PREDICTING THE TEXAS WINDSTORM INSURANCE ASSOCIATION PAYOUT FOR COMMERCIAL PROPERTY LOSS DUE TO IKE BASED ON WEATHER, GEOGRAPHICAL, AND BUILDING VARIABLES

A Thesis

by

KEHUI ZHU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Paul Woods
Committee Members,	Ifte Choudhury
	Faming Liang
Head of Department,	Joe Horlen

May 2013

Major Subject: Construction Management

ABSTRACT

Hurricanes cause enormous loss to life and property worldwide. Predicting the damage caused by hurricane and figuring out what factors are responsible for the damage are important. This study utilizes multiple linear regression models to predict a hurricane – induced Texas Windstorm Insurance Association (TWIA) payout or TWIA payout ratio using independent variables that could affect the hurricane intensity, including distance from the coastline, distance from the hurricane track, distance from the landfall center of Hurricane Ike, proportion in floodplain zone (100 year, 500 year, 100-500 year), building area, proportion in island, number of buildings per parcel, and building age.

The methodology of this study includes Pearson's correlation and multiple linear regressions. First, Pearson's correlation is used to examine whether there are any significant correlations between the dependent and independent variables. For TWIA payout, three independent variables, distance from the coastline, distance from the landfall center, and building area, are correlated to the TWIA payout at the 0.01 level. Distance from the coastline and distance from the landfall center have negative relations with the TWIA payout. The variable, building area, has a positive relation with the TWIA payout. Moreover, the improvement value is correlated to the TWIA payout at the 0.05 level. For TWIA payout ratio, distance from the coastline is correlated to the TWIA payout ratio at the level of 0.01 and distance from the landfall center is correlated to the

TWIA payout ratio at the 0.05 level. These two variables have negative relations to the TWIA payout ratio.

Multiple linear regressions are applied to predict the TWIA payout and payout ratio. A regression model with an Adjusted R Square of 0.264 is presented to predict the TWIA payout. This model could explain 26.4 percent of the variability in TWIA payout using the variables, distance from coastline and building area. A regression model with an Adjusted R Square of 0.121 is presented to predict the TWIA payout ratio.

DEDICATION

To Dr. Paul K. Woods

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Woods. His professionalism in academic and research has broadened my knowledge in construction management and facility management. I would also like to express my gratitude to my committee members, Dr. Choudhury and Dr. Liang, for their guidance and support throughout the course of this research. Thanks also go to Dr. Nicolas, who gave me lots of insights on the data interpolation.

Thanks to my friend and colleague, Ji-Myong Kim who helped and supported me during the study. I also want to extend my gratitude to the Texas Windstorm Insurance Association and Galveston Appraisal District for data provided for this research. Finally, thanks to my mother and father for their encouragement and love.

NOMENCLATURE

TWIA	Texas Windstorm Insurance Association
IMP	Improvement Value
GIS	Geographic Information System
HGAC	Houston – Galveston Area Council
FEMA	Federal Emergency Management Agency
VIF	Variance Inflation Factors
ANOVA	Analysis of Variance

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Topic Introduction

Hurricanes cause enormous loss to life and property worldwide. Within United States, hurricane is one of the costliest natural hazards (Landsea et al. 1999), and the property damages caused by hurricanes have been rising rapidly over the years. For instance, in late October 2012, Hurricane Sandy devastated portions of the Caribbean, Mid – Atlantic and Northeastern United States. It caused an estimated damage of at least $20 billion.

Because of the huge hurricane-induced damage, there have been significant improvements in predicting, tracking, and warning the public of tropical storm events in recent decades (Burton. 2010). Nevertheless, relatively little progress has been made to predict hurricane-induced damage. The proposed research uses the Hurricane Ike as a case study to predict the TWIA payout and TWIA payout ratio. On September 13th 2008, Hurricane Ike made landfall near Galveston, Texas as a Category 2 storm. It resulted in extensive damages, including an estimated 74 deaths statewide and extensive loss in many counties (Zane et al., 2010). Total financial damage from Ike in Texas, Louisiana, and Arkansas is estimated at $29.5 billion dollars – third behind Hurricanes Katrina and

Sandy, respectively. In this research, the TWIA payout and TWIA payout ratio are used to represent the damage caused by Hurricane Ike. Galveston County is the study area.

Research Objective

The purpose of this research is to predict a hurricane-induced TWIA payout or TWIA payout ratio using independent variables that could affect the hurricane intensity. Those independent variables include the distance from the coastline (m), the distance from the hurricane track (m), the distance from the landfall center of hurricane (m), the building area (sq.m), the age of the building (yr), proportion in floodplain zone (A, X, X500), proportion in island (0/1; 1 – on island, 0 – not on island), and the number of buildings per parcel (0/1; 1 – single building, 0 – multiple buildings). In this research, Hurricane IKE is used as an example as the nature of hurricanes remains same. Galveston County is the study area.

As shown in Equation 1, the TWIA payout ratio is TWIA payout ($) divided by the appraised value of improvement ($). The improvement value is the value of the property.

TWIA payout ratio = TWIA payout / 2008 improvement value (1)

Hypothesis

The hypotheses include two regression models. One model is to predict the TWIA payout, and the other is to predict the TWIA payout ratio. For parcels where there is only one property, another model with building age is tested as well.

TWIA Payout Model

The TWIA payout can be predicted using the following multiple linear regression model.

$$Y = \beta 0 + \beta 1x1 + \beta 2x2 + \beta 3x3 + \beta 4x4 + \beta 5x5 + \beta 6x6 + \beta 7x7 + \beta 8x8 + \beta 9x9 + \varepsilon$$

- Y: TWIA Payout
- X1: distance from the hurricane track (m)
- X2: distance from the coastline (m)
- X3: distance from the landfall center of Hurricane (m)
- Categorical variable: location of property within Federal Emergency Management Agency (FEMA) floodplain zone (A, X, X500).
  - X4: located in the Floodplain zone A (0/1)
  - X5: located in the Floodplain zone X500 (0/1)
- X6: building area (sq.m)
- X7: number of buildings per parcel (0/1)

- X8: location of property with regards to Galveston Island (0/1)
- X9: improvement value of the building
- ε: random error

For the categorical variable – FEMA floodplain zone, there are three values: A, X, and X500. These values have the following definitions. Only two values are included in the regression model.

1. Floodplain zone A: areas within the 100-year flood with a 1% annual probability of flooding.
2. Floodplain zone X: areas outside the 500-year flood plain with less than .2% annual probability of flooding.
3. Floodplain zone X500: areas between the 100-year and the 500-year floodplain.

For parcels where there is only one property, another model with building age is tested as well. The TWIA payout can be predicted using the following multiple linear regression model which includes the age of the building as an independent variable.

$$Y= \beta 0 + \beta 1x1 + \beta 2x2 + \beta 3x3 + \beta 4x4 + \beta 5x5 + \beta 6x6 + \beta 7x7 + \beta 8x8 + \beta 9x9 + \varepsilon$$

- Y: TWIA Payout
- X1: distance from the hurricane track (m)
- X2: distance from the coastline (m)

- X3: distance from the landfall center of Hurricane (m)
- Categorical variable: location of property within FEMA floodplain zone (A, X, X500).
  - X4: located in the Floodplain zone A (0/1)
  - X5: located in the Floodplain zone X500 (0/1)
- X6: building area (sq.m)
- X7: location of property with regards to Galveston Island (0/1)
- X8: improvement value of the building
- X9: building age (yr)
- ε: random error

TWIA Payout Ratio Model

The TWIA payout ratio can be predicted using the following multiple linear regression model.

$$Y = \beta0 + \beta1x1 + \beta2x2 + \beta3x3 + \beta4x4 + \beta5x5 + \beta6x6 + \beta7x7 + \beta8x8 + \varepsilon$$

- Y: TWIA Payout Ratio
- X1: distance from the hurricane track (m)
- X2: distance from the coastline (m)

- X3: distance from the landfall center of Hurricane (m)
- Categorical variable: location of property within FEMA floodplain zone (A, X, X500).
    - X4: located in the Floodplain zone A (0/1)
    - X5: located in the Floodplain zone X500 (0/1)
- X6: building area (sq.m)
- X7: No. of buildings per parcel (0/1)
- X8: location of property with regards to Galveston Island (0/1)
- ε: random error

For parcels where there is only one property, another model with building age is tested as well. The TWIA payout ratio can be predicted using the following multiple linear regression model which includes the age of the building as an independent variable.

$$Y= \beta 0 + \beta 1x1 + \beta 2x2 + \beta 3x3 + \beta 4x4 + \beta 5x5 + \beta 6x6 + \beta 7x7+ \beta 8x8 + \varepsilon$$

- Y: TWIA Payout Ratio
- X1: distance from the hurricane track (m)
- X2: distance from the coastline (m)
- X3: distance from the landfall center of Hurricane (m)
- Categorical variable: location of property within FEMA floodplain zone (A, X, X500).

    - X4: located in the Floodplain zone A (0/1)
    - X5: located in the Floodplain zone X500 (0/1)
- X6: building area (sq.m)
- X7: location of property with regards to Galveston Island (0/1)
- X8: building age (yr)
- ε: random error

To achieve the objective of this study, the following research hypotheses are going to be tested.

1. The TWIA payout or TWIA payout ratio increases as the distance from the hurricane track decreases.
2. The TWIA payout or TWIA payout ratio increases as the distance from the coastline decreases.
3. The TWIA payout or TWIA payout ratio increases as the distance from the landfall center of hurricane decreases.
4. Property in Floodplain zone A has the highest TWIA payout or payout ratio compared to property in Floodplain zone X500 and X.
5. The TWIA payout or payout ratio increases as the age of the building increases.
6. The TWIA payout increases as the improvement value of the building increases.
7. The TWIA payout decreases as the building age increases.

8. Property that is on the Galveston Island tends to have more TWIA payout or TWIA payout ratio.
9. Property that has multiple buildings per parcel tends to have more TWIA payout than the parcel that has a single building.

Importance and Expected Benefits

Although hurricanes cause enormous loss to life and property, prediction of hurricane – induced damage is limited. Moreover, windstorm-damage based on statistical analysis is also limited due to the dearth of data (Friedman, 1984). Most researches use damage rating system that is created by the professional judgment. However, this damage rating system is relatively subjective and thus it might not be a good indicator of the hurricane-induced damage.

In this research, the author is able to obtain a large quantity of TWIA payout data for commercial buildings. For instance, the population of payout data in Galveston County is about 1,800. Using the real insurance payout data could make the damage model more objective. Therefore, this insurance payout model would be more objective in predict the hurricane – induced damage than the previous models.

Second, most previous studies focused on the overall buildings rather than a specific type of building. Many studies just investigated the residential properties. Research is very limited in a specific type of buildings, commercial buildings. As this research uses data of commercial building, it would bring more insights not only for researchers, but also for business persons.

The model uses GIS process to collect and process the data. The whole procedure is highly repeatable. If the model is validated, the hurricane track could be altered to areas with more population such as Houston and New York to predict future damage based on these independent variables. Finally, this model would also help local governments to develop policies and plans for hurricane mitigation.

Assumptions

All data collected from public sources and used in this research are accurate and reliable.

Appraised improvement values reasonably reflect actual value of the properties.

Definitions

Parcel: A fundamental cadastral unit. It is a piece of land which can be owned, sold, and developed.

Appraised Improvement Value: The appraised value that is assigned to a structure or building by the Appraisal District Office in US dollars.

Texas Windstorm Insurance Association: The Texas Windstorm Insurance Association was established by legislative mandate to provide wind and hail insurance for Texas Gulf Coast property owners in the event of catastrophic loss.

FEMA Floodplain Zone A: Areas within the 100-year flood with a 1% annual probability of flooding.

FEMA Floodplain Zone X: Areas outside the 500-year flood plain with less than .2% annual probability of flooding.

FEMA Floodplain Zone X500: Areas between the 100-year and the 500-year floodplain.

Geographic Information System: GIS is a system designed to capture, store, manipulate, analyze, and present all types of geographical data.

ArcMap: ArcMap is the main component of Esri's ArcGIS suite of geospatial processing program.

CHAPTER II

LITERATURE REVIEW

Case 1 - Determinants & Characteristics of Damage in Single-Family Island Households from Hurricane Ike

Basic Information

Highfield, W. E., Peacock, W. G., & Van Zandt, S. (2010). Determinants & Characteristics of Damage in Single-Family Island Households from Hurricane Ike. *Association of Collegiate Schools of Planning Conference, Minneapolis, MN*

Authors of this article are Highfield, W. E., Peacock, W. G., and Van Zandt, S. Study date is 2010. Location of this study is Galveston Island and Bolivar Peninsula, which are two major islands that are damaged by Hurricane Ike.

Data

Population of Interest

The population of interest in this stud is detached housing units. However, the number of population is unknown.

Sample

The sample consists of approximately 1,500 detached housing units randomly sampled from parcels on Galveston Island and Bolivar Peninsula. Figure 1 shows the Galveston Island and Bolivar Peninsula study area.

Figure 1. Galveston Island and Bolivar Peninsula Study Area

Variables

Dependent Variables

The dependent variable is called damage index. Four variables including *Foundation and Structural Damage, Roof Damage, Exterior Damage, and Overall Damage* are assessed using four separate five-point scale. Then, these four values are added up to create the damage index.

Independent Variables

Independent variables include two categories: geographic variable and social vulnerability variable. Geographic variables include structure elevation, home age, distance to water, distance to seawall, proportion in FEMA A Zone, proportion in FEMA V Zone, maximum inundation, and Galveston Island. Social vulnerability variables include proportion of Hispanic, proportion of Black, and assessed value of the home. FEMA A Zone is the 1% flood zone. FEMA V Zone is 1% flood zone with velocity or wave action, which represents both flood hazards and wave hazards.

Observational Unit

The observational unit is a detached housing unit.

Method

Data were analyzed in three phases.

1. First, two blocks of independent variables representing geographic and structural characteristics were incrementally loaded into Ordinary Least Squares regression models to explain damage measured by the damage index.
2. Second, the authors explored and modeled the non-linear effects of the age of the home using squared and cubed terms in a third regression block.
3. Finally, the authors added a fourth regression block consisting of social vulnerability variables. OLS models were run using robust standard errors to offset heteroskedastic error structures.

Results

The findings are as follows:

1. Structure elevation is significant (one-tailed) in the final model. It has a negative relationship to the damage index.

