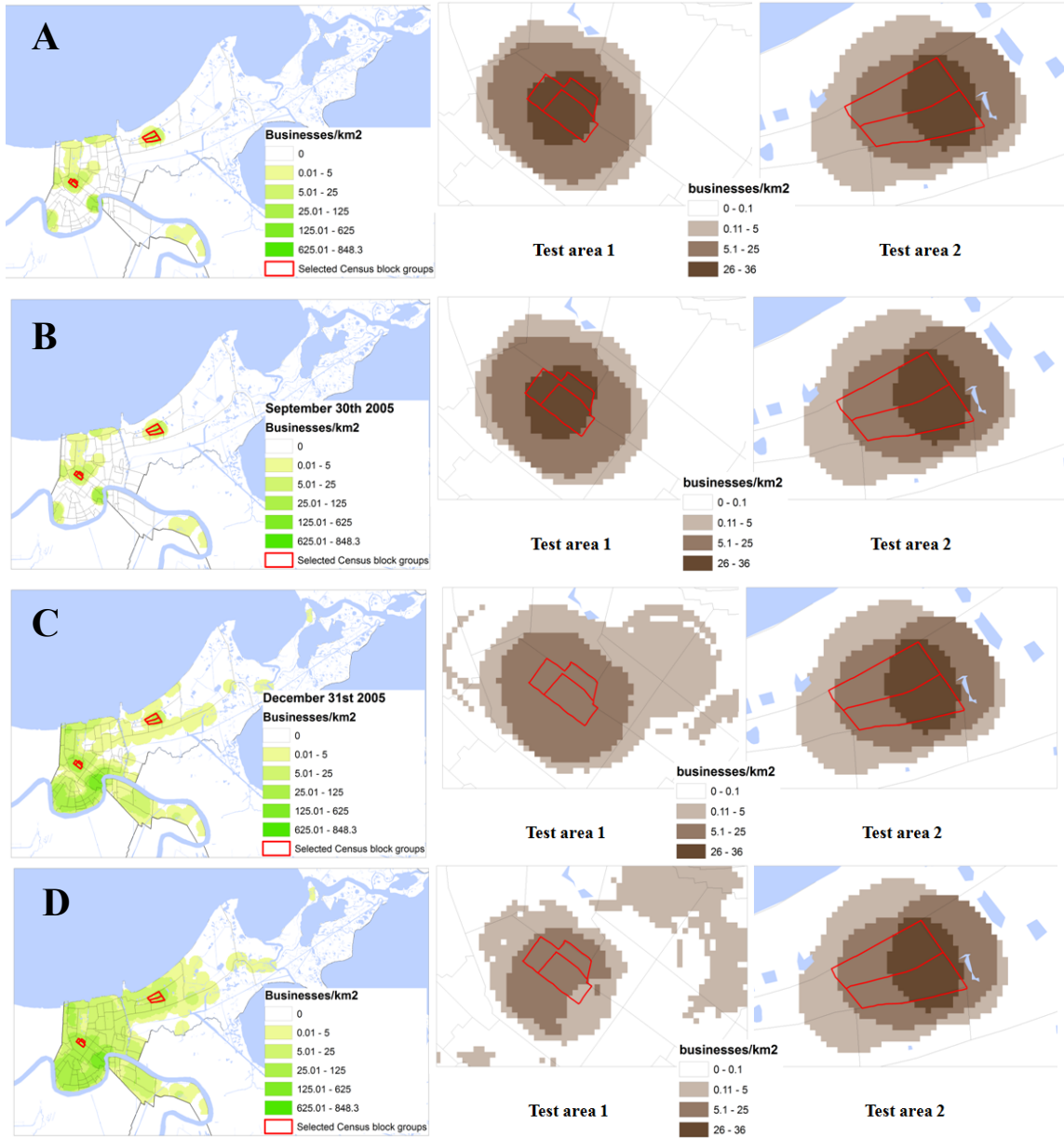


Figure 5-11: Business density for Scenario 3, and density differences between scenario 1 and scenario 3: A) By August 30th 2005, B) By September 30th 2005, C) By December 31st 2005, D) By September 2006.

The results indicate that the effects of the forced reopened businesses start fading after June 2006, when both scenarios start generating similar outputs.