2. Home age is not a simply linear indicator of the hurricane damage. The entrance of quadratic and cubic terms reveals that there is a non-linear relationship between home and damage caused by hurricane. In the final model, damage increases with home age until the first point of inflection at 49.9 years. Damage begins in decrease with home age until reaching the second point of inflection at 97.1 years.
3. Distance to seawall has a negative relationship to the damage index, which suggests that the seawall in Galveston Island is an effective protection.
4. Proportion in FEMA A Zone and proportion in FEMA V Zone are positive indicators. More damage occurred in V Zones compared to A Zones.
5. Maximum inundation is a significant positive predictor of damage.
6. Galveston Island is the strongest predictor of damage. It demonstrates the difference in mean damage between the two locations.
7. Proportion Hispanic and proportion Black are positive indicators. Moreover, the proportion White is negative and statistically significant. The phenomena indicate that minorities are impacted with higher levels of damage.
8. Assessed value of the home is a significant negative predictor.

Case 2 -A Quantitative Method for Estimating Probable Public Costs of Hurricanes

Basic Information

Boswell, M. R., Deyle, R. E., Smith, R. A., & Baker, E. J. (1999). A Quantitative Method for Estimating Probable Public Costs of Hurricanes. *Environment Management*, 23(3), 359-372.

The study date is 1999. Location of this study is Lee County, Florida, USA. The study describes a method for estimating public costs resulting from damage caused by hurricanes and applies the method to a specific local jurisdiction – Lee County, Florida, USA. The method employs a multivariate model developed through multiple regression analysis of an array of variables that measure meteorological, socioeconomic and physical conditions related to the landfall of hurricanes.

Data

Population of Interest

Population of interest is public expenditure.

Sample

Samples include 250 public expenditures of five presidentially declared disasters that occurred in Florida between 1979 and 1995. The data contain a detailed description of applicant jurisdiction, expenditure amounts by expenditure category, damage location, damage facility and a narrative description of the damage.

Variables

Dependent Variable

The dependent variable is total costs for local governments to respond and recover hurricane. Seven federally defined expenditure categories are taken into consideration. These seven categories include:

- Debris
- Protective measures
- Roads, signs, and bridges
- Water control facilities
- Buildings and equipment
- Public utilities
- Parks and recreation, and other

Independent Variables

Independent variables included four categories: tropical cyclone, socioeconomic, development, and physical variables.

Tropical cyclone variables measure the meteorological characteristics of the storm. The tropical cyclone variables include maximum sustained surface wind speed at jurisdiction (miles per hour), forward speed of tropical cyclone (miles per hour), quadrant of on-shore winds (0/1 dichotomous), tropical cyclone angle of approach (degree), entering tropical cyclone (0/1 dichotomous), tropical cyclone surge (0/1 dichotomous), and tropical cyclone landfall (0/1 dichotomous).

Socioeconomic variables measure a set of population and housing value characteristics for community. Socioeconomic variables include population of jurisdiction (persons), population of jurisdiction at risk to storm surge (persons), and median housing unit value (dollars).

Development variables characterize land development of the coastal area of a community. These variables include beachfront low/medium density residential existing land use (linear miles), beachfront high density residential existing land use (linear miles), beachfront commercial existing land use (linear miles), beachfront recreation/conservation existing land use (linear miles) and beachfront vacant existing land use (linear miles).

Physical variables measure the geographic characteristics of community. The physical variables include land area of jurisdiction 1990 (square miles), beachfront length (linear miles), storm wave susceptibility quotient at high tide (percent>moderate), beachfront jurisdiction (0/1 dichotomous), and waterfront jurisdiction (0/1 dichotomous).

Method

Regression analysis is used to develop and test numerous multivariate models for each of the seven expenditure categories of public assistance individually and in various combinations.

Results

The independent variables maximum sustained surface wind speed and populations of jurisdiction consistently meet the t test criterion in all model specifications. The log-log model above explains 74% of the variance in the expenditure data.

The equations for each regression model are specified as follows:

- Linear: $Y=-16547009+183918(WIND)+74.56(POP\_TOT)-211.88(POP\_DENS)$
- Log-log (base 10): $\log Y=-7.77+4.98(\log WIND)+0.90(\log POP\_TOT)$
- Poly A: $Y=-22042821+316.55(WIND)^2+6.58(POP\_TOT)+9.27E\text{-}10(POP\_TOT)$
- Poly B: $Y=-3075930+404.44(WIND)^2+3.13E\text{-}4(POP\_TOT)^3$
- Poly C: $Y=-976532+2.22(WIND)^3+9.81E\text{-}10(POP\_TOT)^3$

CHAPTER III

DATA COLLECTION

Research Data

Population of Interest

The population of interest is all qualified commercial buildings in Galveston County. Here are the criteria of a qualified commercial building.

Criteria for the Admissibility of Data

- The building is among the payout data of Hurricane Ike from TWIA.
- The building is within the Galveston County.
- The building is within the wind field of Hurricane Ike.
- The building could be found in the Galveston Appraisal District and has the 2008 improvement value.
- The building has the record of building area.
- The building has the record of building age.

1,800 data of Galveston County are obtained from TWIA. Because of the criteria of selecting qualified commercial buildings, the population size would be less than 1,800.

Sample Selection

Data is selected using random sampling method. Microsoft Excel spreadsheet is used to generate a random sample. The sample size is 212, which is workable.

Observational Unit

The observational unit is the building in a qualified parcel.

Variables

Dependent Variables

The research uses two dependent variables in different regression models. The first dependent variable is TWIA payout ($). The second dependent variable is TWIA payout ratio, which is TWIA payout ($) divided by appraised value of improvement ($).

Independent variables

Independent variables include distance from the hurricane track (m), distance from the coastline (m), distance from the landfall center of hurricane (m), proportion in floodplain zone (A, X, X500), building area (sq.m), number of buildings per parcel (0/1), Galveston Island (0/1), age of the building* (yr).

The age of the building is collected only when there is only one building per parcel. When there are multiple buildings per parcel, it is hard to adjust multiple different ages of the properties. Therefore, the variable age of the building is included in the regression model when there is only one building per parcel.

Admissibility

Incomplete records will not be accepted.

Data Collection

Data such as TWIA payout, improvement value of the year 2008, the distance from the coastline, the distance from hurricane track, the distance from the landfall center of Hurricane Ike, floodplain zone, and year built for a qualified building are need to be collected to test the hypothesis. The following Table 1 illustrates the nature and source of these data. Table 2 is the data collection form. Incomplete records would not be accepted.

TWIA payout data are obtained via a public data request from the Texas Windstorm Insurance Association. The data provided by TWIA include the address of the commercial building, the date of claim paid, and the amount of claim paid. The parcel information, including improvement value, the year built of the building, building area is obtained from the Galveston Appraisal District. The original unit of measure for building area is square feet and it is converted to square meter.

A shapefile of 2008 Galveston County Parcel is collected and the shapefile illustrates the geographical locations of parcels in the Galveston County. FEMA Floodplain zone and Hurricane Ike track area also in the form of GIS shapefile. The shapefile of floodplain zones is obtained from website of Federal Emergency Management Agency. The shapefile of hurricane track is obtained from Houston-Galveston Area Council. The shapefile of coastline is digitized by the author in ArcMap. After collecting various shapefiles of Galveston County parcel, Hurricane Ike track, and coastline, the landfall center of Hurricane Ike is also digitized using ArcMap. After collecting all the data, geographical data, such as distance from the hurricane track, coastline, and landfall center are measured in the ArcMap. The units of measures for these variables are meter.

Table 1. Data Format and Source

Data	Nature	Source
TWIA payout data	• Address of the commercial building • Date of the claim paid • Amount of the claim paid ($)	TWIA (Public Data Request)
Parcel data	**Shapefile** • Appraised value of improvement of 2008 ($) • Age of the building (yr) • Building size (sqft)	Galveston County Appraisal District http://www.galvestoncad.org/Appraisal/PublicAccess/
Floodplain data	**Shapefile** • Floodplain Zone A • Floodplain Zone X500 • Floodplain Zone X	Federal Emergency Management Agency (FEMA) http://www.tnris.org/get-data#flood
Hurricane IKE track	• Shapefile	Houston-Galveston Area Council (HGAC) http://www.h-gac.com/rds/gis_data/clearinghouse/default.aspx
Coastline	• Shapefile	Digitized in ArcMap

Table 2. Data Collection Form

No.	Property Address	Date of the claim paid (DD/MM/YY)	Amount of the claim paid ($)	Appraised value of improvement($)	TWIA payout ratio	Age of the building* (yr)
1						
2						
3						

Building area (sq.m)	No. of buildings per parcel (0/1)	Located in what floodplain zone (A, X, X500)	Distance from the coastline (m)	Distance from the Hurricane track (m)	Distance from the landfall center of Hurricane (m)

Sample Selection Process

First, 212 qualified TWIA payout data are randomly selected. Second, these data are geocoded according to their addresses in ArcMap (Figure 2). The red dots are 212 qualified buildings. The light purple polygon is the parcel data of Galveston County. The blue line represents the coastline that is digitized by the author. The green line represents the track of Hurricane Ike. The purple point is the landfall center of Hurricane Ike. Figure 3 illustrates the different floodplain zones. Floodplain zone A, X, and X500 are displayed in different colors. The green polygon is the sea.

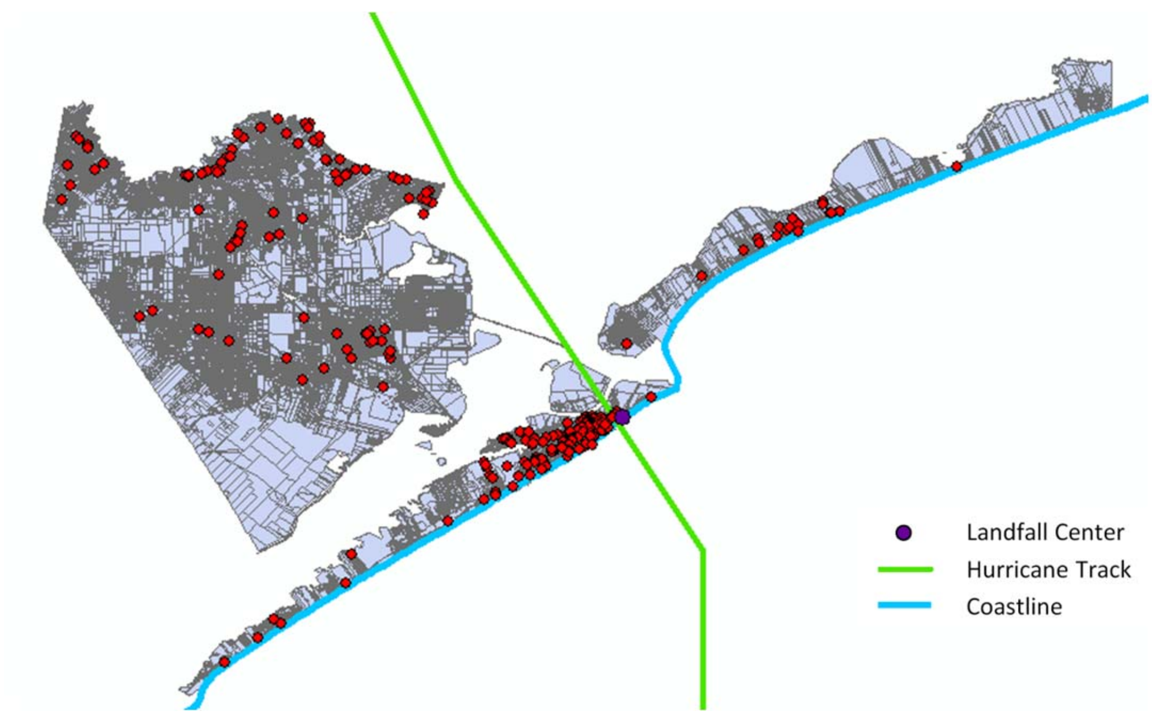

Figure 2. Qualified Buildings in Galveston County

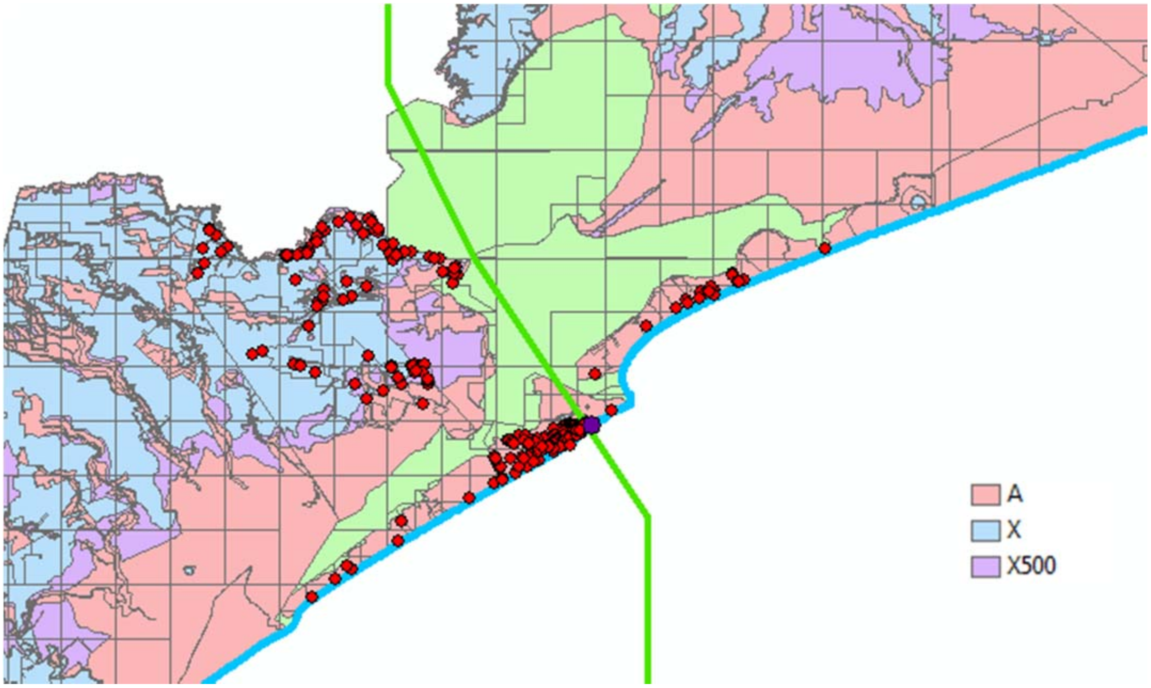

Figure 3. Floodplain Zone

Other variables associated with the property such as the distance from the hurricane track, the distance from the coastline, the distance from the landfall center and the location of property with regards to floodplain are measured in the ArchMap. The improvement value of year 2008, age of building, and building size are obtained from the Galveston Appraisal District (Figure 4, Figure 5). The improvement value includes Improvement HS and Improvement NHS. In other words, the Improvement HS and Improvement NHS are added together to determine the improvement value of the parcel. After collecting all the data, an Excel spreadsheet is created as shown in Table 3.