Chapter 6 : Conclusions

6.1 Summary

The business reopening process in New Orleans is dynamic. The reopening rate is not constant through time. This is due to the fact that the disaster did not affect all businesses in the same way. To study this process we used three telephone surveys carried out in December 2005, June 2006 and October 2007 as described in Lam et al. (2009). Using the information obtained in the third survey we studied the reopening probabilities for businesses in New Orleans after the flood caused by Hurricane Katrina. In the second part of this study we modeled the reopening process using an agent-based technique.

The results of the statistical analysis of the third survey data indicate that the factors that had an effect on the business reopening probabilities were: business size, flood depth, socio-economic characteristics of the residents living in the vicinity such as population density (population by square kilometer), % non white population, % population with age under 18, % population with age over 65, median household income, % renter occupied houses, and % female population, and the influence of neighbor firms.

We used the $\ln(\text{total wages})$ as a proxy for business size. Our results suggest that larger businesses were less affected by the disaster than smaller firms. Larger business showed better reopening probabilities than their smaller peers. According to our analysis this variable was significant until December 2005. After this point business size was no longer a significant predictor.

As expected, flood depth at the business location was found to have a negative effect on the reopening probabilities. Businesses located in non flooded areas had better reopening chances than firms located in flooded areas. This predictor was significant until March 2006.

Certain socio-economic characteristics of the residents in the vicinity of the business were identified as significant predictors. Our analysis indicates that the percentage of non white population and the percentage of population with age under 18 negatively affected the business reopening probabilities. The significance of percentage of non white population was limited to the period between October 1st and December 31st, 2005. The percentage of population with age under 18 was significant in the period between January 1st and March 31st, 2006.

Our results indicate that businesses were positively affected by the percentage of female population in their respective block groups between July 1st and September 30th of 2006 at a significant level. In other words, businesses located in census block groups with higher percentage of female population had higher reopening chances in this period.

Our results also indicate that businesses influence each other. There was a positive influence of neighboring firms reopening until September 30th 2005. Then this variable was not found significant until the period between July 1st and September 30th 2006, when it became a negative predictor. It is possible that after some time with more businesses returning, the competition for suitable locations made it difficult of later returnees to reopen.

The heterogeneity of the businesses community as well as the interactions between firms and the dynamic nature of the reopening process demand a special type of modeling technique. These requirements are fulfilled by the use of an agent-based modeling approach. Based on a literature review it was decided to create a model using Java and open source libraries with spatial capabilities like GeoTools and PostGIS. The model developed in this study represents each business as an individual entity with the business characteristics (business size, location). It also considers the characteristics of the residents living in the

vicinity of the firm, as well as the flood depth at the business location. The model considers the spatial influence of neighboring firms using an inverse distance-weighting function. To estimate the reopening probabilities it was necessary to choose between logistic and probit functions. We used a logistic function to estimate the businesses reopening probabilities in the simulation model because it is easier to calculate on the fly.

The model was calibrated and validated using real data. The model works at a reasonable level. It was able to predict around 65% of the reopening of the businesses at the validation points. The model validation was based on the empirical survey data, allowing its use for more realistic simulated scenarios.

Three scenarios with simulated conditions were evaluated. In Scenario 1 we included in the simulation all the businesses in Orleans Parish. In Scenario 2 we assumed that the flood suffered was of lower magnitude, representing the implementation of measures designed to make buildings and infrastructure more resistant to floods, like the raise in elevation of buildings or better levee/pumping infrastructure. The scenario used a reduced flood depth of 1 meter to simulate the improvements in the city. The results show that a reduced flood depth increases the reopening probabilities. The effect of this change is observable from October 2005 until September 2006. The areas that benefit most from a reduced flood depth would be the Central Business District, Tulane Gravier, and Mid City. The benefit was lower for Tremé/Lafitte, Bayou St. John, Fairgrounds, and part of the Seventh Ward. After June 2006 the scenario with reduced flood depth became more similar to the scenario in which no conditions were changed.