Improvements					
Type	Description	Area	Year Built	Eff Year	Value
C	Commercial				
MA	Main Area	1634	1945	1962	
OP	Open Porch	72	1945	1962	

Figure 4. Parcel Information (Age & Building Size)

Assessment History (R100540)

Situs Address
4107 AVE L
GALVESTON, TX 77550

	2010	2009	2008
Improvement HS	$0	$0	$0
Improvement NHS	$58,820	$108,930	$102,140
Land HS	$0	$0	$0
Land NHS	$16,070	$16,070	$16,070
Agricultural Mkt	$0	$0	$0
Agricultural Use	$0	$0	$0
Timber Market	$0	$0	$0
Timber Use	$0	$0	$0
Market Value	$74,890	$125,000	$118,210
Homestead Limit			
Assessed	$74,890	$125,000	$118,210
Exemptions	EX	EX	EX

Figure 5. Parcel Information (2008 Improvement Value)

Table 3. Example of Data

ID	08 IMP	Payout	Payout Ratio	Floodplain	Distance coastline (m)	Distance track (m)	Distance landfall (m)	Building Area (sqm)	No. of Building	Building Age (yr)	Island
R100643	43,210	4,683	0.11	A	443	906	1,433	588	1	33	1
R101166	148,130	8,484	0.06	A	628	1,095	1,639	4,131	1	48	1
R101178	138,180	38,582	0.28	A	682	1,214	1,762	3,841	1	106	1
R101215	224,930	1,971	0.01	A	805	1,492	2,047	6,055	1	43	1
R101383	106,930	1,171	0.01	A	1,464	3,078	3,678	9,785	1	44	1
R101416	73,760	38,840	0.53	A	1,678	3,676	4,292	1,650	1	80	1
R101716	75,630	13,231	0.17	A	1,236	2,257	2,858	1,704	1	48	1
R101732	404,830	117,121	0.29	A	1,410	2,648	3,264	16,707	1	83	1
R101885	91,550	112,587	1.23	A	1,798	3,762	4,400	3,260	1	38	1
R101975	27,950	50,765	1.82	A	2,102	4,653	5,316	2,250	0	N/A	1
R101998	172,990	91,832	0.53	A	2,388	5,538	6,234	38,567	1	51	1
R102186	261,800	14,405	0.06	A	958	1,323	1,928	5,092	1	23	1

CHAPTER IV

DATA ANALYSIS AND INTERPRETATION

Data Analysis Methodology

Data will be analyzed using statistical methodologies. These methodologies include descriptive statistics, scatterplots, correlation analysis, ANOVA, and multiple linear regression models. Figure 6 illustrates the procedures of analysis.

Descriptive statistics: The descriptive statistic provides summaries of the samples data and it is the basis of the quantitative analysis. The descriptive statistic includes mean, median, standard deviation, and variance.

Scatterplots: Scatterplots are conducted between dependent and independent variables. These scatterplots show relationships among variables visually. Scatterplots among independent variables are also plotted to ensure whether there is collinearity problem among the independent variables.

Correlation analysis: Pearson's correlation analysis is used to identify any correlation between the dependent variables and independent variables. The significance of the correlation is also tested. Correlation analysis is also conducted among independent variables and variance inflation factor (VIF) is checked to see whether there is severe collinearity among independent variables.

ANOVA: ANOVA test is used to check for significant differences among means of variables.

Multiple Linear Regression: OLS linear regression models are conducted. The independent variables are selected by backward elimination function.

Diagnose: Assumptions for the regression including normality, non-constant variance will be checked using the residual plot. If the regression model does not satisfy these assumptions, transformation might be needed.

Transformation: Box-Cox is used to determine to best transformation for this dataset.

Final regression: Final regression is performed after transformation. The assumptions for regression model are verified again.

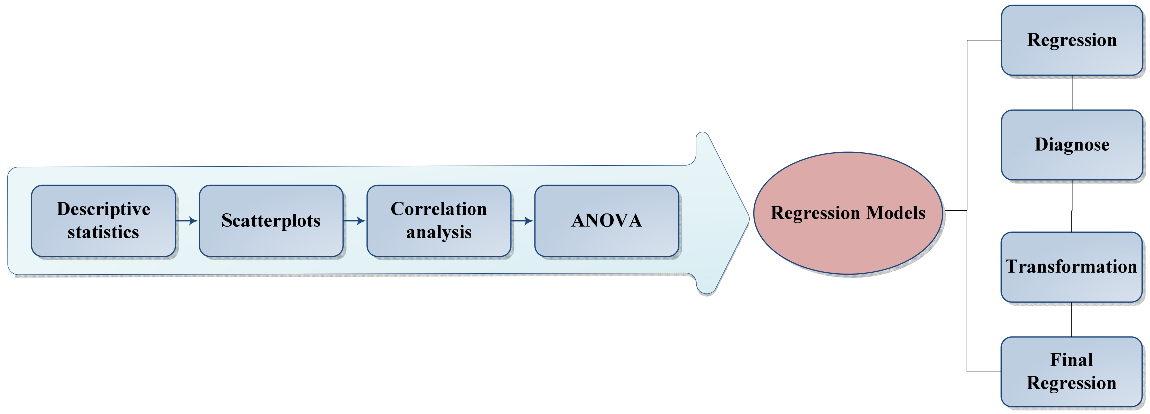

Figure 6. Method of Analysis

Descriptive Statistics

The sample data consist of 212 properties in the Galveston County. Among the 212 properties, 185 parcels have only one building on this parcel while 27 parcels have multiple buildings. In the following analysis, there are two dependent variables, TWIA payout and TWIA payout ratio. TWIA Payout Ratio represents the proportional damage caused by Hurricane Ike. It is the TWIA payout divided by improvement value of the year 2008, which could be calculated using the Equation 2.

TWIA Payout Ratio = (TWIA Payout) / (2008 Improvement Value) (2)

Independent variables are distance from the coastline (m), distance from the Hurricane track (m), distance from the landfall center of Hurricane Ike (m), floodplain zone (A, X, X500), the number of buildings per parcel (0/1), and Galveston Island (0/1), improvement value of 2008 ($).

In terms of floodplain zone, we have three values, A, X, X500. Only two variables A and X500 are used in the regression model. These two variables are coded as A (0/1) and X500 (0/1). For the variable floodplain zone A, 1 represent the property is within the floodplain zone A while 0 represent the property falls outside of the floodplain zone A. Similarly, 1 of zone X500 means the building is within the floodplain zone X500 while 0 means the building is outside of floodplain zone X500. Therefore, we have (A, X500) for each property. (1, 0) means that the property is within the floodplain zone A; (0, 1) means that the property is within the floodplain zone X500; (0, 0) means that the property is within the floodplain zone X. For the categorical variable the number of buildings per parcel, 1 represents that there is just one building in the parcel while 0 represents that there are multiple buildings in the parcel. Similarly, for the categorical variable the Galveston Island, 1 represents that the building is on the Galveston Island while 0 represents that the building is not on the Galveston Island.

Figure 7 shows the qualified buildings based on their geographic location. For each qualified building, X and Y coordinates illustrate the geographic location. The Z values or heights of these building points represent value of TWIA payout ratio, which is one of the dependent variable in this study. There are several hot spots in the map and these hot spots are displayed using darker color.

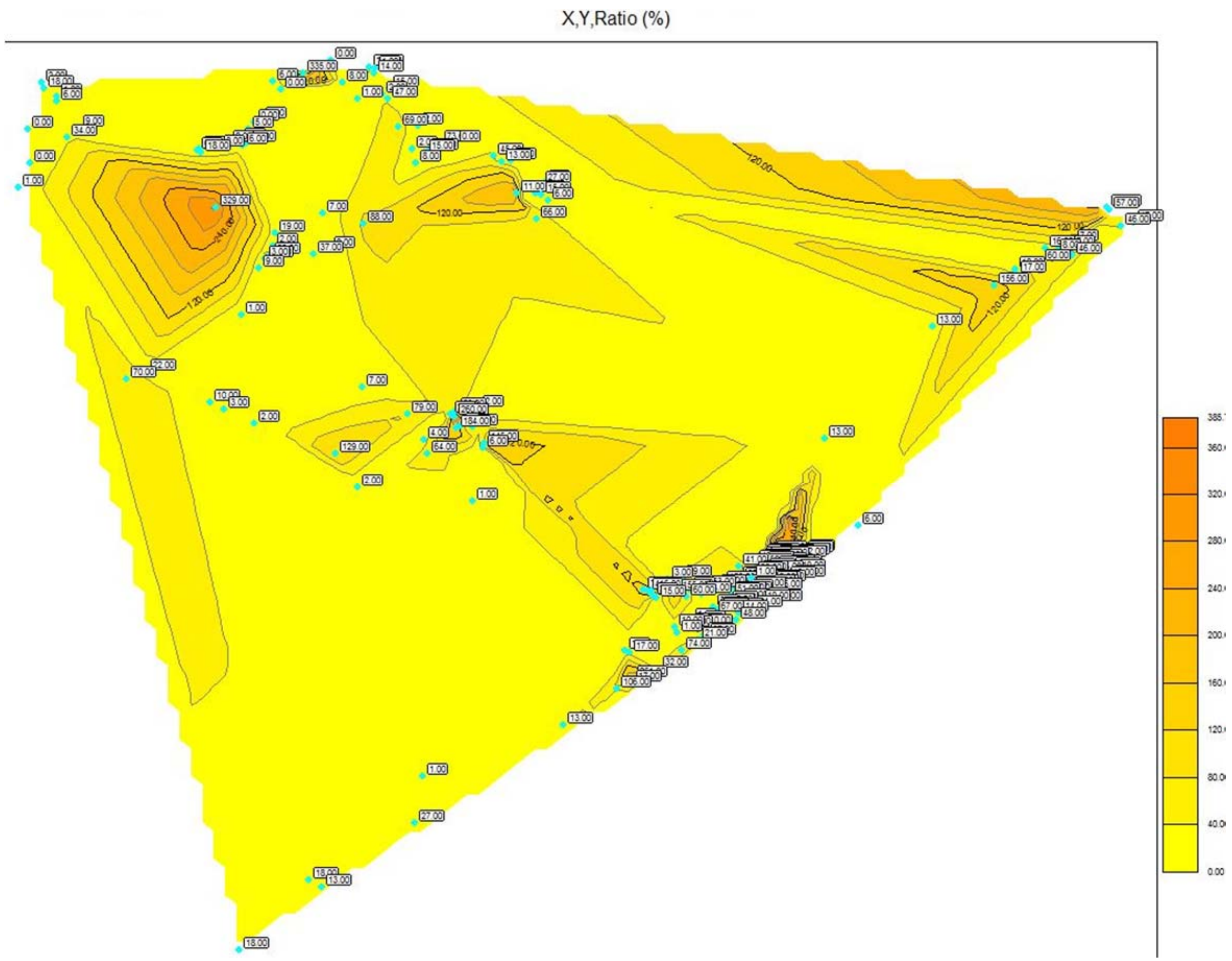

Figure 7. Locations and Ratio Value of Buildings

A preliminary summary of the data are given by the descriptive statistic (Table 4). The dependent variable, TWIA payout, has a mean of 74,796.60 and a median of 17,680. TWIA Payout Ratio has a mean of 0.493 and a median of 0.15. Mean of the improvement value is 291,012.18 and its median is 131,715. Mean of the distance from the coastline is 12,556.07 meters and its median is 1,911.67 meters. Mean of the distance from hurricane track is 10,051.09 meters and its median is 7,568.15 meters. Mean of the distance from landfall center of Hurricane Ike is 17,899.49 meters and its median is 13,584 meters. The average building area is 1,320.34 square meters and its median is 407.52 square meters. For the categorical data – floodplain zone, 58 percent of the parcels are within Floodplain Zone A, 27.4 percent are within Floodplain Zone X, and 14.6 percent are within Floodplain Zone X500 (Table 5). Regarding the variable – number of buildings per parcel, 87.3 percent of the parcels only have one building. 12.7 percent of the parcels have more than one property (Table 6). 59.4 percent of all buildings are located in the Galveston Island (Table 7).

Table 4. Descriptive Statistics

	Mean	Std. Deviation	Median	N
TWIA Payout ($)	74796.60	157925.57	17680	212
IMP ($)	291012.18	555421.76	131715	212
TWIA Payout Ratio	0.493	1.115	0.150	212

Table 5 Continuted. Descriptive Statistics

Distance from the Coastline (m)	12556.07	14775.726	1911.67	212
Distance from the Hurricane Track (m)	10051.09	8116.79	7568.15	212
Distance from the landfall center (m)	17899.49	14522.84	13584	212
Building Area (sq.m)	1320.34	2987.95	407.52	212

Table 6. Descriptive Statistic of Floodplain Zone

Floodplain Zone	Frequency	Percent	Cumulative Percent
A	123	58.0	58.0
X	58	27.4	85.4
X500	31	14.6	100.0
Total	212	100.0	

Table 7. Descriptive Statistic of No. of Buildings

No. of Buildings	Frequency	Percent	Cumulative Percent
0	27	12.7	12.7
1	185	87.3	100.0
Total	212	100.0	

Table 8. Descriptive Statistic of Galveston Island

Galveston Island	Frequency	Percent	Cumulative Percent
0	86	40.6	40.6
1	126	59.4	100.0
Total	212	100.0	

One issue regarding the preliminary descriptive analysis is that 28 out of 212 buildings have the TWIA payout ratio beyond 1 (Figure 8). These points are larger than Q3 + 1.5 IQR. Numbers that are larger than Q3 + 1.5 IQR are outliers. As shown in the Box Plot (Figure 9), there are many points that do not fall in the inner fences of the box. The circles and asterisks are outliers and asterisks are extreme outliers.

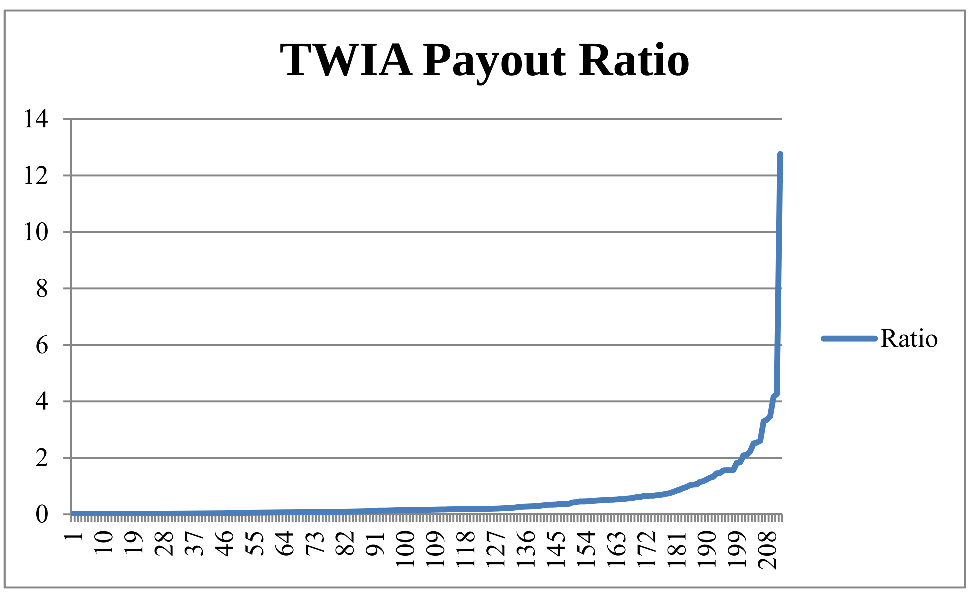

Figure 8. Distribution of TWIA Payout Data

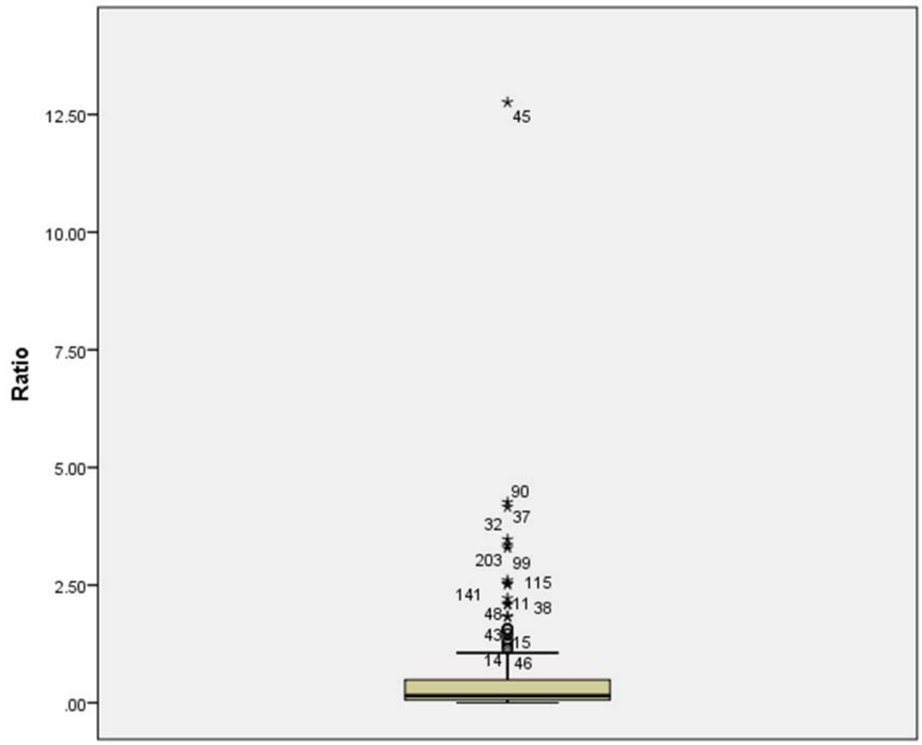

Figure 9. Box plot of TWIA Payout Ratio

There are numerous reasons why payments on these properties may exceed the improvement value obtained from the appraisal district. After investigating these properties and interviewing representatives from TWIA, I came up with several reasons that might be responsible for this issue.

- The market value of a building fluctuates based on many conditions and is not always an accurate proxy for the cost to rebuild the structure if it were destroyed.

This cost to rebuild or replace the building is normally the basis for insurance policy limits and indemnity payments.

- The payments include all property loss at the specified location and exclude other coverage such as business interruption, per the requestor's original instructions. TWIA defines property loss to include loss to the contents of a structure in addition to the structure itself. Payments made for contents losses might cause the total payments to exceed the value of only the structure.

- Insured can also receive up to 10% of the structure limit in coverage for additional structures, or outbuildings, located at the insured location. This can also increase total property payments beyond the value of the primary structure.

- Some of these claims may have been litigated or have some level of attorney involvement. In these situations, a settlement is often reached with the plaintiff which may or may not include consideration of other issues in addition to disputed property loss. These payments are treated as loss payments by TWIA and it is impossible to remove any potential other damages from the property damage for these payments.

Based on the theory of outlier and realistic situations, these buildings are excluded in the following analysis. After excluding these outliers, we have a new descriptive statistics (Table 8, 9, 10, 11).

Table 9. New Descriptive Statistics

	Mean	Std. Deviation	Median	N
TWIA Payout ($)	47020.91	122906.44	13376.5	184
IMP ($)	313870.38	590337.908	147500	184
TWIA Payout Ratio	0.202	0.223	0.13	184
Distance from the Coastline (m)	10259.74	15113.95	1929.78	184
Distance from the Hurricane Track (m)	10259.74	8306.4	7568.15	184
Distance from the landfall center (m)	18544.73	14819.73	18959	184
Building Area (sq.m)	1427.63	3187.56	402.78	184

Table 10. New Descriptive Statistic of Floodplain Zone

Floodplain Zone	Frequency	Percent	Cumulative Percent
A	102	55.4	55.4
X	55	29.9	85.3
X500	27	14.7	100.0
Total	184	100.0	

Table 11. New Descriptive Statistic of No. of Buildings

No. of Buildings	Frequency	Percent	Cumulative Percent
0	22	12.0	12.0
1	162	88.0	100.0
Total	184	100.0	

Table 12. New Descriptive Statistic of Galveston Island

Galveston Island	Frequency	Percent	Cumulative Percent
0	78	42.4	42.4
1	106	57.6	100.0
Total	184	100.0	

Scatter Plots

The scatterplots for TWIA payout and TWIA payout ratio versus the distance from the coastline indicate a negative relationship (Figure 10, 11), which meets our expectation. As the building's distance from the coastline increases, the TWIA payout and TWIA payout ratio decrease.

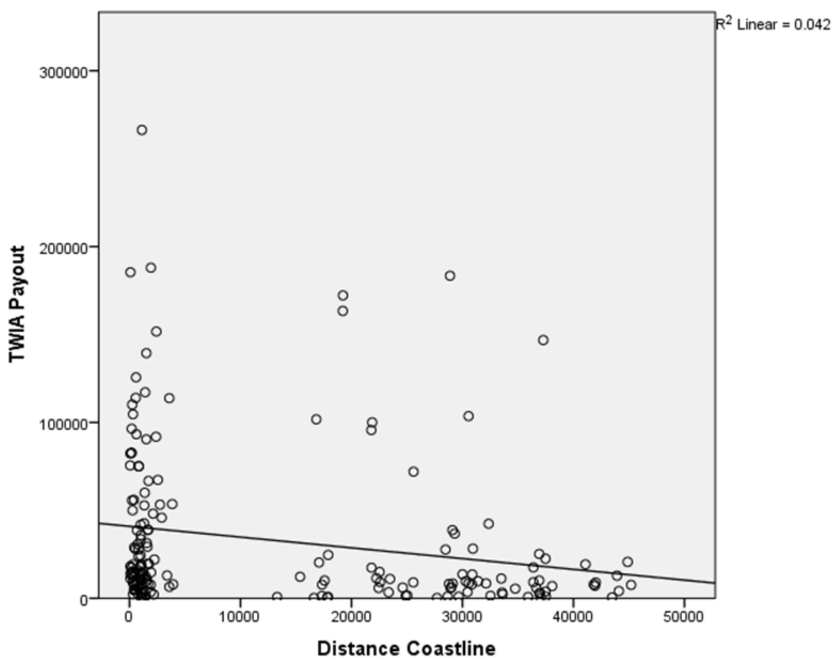

Figure 10. Scatterplot of TWIA Payout Versus Distance from the Coastline

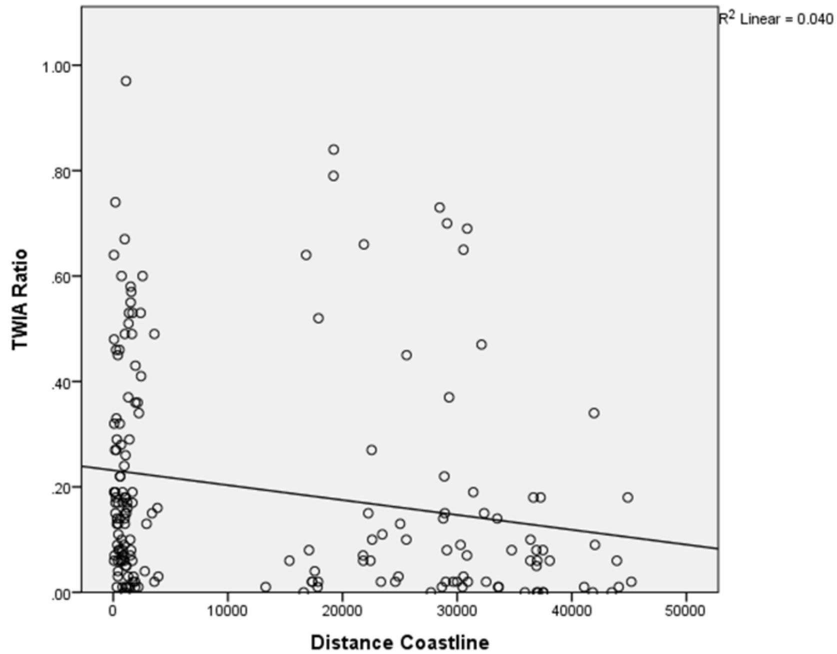

Figure 11. Scatterplot of TWIA Payout Ratio Versus Distance from the Coastline

The scatterplots for TWIA payout and TWIA payout ratio versus the distance from the hurricane track indicate a negative relationship (Figure 12, 13). It also meets our expectation. As the building's distance from the hurricane track increases, the TWIA payout and payout ratio decrease.

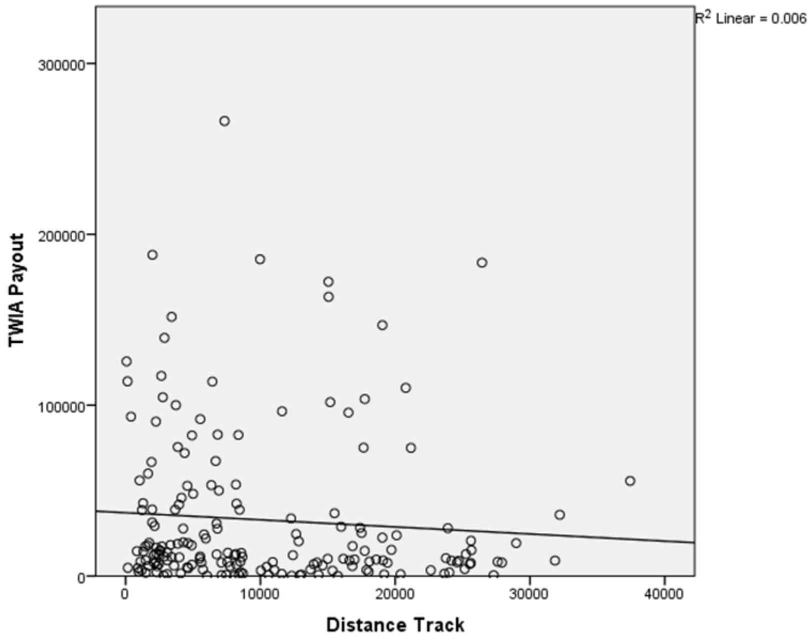

Figure 12. Scatterplot of TWIA Payout Versus Distance from the Track

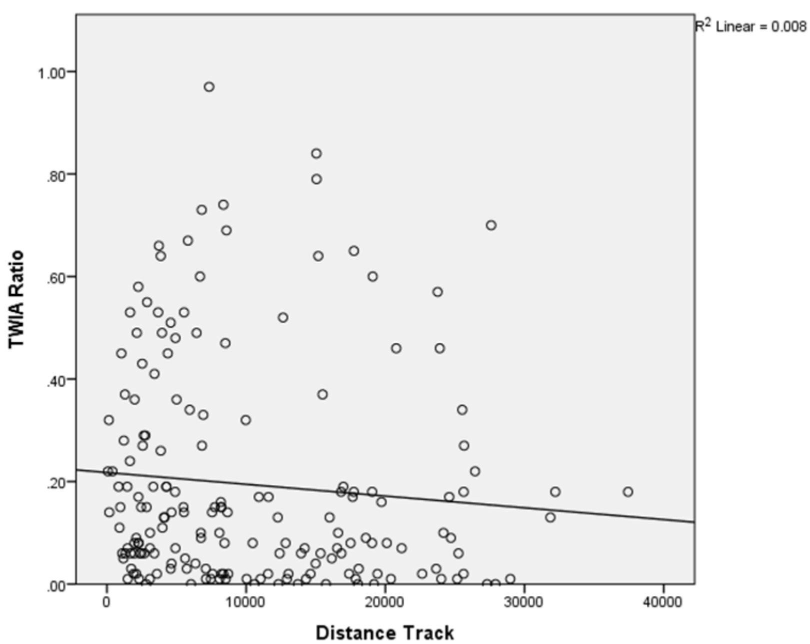

Figure 13. Scatterplot of TWIA Payout Ratio Versus Distance from the Track

The scatterplots for TWIA payout and payout ratio versus the distance from the landfall center of the hurricane also indicate a negative relationship (Figure 14, 15). As the distance from the landfall center increases, the TWIA payout and payout ratio decrease.