In Scenario 3 we evaluated the effect of an aid strategy that would make certain businesses totally resilient to the flood. This increase in resilience was implemented in the

model by making their status to “open” immediately after the disaster. The purpose of this simulation was to observe the spatial effect that opened businesses had on their neighbors. For this scenario we selected two areas in the study area in which we randomly selected a number of businesses that would benefit from the “increase resilience policy”. In the first selected area, the policy had a positive impact in the period between October and December 2005. After this period the effect tends to stabilize until June 2006, when the impact of the artificially reopened businesses began to decrease. In the second site the increase of reopened businesses had no impact. The results suggest that the effect of making more business to open was compromised by the business density and the socio-economic conditions of the area.

The agent-based model allows the evaluation of the business reopening process using simulating conditions different from reality. It also allows the study of the spatial effect resulting from changing the values of any relevant variable at any specific location. In Scenario 3 we increased the resilience of a certain group of businesses to the disaster. In a similar fashion it is possible to evaluate the effect of socio-economic changes in the characteristics of the resident population in specific parts of the study area. The microscopic nature of this type of models makes them extremely attractive for simulation in urban systems with environments composed by heterogeneous components, specially nowadays when detailed information is becoming more available.

6.2 Contributions

This research contributes to the fields of dynamic spatial modeling and disaster and vulnerability science. This research is innovative in four aspects.

1) Although there is some research where an agent-based approach has been used to study of post-disaster recovery processes (Oh et al. 2006), the simulation of the business

reopening process using businesses as agents is novel. To our best knowledge there is no previous research on simulating the business recovery process using agents to represent businesses in an agent-based simulation model. The model architecture allows researchers to represent with high detail very heterogeneous populations such as the business community in a city.

2) This research is also innovative in the use of real data for the calibration and validation procedures. The validation and calibration process make this model reliable. Although its use in other applications would require additional calibration, the simulation model as an architecture can be easily adapted to other spatial locations and disasters of different nature.

3) This research considers the spatial effect of the decisions of neighboring firms on a business reopening decision. The distance was considered as a weighting factor where closer firms have a stronger influence than farther neighbors. By considering the spatial dimension of the businesses decisions, the realism of this model increases.

4) The model uses XML initialization files, which provide the model with great flexibility. The model can be easily modified to simulate diverse scenarios, allowing the researchers to modify specific variables without having to change the simulation model as a whole.

6.3 Future Research

This research demonstrates the feasibility of an agent-based approach for the simulation of the post-disaster business reopening process. This model has been calibrated and validated. However the model can be improved by including other variables. The results of the logit and probit regression analysis indicate that the resident return influences the

businesses reopening probabilities. After a disaster, population return is difficult to estimate, but future research should consider proxies such as mail delivery and school enrollment among others (Plyer et al. 2010).

In our validation process we found that the model inaccuracies were located mostly in small areas. Although we suggested some reasons for which these areas might behave differently than the rest of the study area, further research is required. Future studies could consider other variables, such as the building type from which the business operates and the economic relationships between businesses. In addition, distance scales other than the 1-km criterion used in this study may need to be experimented to explore the scale effects on modeling business reopening.

The calibration of this model was done with data from the third telephone survey carried out in October 2007. Previous telephone surveys (December 2005 and June 2006) did not gather information about the number of weeks the businesses were closed after the disaster. Future surveys in this field should obtain temporal information for a better statistical analysis.

References

- Adger, W. N., T. P. Hughes, C. Folke, S. R. Carpenter & J. Rockstrom (2005). Social-ecological resilience to coastal disasters. *Science*, 309(5737): 1036-1039.
- Aldrich, J. H. & F. D. Nelson. (1984). *Linear Probability, Logit, and Probit Models*. Beverly Hills: SAGE.
- Alesch, D. J., J. N. Holly, E. Mittler & R. Nagy. (2001). Organizations at risk: What happens when small businesses and not-for-profits encounter natural disasters. Fairfax, Virginia: Public Entity Risk Institute.
- An, L., M. Linderman, J. Qi, A. Shortridge & J. Liu (2005). Exploring complexity in a human-environment system: An agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers*, 95: 54-79.
- Anwar, M., C. Jeanneret, L. Parrot & D. Marceau (2007). Conceptualization and implementation of a multi-agent model to simulate whale-watching tours in the St. Lawrence Estuary in Quebec, Canada. *Environmental Modelling & Software*, 22(12): 1775-1787.
- Baker, E. J. (1995). Public response to hurricane probability forecasts. *The Professional Geographer*, 47(2): 137-147.
- Batty, M. (2006). *Model Cities*. Liverpool: Centre for Advanced Spatial Analysis, University College London.
- . (2007). *Cities and Complexity Understanding Cities With Cellular Automata, Agent-Based Models, and Fractals*. The MIT Press.
- Batty, M., J. Desyllas & E. Duxbury (2003). The discrete dynamics of small-scale spatial events: Agent-based models of mobility in carnivals and street parades. *International Journal of Geographical Information Science*, 17(7): 673-697.
- Benenson, I. (2009). Status of OBEUS project. Personal communication, February 21st 2009.
- Benenson, I., I. Omer & E. Hatna (2002). Entity-based modeling of urban residential dynamics the case of Yaffo, Tel Aviv. *Environment and Planning B: Planning and Design*, 29(4): 491-512.
- Benenson, I. & P. M. Torrens. (2004). *Geosimulation Automata-Based Modeling of Urban Phenomena*. West Sussex, England: John Wiley&Sons Ltd.
- Billari, F. C., T. Fent, A. Prskawetz & J. Scheffran. (2006). Agent-based computational modelling : An introduction. In *Agent -Based Computational Modelling*, eds. F. C. Billari, T. Fent & J. Scheffran, 1-16. Germany: Springer.

- Campanella, R. (2006). Street survey of business reopenings in post-Katrina New Orleans. *Whitepaper funded by National Science Foundation Award 0554937*.
- . (2008). *Bienville's Dilemma: A Historical Geography of New Orleans*. Lafayette, LA: University of Louisiana Press, Lafayette.
- Chaiken, J. M. (1975). Hypercube queuing model: Executive summary. The New York City Rand Institute.
- Chang, S. (2006). Do disaster areas benefit from disasters? *Growth and Change*, 15(4): 24-31.
- Clarke, K. C. (2008). Mapping and modelling land use change: an application of the SLEUTH Model. In *Landscape Analysis and Visualisation*, eds. C. Pettit, W. Cartwright, I. Bishop, K. Lowell, D. Pullar & D. Duncan, 353-366. Springer Berlin Heidelberg.
- CIMSS (2011). Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison/Space Science and Engineering Center website <http://cimss.ssec.wisc.edu/> (last accessed 25 March 2011).
- Cutter, S. L. & C. T. Emrich (2006). Moral hazard, social catastrophe: The changing face of vulnerability along the hurricane coasts. *The Annals of the American Academy of Political and Social Science*, 604(1): 102-112.
- Cutter, S. L., C. T. Emrich, J. T. Mitchel, B. J. Boruf, M. Gall, M. Schmidlein, C. G. Burton & G. Melton (2006). The long road home: Race, class and recovery from Hurricane Katrina. *Environment*, 48(2): 8-20.
- Cutter, S. L. & C. Finch (2008). Temporal and spatial changes in social vulnerability to natural hazards. *Proceedings of the National Academy of Sciences of the United States of America*, 105(7): 2301-2306.
- Cutter, S. L., J. T. Mitchel & M. S. Scott (2004). Revealing the vulnerability of people and places: A case study of Georgetown county, South Carolina. *Annals of the Association of American Geographers*, 90(4): 713-737.
- Dahlhamer, J.M. & M.J.D. D'Souza (1997). Determinants of business-disaster preparedness in two U.S. metropolitan areas. *International Journal for mass emergencies and Disasters*, 15(2) 265-281.
- Dahlhamer, J. M. & K. J. Tierney (1998). Rebounding from disruptive events: Business recovery following the Northridge Earthquake. *Sociological Spectrum: Mid-South Sociological Association*, 18(2): 121-141.
- Dietzel, C. & K. C. Clarke (2004). Replication of spatio-temporal land use patterns at three levels of aggregation by an urban cellular automata. *Lecture Notes in Computer Science*, 3305: 523-532.