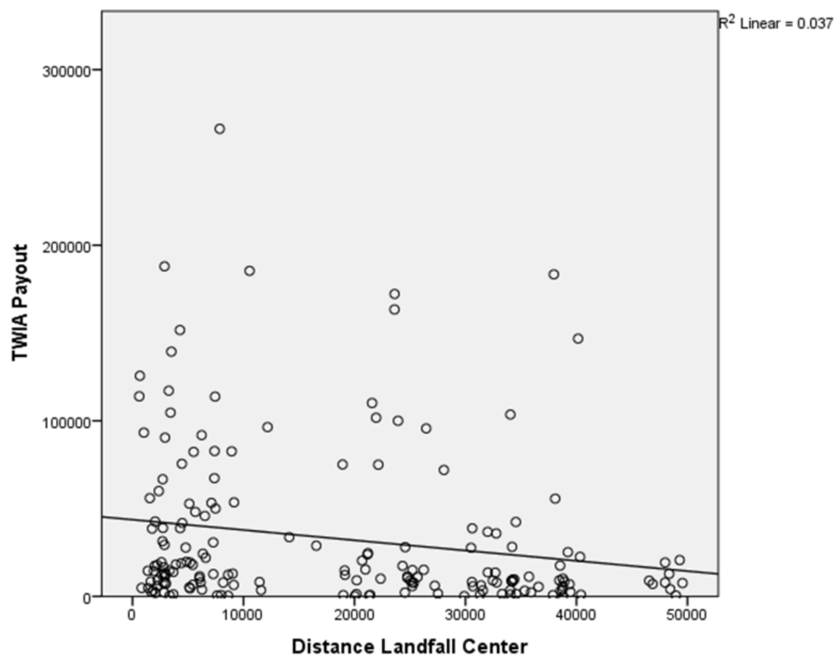

Figure 14. Scatterplot of TWIA Payout Versus Landfall Center of Hurricane

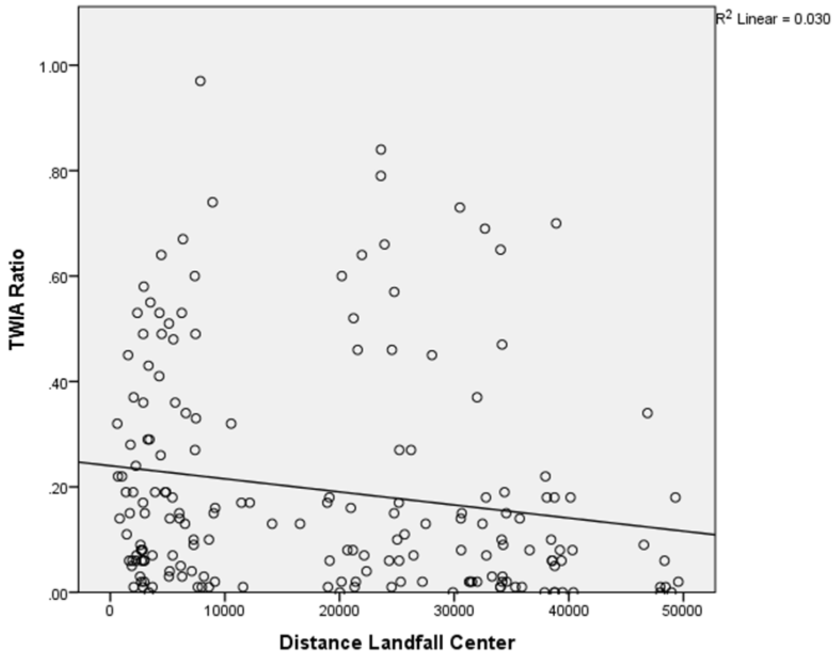

Figure 15. Scatterplot of TWIA Payout Ratio Versus Landfall Center of Hurricane

The scatterplot for TWIA payout versus the building area indicates a positive relationship (Figure 16). As the building area increases, the TWIA payout increases. However, the scatterplot for TWIA payout ratio versus the building area indicates a negative relationship (Figure 17). As the building area increases, the TWIA payout ratio decreases. However, many points cluster together, making the interpolation difficult.

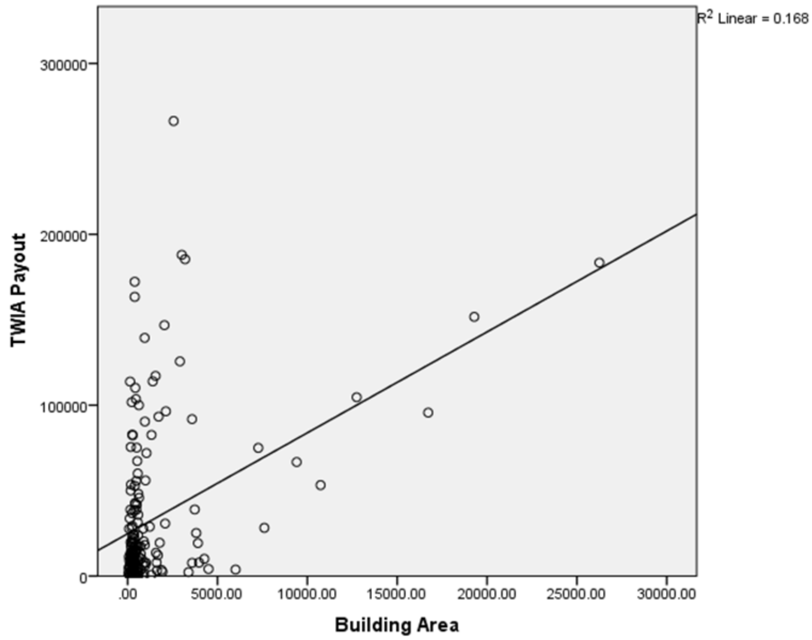

Figure 16. Scatterplot of TWIA Payout Ratio Versus Building Area

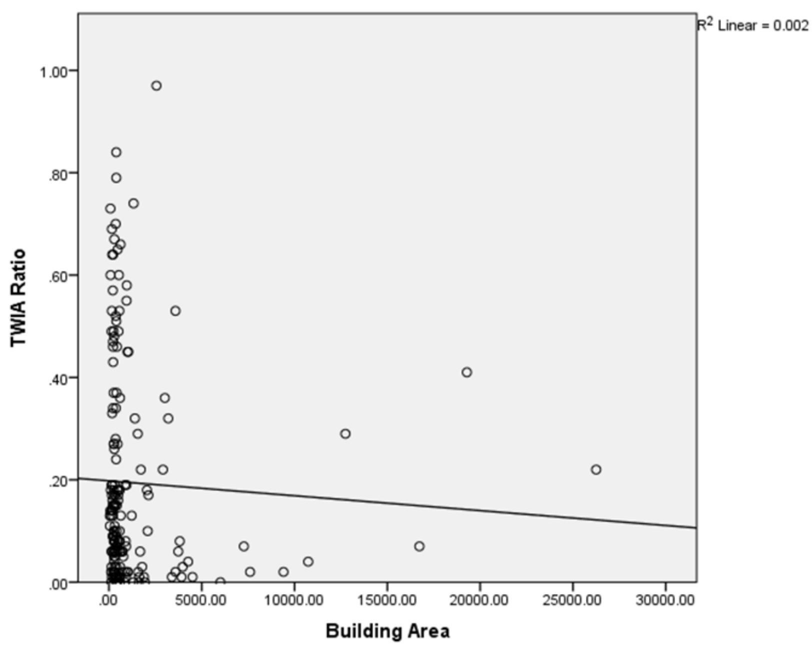

Figure 17. Scatterplot of TWIA Payout Ratio Versus Building Area

The scatterplots for TWIA payout and payout ratio versus floodplain zone indicate that buildings in the Floodplain Zone A tend to receive more payout or payout ratio than those in the Floodplain Zone X (Figure 18, 19). It satisfies our expectation.

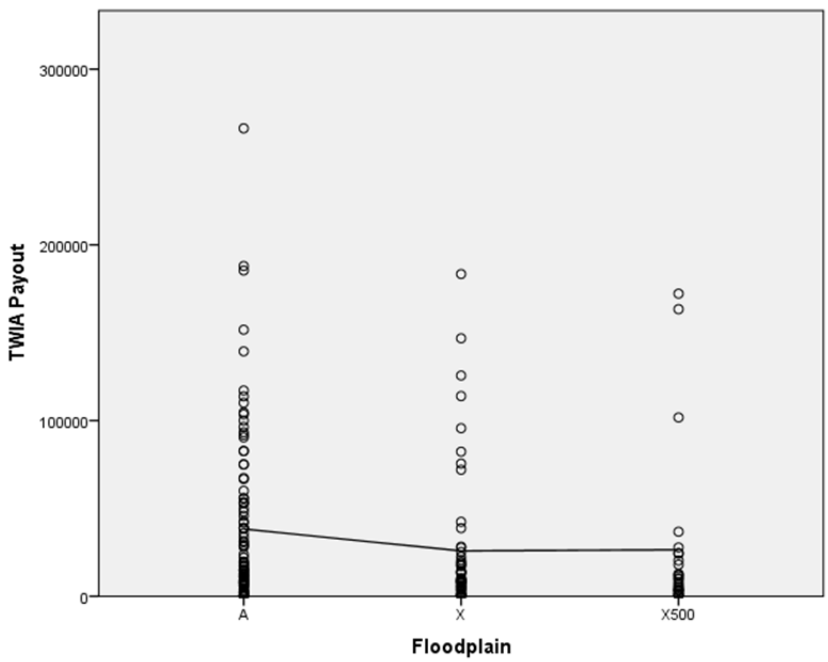

Figure 18. Scatterplot of TWIA Payout Versus Floodplain Zone

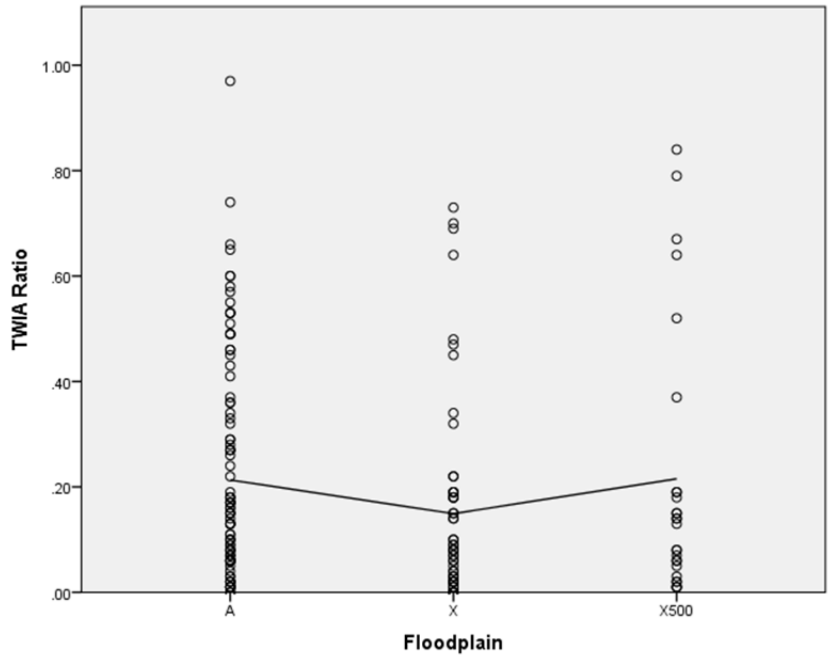

Figure 19. Scatterplot of TWIA Payout Ratio Versus Floodplain Zone

The scatterplots for TWIA payout and payout ratio versus Galveston Island indicate that building in the Galveston Island (1) tends to receive more payout or payout ratio (Figure 20, 21).

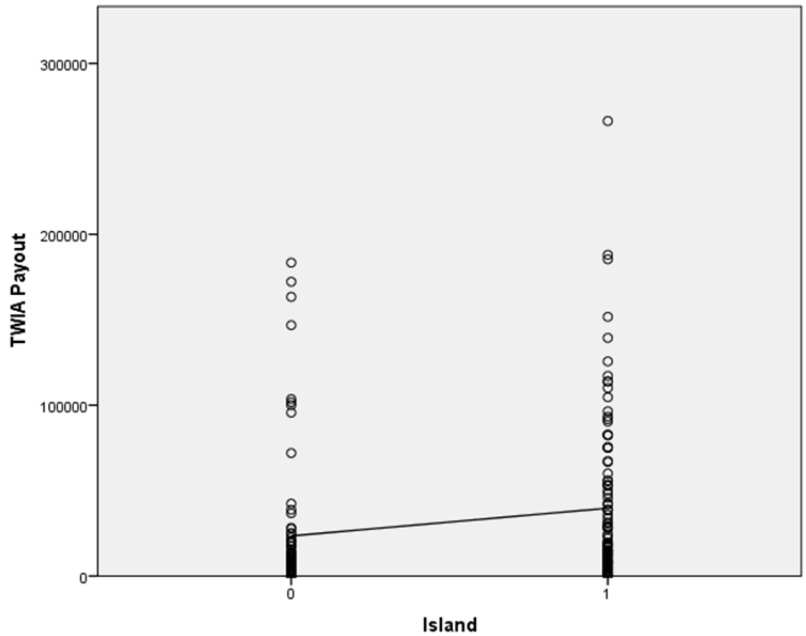

Figure 20. Scatterplot of TWIA Payout Ratio Versus Galveston Island

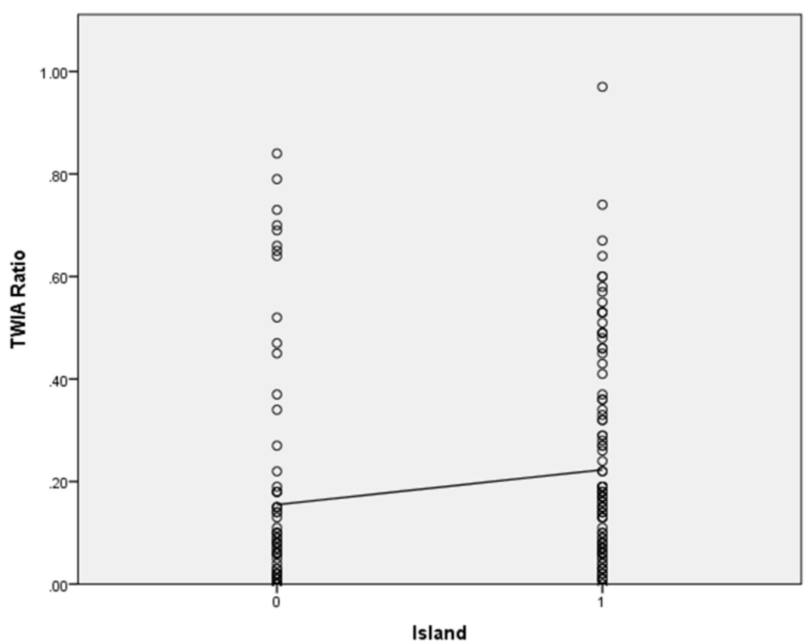

Figure 21. Scatterplot of TWIA Payout Ratio Versus Galveston Island

The scatterplot for TWIA payout versus No. of building per parcel indicates that the parcel with one building seem to receive more TWIA payout (Figure 22, 23), which is not what we expect. However, because of the few points for multiple buildings per parcel, it is difficult to make interpolation based on the scatterplots.