- Dubin, R. (2009). Spatial weights . In *The SAGE Handbook of Spatial Analysis*, ed. A. S. Fotheringham & Rogerson P. A., 125-157. Wiltshire, GB: SAGE Publications Inc.
- Edmonds, B. (1999). Modeling bounded rationality in agent-based simulation using the evolution of mental models. In *Computational Techniques for Modeling Learning in Economics*, ed. B. T., 305-332. Boston: Kluwer Academic Publishers.
- Elliott, J. R. & J. Pais (2006). Race, class, and Hurricane Katrina: Social differences in human responses to disaster. *Social Science Research*, 35(2): 295-321.
- Epstein, J. M. (2006). *Generative social science*. Princeton, New Jersey: Princeton University Press.
- Ewing, B. T., J. Brown Kruse & D. Sutter (2007). Hurricanes and economic research: An introduction to the Hurricane Katrina Symposium. *Southern Economic Journal*, 74(2): 315-325.
- Fiddaman, T. S. (2002). Exploring policy options with a behavioral climate-economy model. *System Dynamics Review*, 18(2): 243-267.
- Frey, W. H. & A. Singer (2006). Katrina and Rita impacts on gulf coast populations: First census findings. Washington: The Brookings Institution Metropolitan Policy Program.
- Frey, W. H., A. Singer & D. Park (2007). Resettling New Orleans: The first full picture from the Census. The Brookings Institution Metropolitan Policy Program.
- Gilbert, N. (2008). *Agent-Based Models*. Los Angeles: SAGE Publications, Inc.
- Gilbert, N. & K. G. Troitzsch (2005). *Simulation for the Social Scientist*. Berkshire, England: Open University Press.
- GNOCDC (2011) Greater New Orleans Community Data Center www.gnocdc.org (last accessed 25 March 2011)
- Goudeau, D. A., and W. C. Conner (1968). Storm surge over the Mississippi river delta accompanying Hurricane Betsy, 1965. *Monthly Weather Review*, 96 (2):118-124.
- Graumann, A., T. Houston, J. Lawrimore, D. Levinson, N. Lott, S. McCown, S. Stephens, D. Wuertz (2005). Hurricane Katrina a climatological perspective. Asheville, North Carolina: NOAA's National Climatic Data Center.
- Gutmann, M. P. & V. Field (2010). Katrina in historical context: environment and migration in the U.S. *Population and Environment*, 31: 3-19.
- Heath, B., R. Hill & F. Ciarallo (2009). A Survey of agent-based modeling practices (January 1998 to July 2008). *Journal of Artificial Societies and Social Simulation*, 12(4) 9.

- Horstmann, C. (1999). *Computing Concepts With Java2 Essentials*. New York: John Wiley & Sons, Inc.
- Kates, R. W., C. E. Colten, S. Laska & S. P. Leatherman (2006). Reconstruction of New Orleans after Hurricane Katrina: A research perspective. *Proceedings of the National Academy of Sciences*, 103(40): 14653-14660.
- Kelman, A. (2003). *A River and Its City: The Nature of Landscape in New Orleans*. Berkeley: University of California Press.
- Kroll, C. A., J. D. Landis, Q. Shen & S. Stryker. (1991). Economic Impacts of the Loma Prieta Earthquake: A Focus on Small Business. The University of California, Transportation Center.
- L.S.U. (2005). Hurricane Katrina & Rita Clearinghouse Cooperative. Baton Rouge. <http://katrina.lsu.edu> (last accessed 5 February 2011).
- Lam, N., K. Pace, R. Campanella & J. LeSage (2007). Collaborative research:(DRU) Modeling business return amid post-disaster uncertainty: New Orleans after Katrina. Louisiana State University, Tulane University, Texas State University. National Science Foundation Research Proposal.
- Lam, N., K. Pace, R. Campanella, J. LeSage & H. Arenas (2009). Business return in New Orleans: Decision making amid post-Katrina uncertainty. *PLoS ONE*, 4 (8).
- Landry, C. E., O. Bin, P. Hindsley, J. C. Whitehead & K. Wilson (2007). Going home: Evacuation-migration decisions of Hurricane Katrina survivors. *Southern Economic Journal*, 74(2): 326-343.
- Larson, R. C. (1975). Hypercube queuing model: program description. The New York city Rand Institute.
- LeSage, J. & K. Pace (2009). *Introduction to spatial econometrics*. Boca Raton, FL: CRC Press Taylor & Francis Group.
- LeSage, J., K. Pace, N. Lam, R. Campanella & X. Liu (2010). New Orleans business recovery in the aftermath of Hurricane Katrina.
- Liang, W. (2009). An Integrated agent-based microsimulation model for hurricane evacuation in New Orleans. Ph.D. Dissertation. Louisiana State University.
- Liao, T. F. (1994). *Interpreting Probability Models Logit, Probit and Other Generalized Linear Models*. Thousand Oaks, CA: SAGE Publications, Inc.
- Longley, P. A., M. F. Goodchild, D. J. Maguire & D. W. Rhind (2005). *Geographic Information Systems and Science*. West Sussex, England: John Wiley & Sons Ltd..
- Malthus, T. R. (1798). *An essay on the principle of population*. London.

- Masozera, M., M. Bailey & C. Kerchner (2007). Distribution of impacts of natural disasters across income groups: A case study of New Orleans. *Ecological Economics*, 63(2-3): 299-306.
- Miller, J. H. & S. E. Page (2007). *Complex Adaptive Systems*. Princeton, New Jersey: Princeton University Press.
- MSNBC (2011). No quick fix for New Orleans breached levees
http://www.msnbc.msn.com/id/9130254/ns/us_news-katrina_the_long_road_back/t/no-quick-fix-new-orleans-breached-levees/ (last accessed 25 March 2011)
- NASA (2007). Global Change Master Directory website.
<http://gcmd.nasa.gov/records/SLEUTH.html> (last accessed 19 March 2011).
- North, M. J. & C. M. Macal (2007). *Managing Business Complexity Discovering Strategic Solutions With Agent-Based Modeling and Simulation*. New York ,NY: Oxford University Press.
- Norusis, M. J. (2003). *SPSS 12.0 Statistical Procedures Companion*. Harlow, GB: Prentice Hall.
- O'Sullivan, D. (2008). Geographical information science: Agent-based models. *Progress in Human Geography*, 32(4): 541-550.
- O'Sullivan, D. & P. M. Torrens (2000). Cellular models of urban systems. In *CASA Working Papers 22*. London: Centre for Advanced Spatial Analysis (UCL)
- Oh, J., J. Hwang & S. F. Smith (2006). Agent Technologies for Post-disaster Urban Planning. In *Proceedings of the First International Workshop on Agent Technology for Disaster Management* ,Hakodate, Japan 2006, 24-31.
- Omer, I. (1999). Demographic processes and ethnic residential segregation. *Discrete Dynamics in Natura and Society*, 3: 117-184.
- OSGeo. (2011). Open Source Geospatial Foundation website. <http://www.osgeo.org/> (last accessed 19 March 2011).
- Palm, R. (1995). Catastrophic earthquake insurance: Patterns of adoption. *Economic Geography*, 71: 119-131.
- Palm, R. & M. E. Hodgson (1993). Natural hazards in Puerto Rico. *The Geographical Review*, 83(3): 280-289.
- Peng, J. C.-Y. & T.-S. H. So (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics*, 1(1): 31-70.

- Pielke, R. A. (1997). *Hurricanes: Their Nature and Impacts on Society*. West Sussex, England: Wiley.
- Pielke, R. A. & C. W. Landsea (1998). Normalized hurricane damages in the United States 1925 -95. *Weather and Forecasting*, 13: 621-631.
- Plyer, A., J. Bonaguro & K. Hodges (2010). Using administrative data to estimate population displacement and resettlement following a catastrophic U.S. disaster. *Population and Environment*, 31(1-3): 150-175.
- Plyer, A. & K. Hodges (2008). Using U.S. postal service delivery statistics to track population shifts following a major U.S. disaster. In *Population Association of America 2008 Annual Meeting*. New Orleans, U.S.
- Port of New Orleans (2011). Port of New Orleans website, <http://www.portno.com> (last accessed 25 March 2011).
- Ramsey, P. (2011). Introduction to Postgis. Victoria-British Columbia: Refractions Research.
- Rappaport, J. & J. D. Sachs (2003). The United States as a Coastal Nation. *Journal of Economic Growth*, 8(10): 5-46.
- Sargent, R. G. (1998). Verification and Validation of Simulation Models. In *Proceedings of the 30th conference on Winter Simulation*, eds. D.J. Medeiros, Edward F. Watson, John S. Carson and Mani S. Manivannan. IEEE Computer Society Press, Los Alamitos, CA, USA 121-130, Washington D.C. U.S.
- Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1: 143-186.
- Schreinemachers, P. (2006). The (Ir)relevance of the crop yield gap concept to food security in developing countries. Doctoral dissertation, Rheinischen Friedrich-Wilhelms-Universität. Bonn, Germany.
- Schreinemachers, P. & T. Berger (2006). Land-use decision in developing countries and their representation in multi-agent systems. *Journal of Land Use Science*, 1(1): 29-44.
- Sharif, A. A., A. Zaharim & K. Sopian (2009). The comparison logit and probit regression analyses in estimating the strenght of gear teeth. *European Journal of Scientific Research*, 27(4): 548-553.
- Simon, H. (1997). *Models of Bounded Rationality Vol.3: Empirically Grounded Economic Reason*. Massachusetts: MIT Press.
- Skidmore, M. & H. Toya (2002). Do natural disasters promote long-run growth? *Economic Enquiry*, 40(4): 664-687.