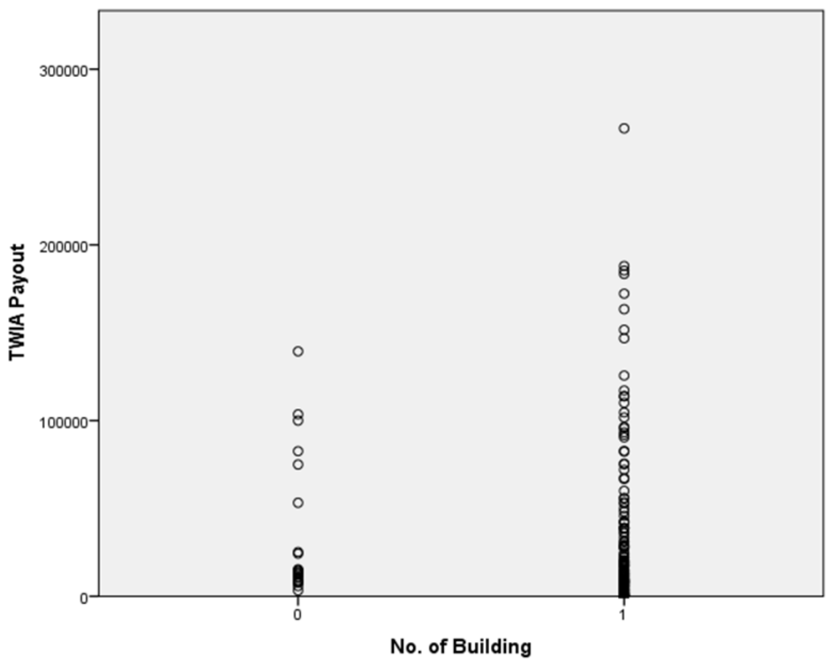

Figure 22. Scatterplot of TWIA Payout Versus No. of Building per Parcel

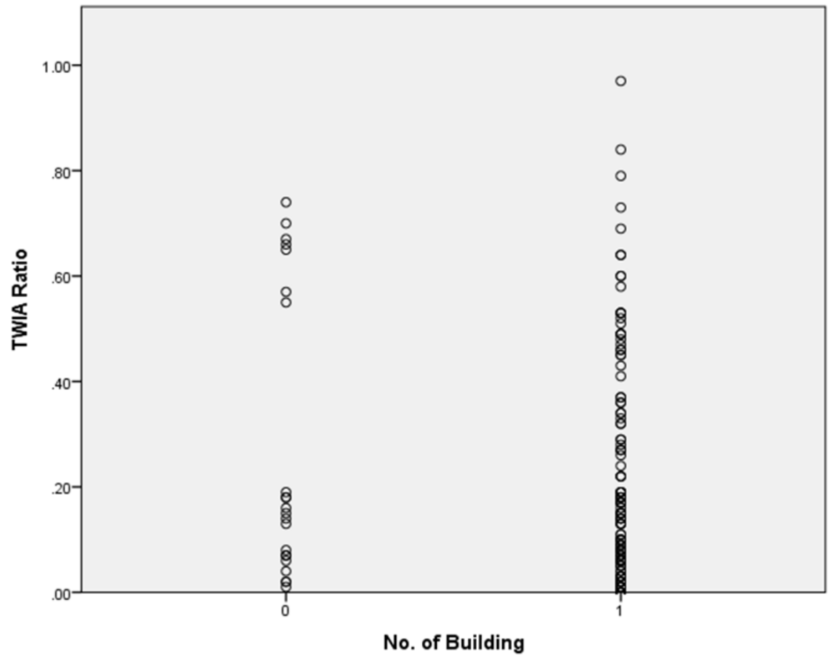

Figure 23. Scatterplot of TWIA Payout Ratio Versus No. of Building per Parcel

Correlation Analysis

Correlation analysis is conducted between dependent variables and independent variables. Note that the correlations are not computed with the variables Floodplain A, Floodplain X500, number of building, and Galveston Island. All of these variables are indicator variables and their correlations with the other variables are not meaningful. The Person's Correlation results are shown in Table 12. Correlations that are significant at the 0.05 level (2-tailed) are highlighted in yellow. Correlations that are significant at the 0.01 level (2-tailed) are highlighted in red.

As shown in the Person's Correlation results, three independent variables, distance from the coastline, distance from the landfall center and building area, are correlated to the TWIA payout, which is significant at the 0.01 level. Distance from the coastline and distance from the landfall center have negative relations with the TWIA payout. As the distance from the coastline or distance from the landfall center increases, the TWIA payout decreases. The variable, building area, has a positive relation with the TWIA payout. As the building area increases, the TWIA payout also increases. Moreover, the improvement value is correlated to the TWIA payout that is significant at the 0.05 level. The TWIA payout increases as the improvement value increases.

Distance from the coastline is correlated to the TWIA payout ratio at the level of 0.01 and distance from the landfall center is correlated to the TWIA payout ratio. The correlations are significant at the 0.05 level. These two variables have negative relations to the TWIA payout ratio, which means the TWIA payout ratio decreases as the distance from the coastline or landfall center increases.

Correlation is also conducted to detect any collinearity problem among explanatory variables. There appears to be some collinearity problems between some variables. In particular, improvement value and building area has a correlation of 0.474, which is significant at the 0.01 level. Distance from the coastline and distance from the hurricane track have a positive relationship of 0.499. These correlations are also significant at the 0.01 level. Distance from the coastline and distance from the landfall center has a very strong relationship of 0.910 at the 0.01 level. As 0.910 is beyond our threshold value of 0.9, there is a severe collinearity problem between these two variables. Either distance from the coastline or distance from the land center would be dropped in the regression model. For the collinearity problem, VIF is calculated in the Table 15 as well.

Table 13. Correlation Results

		TWIA Payout	IMP	TWIA Ratio	Distance Coastline	Distance Track	Distance Landfall	Building Area
TWIA Payout	Pearson Correlation	1	.172*	.605**	-.205**	-.077	-.193**	.410**
	Sig. (2-tailed)		.021	.000	.006	.304	.009	.000
IMP	Pearson		1	-.232**	.164*	.174*	.157*	.474**
	Sig. (2-tailed)			.002	.027	.019	.035	.000
TWIA Ratio	Pearson			1	-.201**	-.091	-.174*	-.043
	Sig. (2-tailed)				.007	.226	.019	.566
Distance Coastline	Pearson				1	.499**	.910**	.042
	Sig. (2-tailed)					.000	.000	.577
Distance Track	Pearson					1	.779**	.086
	Sig. (2-tailed)						.000	.251
Distance Landfall	Pearson						1	.036
	Sig. (2-tailed)							.631
Building Area	Pearson							1
	Sig. (2-tailed)							

*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

Regression Analysis

The hypotheses include two regression models. One model is to predict the TWIA payout, and the other is to predict the TWIA payout ratio.

Original Payout Model – Full Model

For the full model of TWIA payout, the dependent variable is TWIA Payout and the predictors include improvement value of 2008, Floodplain A, Floodplain x500, number of buildings per parcel, distance from the hurricane track, building area, Galveston Island, distance from the coastline, and distance from the landfall center of Hurricane Ike. The full model has an R Square of 0.220 and an Adjusted R Square of 0.179 (Table 13). The ANOVA test indicates that the model is significant at the level 0.001 (Table 14). However, as shown in Table 15, there is severe collinearity problem in this model. The VIF of distance from the landfall center is 37.044 and the VIF of distance from coastline is 26.205. Both of these variables are large than 10, which indicate a collinearity problem. One of the two variables would be dropped in the next analysis.

Table 14. Payout Model - Full Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.469[a]	.220	.179	40919.472

Table 15. Payout Model - Full Model ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	80895432095.585	9	8988381343.954	5.368	.000[b]
Residual	286322946622.968	171	1674403196.626		
Total	367218378718.552	180			

Table 16. Coefficients and Collinearity Statistics for Payout Model – Full Model

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	44158.853	23351.117		1.891	.060		
Floodplain A	3234.828	9036.559	.036	.358	.721	.458	2.182
Floodplain X500	-491.607	10867.254	-.004	-.045	.964	.637	1.570
Distance Coastline	-.490	1.032	-.164	-.475	.635	.038	26.205
Distance Track	.532	1.057	.099	.503	.616	.119	8.414
Distance Landfall Center	-.688	1.248	-.226	-.551	.582	.027	37.044
No. of Building	-1087.041	9498.671	-.008	-.114	.909	.960	1.041
Island	-11470.347	18684.430	-.126	-.614	.540	.108	9.225
Building Area	5.824	1.124	.404	5.183	.000	.751	1.332
IMP	.001	.008	.014	.172	.863	.718	1.393

Payout Model – Backward Elimination Selection

Backward elimination is used to select the regression model. The criteria for backward elimination are: when the probability of F less than or equal to 0.01, enter the variable; when the probability of F is larger than 0.01, remove the variable.

Because of the collinearity problem between the distance from the coastline and the distance from the landfall center, the variable - distance from the landfall center is eliminated in the regression model. After conducting backward elimination, we came up with 7 models (Table 16) and all of these models are significant at the level of 0.01 (Table 17). Based on the Adjusted R Square, the best regression model is model 7. In model 7, the dependent variable is TWIA Payout and independent variables are building area and distance from coastline. The Adjusted R Square for this model is 0.208 (Table 16), which means 20.8% of the variability could be explained by this model.

Table 17. Payout Model Summary - Backward

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.468[a]	.219	.183	40836.554
2	.468[b]	.219	.187	40718.358
3	.468[c]	.219	.192	40601.183
4	.468[d]	.219	.197	40485.404
5	.468[e]	.219	.201	40376.913
6	.467[f]	.218	.205	40273.028
7	.466[g]	.217	.208	40188.810

a. Predictors: (Constant), IMP, No. of Building, Floodplain A, Distance Track, Building Area, Floodplain X500, Island , Distance Coastline

b. Predictors: (Constant), IMP, No. of Building, Floodplain A, Distance Track, Building Area, Island , Distance Coastline

c. Predictors: (Constant), IMP, No. of Building, Floodplain A, Building Area, Island , Distance Coastline

d. Predictors: (Constant), IMP, Floodplain A, Building Area, Island , Distance Coastline

e. Predictors: (Constant), Floodplain A, Building Area, Island , Distance Coastline

f. Predictors: (Constant), Building Area, Island , Distance Coastline

g. Predictors: (Constant), Building Area, Distance Coastline

h. Dependent Variable: TWIA Payout

Table 18. Payout Model ANOVA - Backward

Model		Sum of Squares	df	Mean Square	F	Sig.
	Regression	80387024701.268	8	10048378087.658	6.026	.000[b]
1	Residual	2868E+11	172	1667624151.263		
	Total	3672 E+11	180			
	Regression	80387023396.879	7	11483860485.268	6.926	.000[c]
2	Residual	2868 E+11	173	1657984712.842		
	Total	3672 E+11	180			
	Regression	80387018552.798	6	13397836425.466	8.128	.000[d]
3	Residual	2868E+11	174	1648456092.907		
	Total	3672 E+11	180			
	Regression	80381491061.364	5	16076298212.273	9.808	.000[e]
4	Residual	2868E+11	175	1639067929.470		
	Total	3672 E+11	180			
	Regression	80286444501.065	4	20071611125.266	12.312	.000[f]
5	Residual	2868E+11	176	1630295080.781		
	Total	3672 E+11	180			
	Regression	80139101697.895	3	26713033899.298	16.470	.000[g]
6	Residual	2868E+11	177	1621916819.326		
	Total	3672 E+11	180			
	Regression	79723372537.244	2	39861686268.622	24.680	.000[h]
7	Residual	2868E+11	178	1615140484.165		
	Total	3672 E+11	180			

a. Dependent Variable: TWIA Payout

b. Predictors: (Constant), IMP, No. of Building, Floodplain A, Distance Track, Building Area, Floodplain X500, Island , Distance Coastline

c. Predictors: (Constant), IMP, No. of Building, Floodplain A, Distance Track, Building Area, Island , Distance Coastline

d. Predictors: (Constant), IMP, No. of Building, Floodplain A, Building Area, Island , Distance Coastline

e. Predictors: (Constant), IMP, Floodplain A, Building Area, Island , Distance Coastline

f. Predictors: (Constant), Floodplain A, Building Area, Island , Distance Coastline

g. Predictors: (Constant), Building Area, Island , Distance Coastline

h. Predictors: (Constant), Building Area, Distance Coastline

Before we proceed to any further analysis, it is essential to check model assumptions. The assumptions for regression model consist of zero expectation, constant variance, normality, and independence.

- Zero expectation: E (εi) = 0 for all i.
- Constant variance: V (εi) = $\sigma 2\varepsilon$ for all i.
- Normality: εi is normally distributed.
- Independence: the εi are independent.

As shown in the Histogram and Normal P-P Plot of regression standardized residual (Figure 24, 25), the model does not satisfy the normality condition. Furthermore, it does not meet the constant variance assumption (Figure 26). Therefore, transformation is needed.

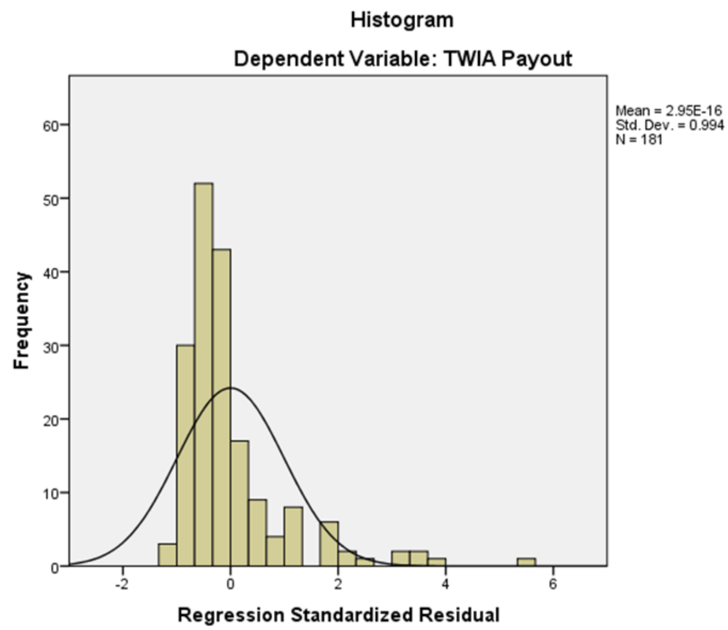

Figure 24. Histogram Plot of Standardize Residual for Payout Model – Backward

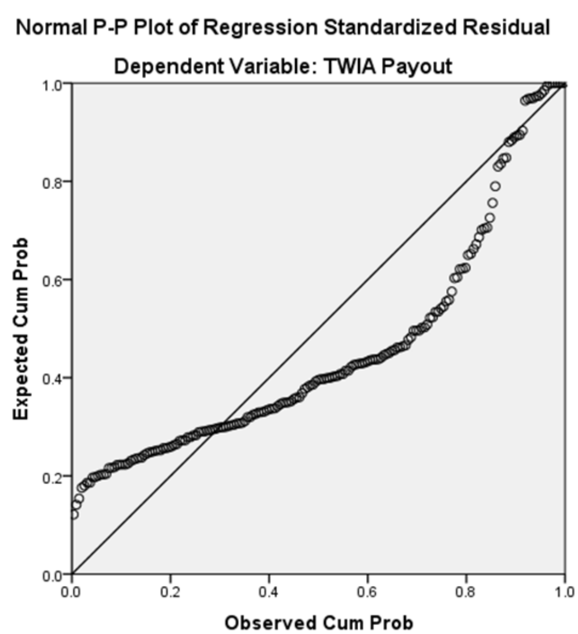

Figure 25. Normal Probability Plot of Standardized Residual Plots - Backward

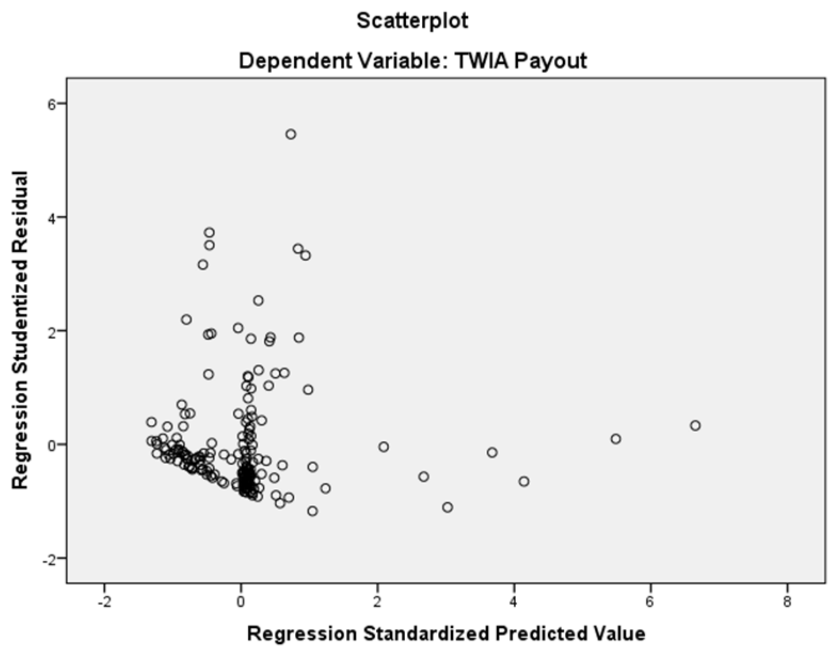

Figure 26. Studentized Residual of Payout Model - Backward

Transformation

Because the original regression model residuals are not normally distributed, Box-Cox transformation is used to determine an ideal transformation. A grid of values of λ (-2, -1.75, -1.5, -1.25, -1.0, -0.75, -0.5, -0.25, 0, 0.25, 0.50, 0.75, 1.0, 1.25, 1.5, 1.75, 2, log) is used to transform the dataset.

When taking the normality and MS (Residual) into consideration, the transformation of $\lambda=0.25$ receives the best result. The dependent variable TWIA payout is transformed using $\lambda=0.25$ transformation (Equation 3). After transforming the dependent variable, the transformed model still does not satisfy the non-constant variance assumption. Then, the independent variables are also transformed (Equation 4, 5).

Transformed (TWIA payout) = (TWIA payout) $^{0.25}$ (3)

Transformed (Building area) = (Building area) $^{0.25}$ (4)

Transformed (Distance from coastline) = (Distance from coastline) $^{0.25}$ (5)

	Payout	Payout_0.25	Area_sqm	Area_0.25	Coast_m	Coast_0.25
1	1570	6.29	37.90	2.48	25031	12.58
2	4683	8.27	54.63	2.72	443	4.59
3	11220	10.29	68.38	2.88	33496	13.53
4	27693	12.90	79.43	2.99	28475	12.99
5	9138	9.78	89.19	3.07	725	5.19
6	5806	8.73	97.36	3.14	36652	13.84
7	850	5.40	111.48	3.25	17871	11.56
8	33687	13.55	115.57	3.28	1011	5.64
9	103	3.19	120.03	3.31	37075	13.88
10	6040	8.82	127.09	3.36	752	5.24
11	4831	8.34	135.82	3.41	581	4.91
12	1378	6.09	139.35	3.44	24872	12.56
13	113779	18.37	139.35	3.44	3586	7.74
14	38840	14.04	153.29	3.52	1678	6.40
15	11476	10.35	154.96	3.53	996	5.62

Figure 27. Transformation of TWIA Payout Model

Final Payout Model – After Transformation

After transformation, we have a new model with an Adjusted R Square of 0.264 (Table 18), which means that 26.4 percent of variability in transformed TWIA payout could be explained by this model. The Adjusted R Square has been improved through the transformation. Moreover, the model is significant at the level of 0.01 (Table 19). Two independent variables, distance from coastline and building area, are also significant at the level of 0.01 (Table 20). As shown in Table 20, VIF for the distance from the coastline and building area are around 1. It means there is no collinearity problem in this regression model.

Regarding the assumptions of regression, the new histogram plot of standardize residual and Normal Probability Plot of Standardized Residual indicate that the transformed regression satisfy the normality condition (Figure 28, 29). The Kolmogorov-Smirnov value and Shapiro-Wilk value also confirm that the residuals are normally distributed because the p values are larger than 0.05 (Table 21). Furthermore, the residual plot is a null plot and it does not show any violation of constant variance (Figure 30). Breusch – Pagan (BP) test is also performed to test the homogeneous variances for the regression model.

The BP tests the hypotheses:

H_0: Homogeneous variances

H_a: Heterogeneous variances

From Table 19 we obtain SS (Residual) = 2349.12 and from the Table 22 we obtain SS (Regression)* = 289.769. Then, BP statistic is computed as follows.

BP = [SS (Regression)*/2] ÷ [SS (Residuals)/n]^2 = (289.769/2) ÷ (2349.12/181)^2 = 0.86

Based on the chi-squared table, we obtain that P value is 0.35. Because P value is larger than 0.05, we cannot reject H_0: homogenous variances and conclude that there is no non-constant variance in this model. Therefore, the transformed regression model meets all the assumptions.

The coefficients of (Distance from coastline)$^{0.25}$ and (Building area)$^{0.25}$ are -0.348 and 0.979 (Table 20). The transformed TWIA payout model could be established using the Equation 6. The TWIA payout could be predicted by distance from coastline and building area using the Equation 7.

(TWIA Payout)$^{0.25}$ = 9.421 - 0.348 * (Distance from coastline)$^{0.25}$ + 0.979* (Building Area)$^{0.25}$ (6)

TWIA Payout = [9.421 - 0.348 * (Distance from coastline)$^{0.25}$ + 0.979* (Building Area)$^{0.25}$]$^{(1/0.25)}$ (7)

The coefficients of independent variables mean:

- As the distance from the coastline increase, the TWIA payout decreases.
- As the building area increases, the TWIA payout increases.

Table 19. Final Payout Model Summary - after Transformation

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.522[a]	.273	.264	3.47918

a. Predictors: (Constant), $(\text{Building Area})^{0.25}$, $(\text{Distance Coastline})^{0.25}$

b. Dependent Variable: $\text{Payout}^{0.25}$

Table 20. Final Payout Model ANOVA - after Transformation

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	735.951	2	367.976	27.883	.000[b]
Residual	2349.120	178	13.197		
Total	3085.071	180			

Table 21. Coefficients of Payout Model – after Transformation

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	9.421	.959		9.822	.000		
(Distance Coastline)0.25	-.348	.067	-.337	-5.229	.000	.996	1.004
(Building Area)0.25	.979	.150	.420	6.518	.000	.996	1.004

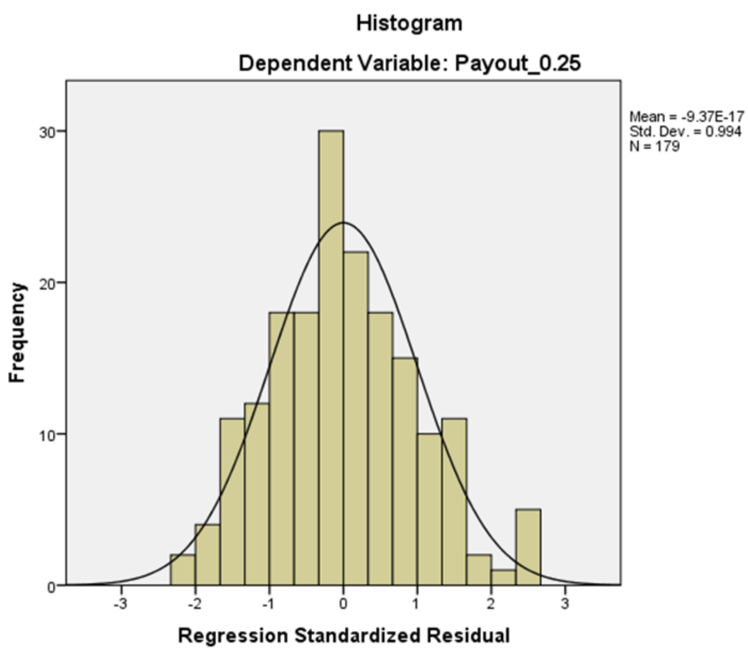

Figure 28. Histogram Plot of Standardize Residual for Payout Model – after Transformation

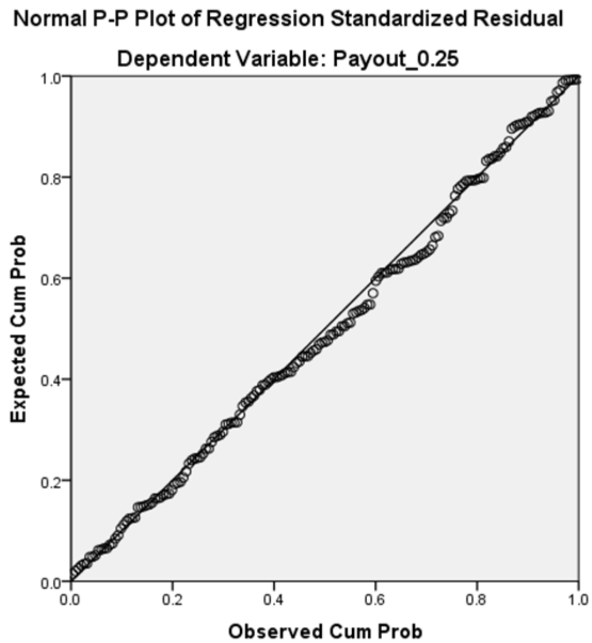

Figure 29. Normal Probability Plot of Standardized Residual Plots

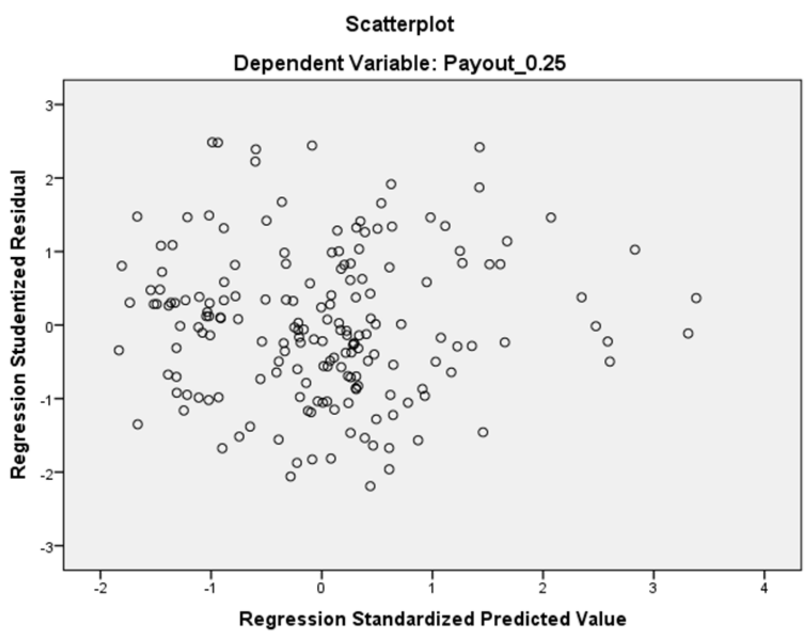

Figure 30. Studentized Residual of Payout Model – after Transformation

Table 22. Test for Normality - after Transformation

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Studentized Residual	.052	179	.200*	.988	179	.152

Table 23. Residual_SQ Regression Model ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	289.769	2	144.884	.416	.660[b]
Residual	61946.887	178	348.016		
Total	62236.656	180			

a. Dependent Variable: Residual_SQ

b. Predictors: (Constant), (Building Area)$^{0.25}$, (Distance Coastline)$^{0.25}$

Validity of Final Payout Model

The Adjusted R Square of the final payout model is 0.264. It means that 26.4 percent of variability in transformed TWIA payout could be explained by this model. Moreover, there is no collinearity problem of independent variables in this model.

To illustrate the validity of this model, the scatterplot of actual TWIA payout versus predicted TWIA payout is shown in the Figure 31.

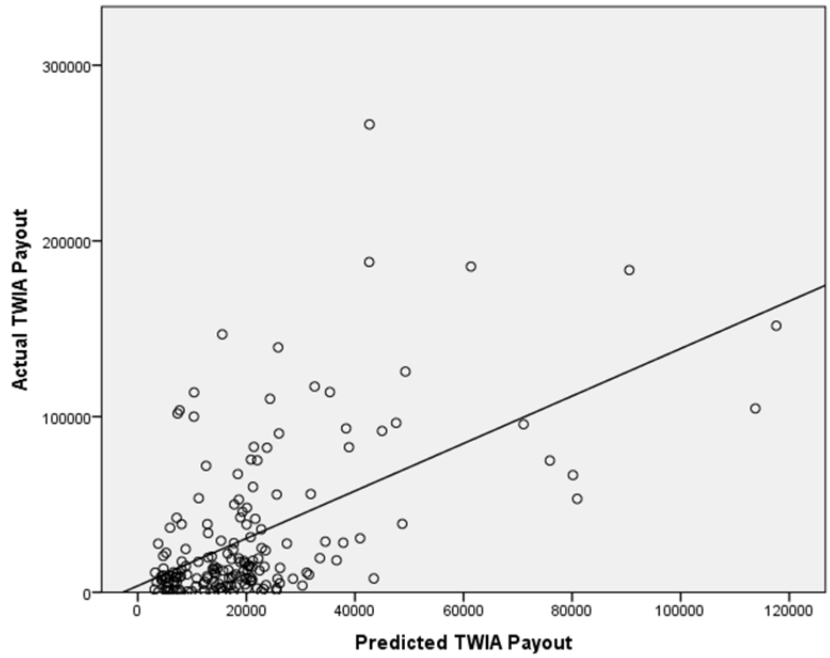

Figure 31. Actual Versus Predicted TWIA Payout

For parcels where there is only one property, another model with building age is also tested. After conducting backward elimination, the building age is not significant.