- Spicer, J. (2005). *Making Sense Of Multivariate Data Analysis*. Thousand Oaks, CA: Sage Publications.
- Stewart, T. R. (2002). Extreme Event Decision Making Workshop report. Center for Policy Research, University of Albany. Albany, NY.
- Stringfield, J. D. (2010). Higher ground: an exploratory analysis of characteristics affecting returning populations after Hurricane Katrina. *Population and Environment*, 31(1-3): 43-63.
- The Brookings Institute (2008). The New Orleans Index: Three Years After Katrina.
- Torrens, P. M. (2000). How cellular models of urban systems work. London: Centre for Advanced Spatial Analysis, University College London.
- . (2001). A hybrid geocomputation model for operational land-use and transport. In *Association of American Geographers Annual Meeting*, New York, NY.
- Torrens, P. M. & I. Benenson (2005). Geographic automata systems. *International Journal of Geographical Information Science*, 19(4): 385-412.
- Turner, B. L., R. E. Kasperson, P. A. Matson, J. J. McCarthy, R. W. Corel, L. Christensen, N. Eckley, J. X. Kasperson, A. Luers, M. L. Martello, C. Polsky, A. Pulsipher & A. Shiller (2003). A Framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences of the United States of America*, 100(14): 8074-8079.
- Vale, L. J. & T. J. Campanella. (2005). *The Resilient City How Modern Cities Recover From Disaster*. New York, NY: Oxford University Press.
- Ventana Systems inc. (2010). Vensim website <http://www.vensim.com/> (last accessed 25 March 2011).
- Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118(2972): 558-560.
- Waugh, W. L. & R. B. Smith (2006). Economic development and reconstruction on the Gulf after Katrina. *Economic Development Quarterly*, 20(3):211-218.
- Webb, G. R., K. J. Tierney, and J. M. Dahlhamer (2002). Predicting long-term business recovery from disaster: a comparison of the Loma Prieta earthquake and Hurricane Andrew. *Environmental Hazards*, 4(2-3):45-58.
- Webb, G. F. (2008). Population models structured by age, size and spatial position. In *Structured Population Models in Biology and Epidemiology*, eds. P. Magal & S. Ruan, 1-49. Springer-Verlag Berlin Heidelberg.

Worsley, J. C. & J. D. Drake (2002). *Practical PostgreSQL*. Sebastopol, CA: O'Reilly & Associates, Inc.

Zhang, Y., M. K. Lindell & C. S. Prater (2009). Vulnerability of community businesses to environmental disasters. *Disasters*, 33(1): 38-57.

Vita

Helbert was born in Peru, his field of interest is the use of Geographic Information Systems to solve urban problems. He got his bachelor degree in architecture from San Agustin National University in Arequipa, Peru in 2000. Shortly after his graduation he started working in Organismo de Formalización de la Propiedad Informal a country wide urban cadaster project. It is in this project that he had his first contact with Geographic Information Systems. He completed the second speciality in Systems Engineering in San Agustin University in 2002. The same year he got a scholarship from the Netherlands Fellowship Program to study in the master of science program in Geoinformatics in the International Institute for Geo-Information Science and Earth Observation, in Enschede, The Netherlands. After obtaining his Masters degree he returned to Peru where he worked in urban vulnerability assessment projects for Peruvian cities. In 2006 he enrolled in the doctoral program in geography in LSU. In the year 2010 he got the Doctoral Dissertation Research Improvement Grant from the National Science Foundation.