Original Payout Ratio Model – Full Model

For the full model of TWIA payout ratio, the dependent variable is TWIA payout ratio and predictors include Floodplain A, Floodplain X500, numbers of building per parcel, distance from the hurricane track, building area, island, distance from the coastline, and distance from the landfall center of Hurricane Ike. The full model of TWIA payout ratio has an R square of 0.061 (Table 23). The ANOVA test indicates that the model is not significant (Table 24). As shown in Table 25, there are also severe collinearity problems in this model. The VIF of distance from the landfall center is 36.761 and the VIF of distance from coastline is 25.769. Both of these variables are large than 10, which indicate a potential collinearity problem. One of the two variables would be dropped in the next analysis.

Table 24. Payout Ratio Model - Full Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.246[a]	.061	.018	.22080

Table 25. Payout Ratio Model - Full Model ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	.552	8	.069	1.415	.193[b]
Residual	8.532	175	.049		
Total	9.084	183			

Table 26. Coefficients and Collinearity Statistics for Payout Ratio - Full Model

Payout Ratio – Full Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	.374	.126		2.981	.003		
Floodplain A	.007	.049	.017	.154	.878	.453	2.205
Floodplain X500	.046	.057	.074	.806	.421	.645	1.551
Distance Coastline	-5.370E-006	.000	-.364	-.980	.329	.039	25.769
Distance Track	-4.107E-007	.000	-.015	-.073	.942	.121	8.245
Distance Landfall Center	5.570E-007	.000	.037	.083	.934	.027	36.761
No. of Building	-.085	.051	-.125	-1.670	.097	.962	1.039
Building Area	-2.749E-007	.000	-.004	-.052	.958	.945	1.059
Island	-.074	.100	-.165	-.741	.460	.108	9.284

Payout Ratio Model – Backward Elimination Selection

Backward elimination is used to select the regression model. The criteria for backward elimination is the probability of F to enter <=0.01 and probability of F to remove >0.01.

Because of the collinearity problem between the distance from the coastline and the distance from the landfall center, the variable - distance from the landfall center is eliminated. After conducting backward elimination, we came up with 6 models (Table 26). Five out of six models are significant at the level of 0.05 (Table 27). Based on the Adjusted R Square, the best regression model is model 5. However, in model 5, the VIF of Galveston Island and distance from coastline are 7.884 and 7.882. Even though these two VIF are not larger than 10, there is a potential problem of collinearity. Based on the R Square and VIF, Model 6 is selected. Model 6 has two independent variables, distance from the coastline and No. of building. However, the Adjusted R Square for this model is 0.051 (Table 26), which is pretty small.

Table 27. Payout Ratio Model Summary - Backward

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.266[a]	.071	.033	.20853
2	.266[b]	.071	.039	.20793
3	.266[c]	.071	.044	.20735
4	.265[d]	.070	.049	.20682
5	.262[e]	.069	.053	.20640
6	.249[f]	.062	.051	.20656

a. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Distance Track, Floodplain A, Distance Coastline

b. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Floodplain A, Distance Coastline

c. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Distance Coastline

d. Predictors: (Constant), Island , No. of Building, Building Area, Distance Coastline

e. Predictors: (Constant), Island , No. of Building, Distance Coastline

f. Predictors: (Constant), No. of Building, Distance Coastline

Table 28. Payout Ratio Model ANOVA - Backward

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.574	7	.082	1.885	.075[b]
	Residual	7.523	173	.043		
	Total	8.097	180			
2	Regression	.574	6	.096	2.211	.044[c]
	Residual	7.523	174	.043		
	Total	8.097	180			
3	Regression	.573	5	.115	2.664	.024[d]
	Residual	7.524	175	.043		
	Total	8.097	180			
4	Regression	.569	4	.142	3.324	.012[e]
	Residual	7.528	176	.043		
	Total	8.097	180			
5	Regression	.556	3	.185	4.351	.006[f]
	Residual	7.541	177	.043		
	Total	8.097	180			
6	Regression	.502	2	.251	5.882	.003[g]
	Residual	7.595	178	.043		
	Total	8.097	180			

a. Dependent Variable: TWIA Ratio

b. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Distance Track, Floodplain A, Distance Coastline

c. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Floodplain A, Distance Coastline

d. Predictors: (Constant), Island , No. of Building, Building Area, Floodplain X500, Distance Coastline

e. Predictors: (Constant), Island , No. of Building, Building Area, Distance Coastline

f. Predictors: (Constant), Island , No. of Building, Distance Coastline

g. Predictors: (Constant), No. of Building, Distance Coastline

Once again, it is essential to check model assumptions before we proceed to the further analysis. As shown in the Histogram and Normal P-P Plot of regression standardized residual (Figure 32, 33), the model does not satisfy the normality condition. Furthermore, the residual plot does not seem to be random (Figure 34). Therefore, transformation is needed for this model.

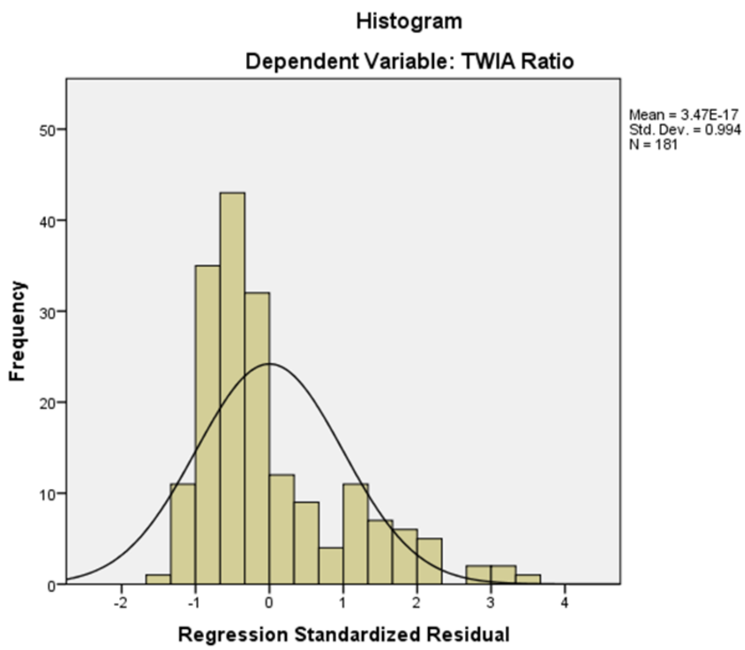

Figure 32. Histogram Plot of Residual Plot for Payout Ratio Model – Backward

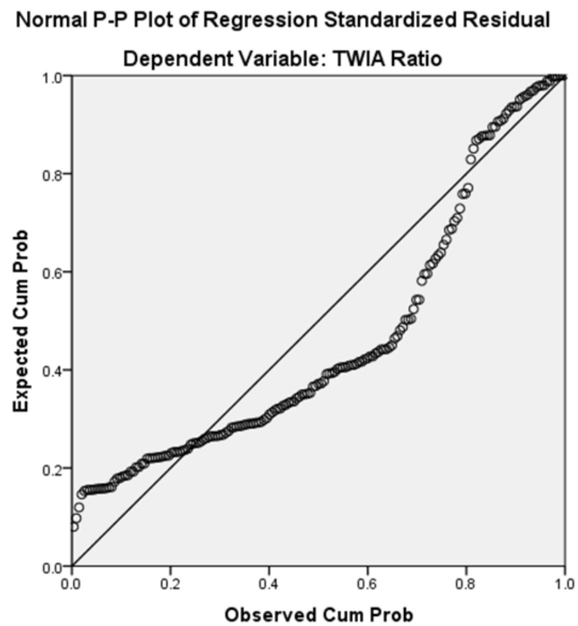

Figure 33. Normal Probability Plot of Standardized Residual Plots – Backward

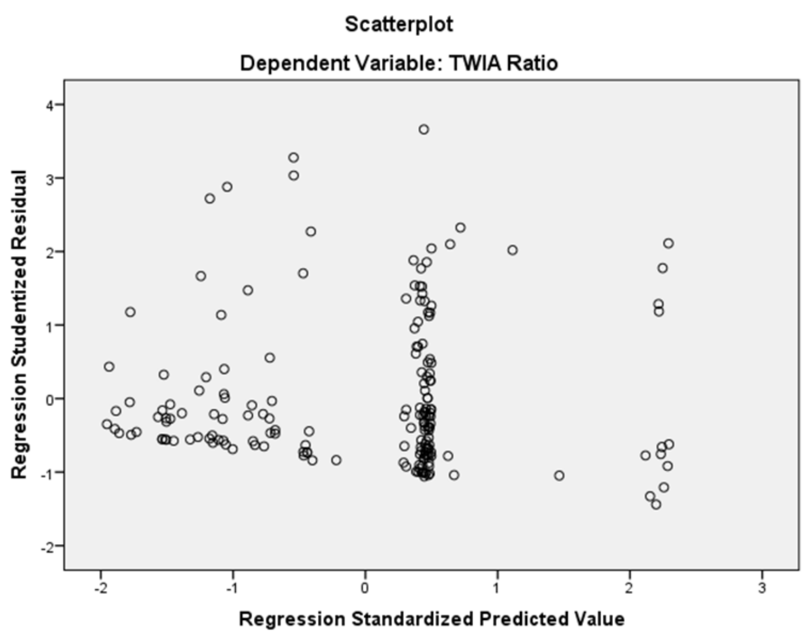

Figure 34. Studentized Residual of Payout Model - Backward

Transformation

Because the original regression model residuals are not normally distributed, Box-Cox transformation is also used to transform the dataset. Form a grid of values of λ (-2, -1.75, -1.5, -1.25, -1.0, -0.75, -0.5, -0.25, 0, 0.25, 0.50, 0.75, 1.0, 1.25, 1.5, 1.75, 2, log), the transformation of λ=0.25 receives the best result when taking the normality and MS(Residual) into consideration.

The dependent variable TWIA payout ratio is transformed using λ=0.25 transformation (Equation 8). Because of the non-constant variance, the independent variable, distance from coastline, is also transformed (Equation 9).

$$\textbf{Transformed (TWIA payout ratio)} = \textbf{(TWIA payout ratio)}^{0.25} \qquad \textbf{(8)}$$

$$\textbf{Transformed (Distance from coastline)} = \textbf{(Distance from coastline)}^{0.25} \qquad \textbf{(9)}$$

	Ratio	Ratio0.25	Coast_m	Coast_0.25
1	.97	.99	1118	5.78
2	.92	.98	1107	5.77
3	.88	.97	29200	13.07
4	.84	.96	19232	11.78
5	.79	.94	19214	11.77
6	.74	.93	186	3.70
7	.73	.92	28475	12.99
8	.70	.91	29124	13.06
9	.69	.91	30878	13.26
10	.67	.90	996	5.62
11	.66	.90	21866	12.16
12	.65	.90	30551	13.22
13	.64	.89	60	2.79
14	.64	.89	16828	11.39
15	.60	.88	2556	7.11

Figure 35. Transformation of TWIA Payout Ratio Model

Final Payout Ratio Model – After Transformation

After transformation, we have a new model with an Adjusted R Square of 0.112 (Table 28), which means that 11.2 percent of variability in transformed TWIA payout ratio could be explained by this model. Moreover, the model is significant at the level of 0.01 (Table 29). As shown in Table 30, VIF for the distance from the coastline and building area is around 1. Therefore, there is no collinearity problem in this model.

Regarding the assumptions of regression, the new histogram plot of standardize residual and Normal Probability Plot of Standardized Residual indicate that the transformed regression satisfy the normality condition (Figure 36, 37). The Kolmogorov-Smirnov value also confirms that the residuals are normally distributed. P value for Kolmogorov-Smirnov test of normality is 0.2 and P value for Shapiro-Wilk is 0.094. Both of these numbers are larger than 0.05 (Table 31). The transformed regression model meets the normality assumption.

The residual plot (Figure 38) seems to satisfy the constant variance assumption. However, there is a violation of constant variance after performing BP test. Therefore, the OLS regression models are run using the robust standard errors to offset the heteroskedastic error. Results for this new model are shown in Table 32. R Square is 0.121.

In the new model, the coefficients of (Distance from coastline) $^{0.25}$ and No. of Building are -0.018 and -0.099 (Table 32). The transformed TWIA payout ratio model could be established using the Equation 10. The TWIA payout ratio could be predicted by distance from coastline and No. of building using the Equation 11.

(TWIA Payout Ratio) $^{0.25}$ = 0.812 - 0.018 * (Distance from coastline) $^{0.25}$ - 0.099 * (No. of Building) (10)

TWIA Payout Ratio= [0.812 - 0.018 * (Distance from coastline) $^{0.25}$ - 0.099 * (No. of Building)] $^{(1/0.25)}$ (11)

The coefficients of independent variables mean:

- As the distance from the coastline increase, the TWIA payout ratio decreases.
- Parcels that have multiple buildings tend to have larger TWIA payout ratio than those that have only one building.

Table 29. Final Payout Ratio Model Summary - after Transformation

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.348[a]	.121	.112	.20986

a. Predictors: (Constant), Coast $^{0.25}$, No. of Building
b. Dependent Variable: Ratio $^{0.25}$

Table 30. Final Payout Ratio Model ANOVA - after Transformation

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.084	2	.542	12.302	.000[b]
Residual	7.839	178	.044		
Total	8.923	180			

a. Dependent Variable: $Ratio^{0.25}$

b. Predictors: (Constant), $Coast^{0.25}$, No. of Building

Table 31. Coefficients of Payout Ratio Model – after Transformation

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	.814	.057		14.162	.000		
No. of Building	-.099	.048	-.146	-2.071	.040	.999	1.001
$Coast^{0.25}$	-.018	.004	-.322	-4.576	.000	.999	1.001

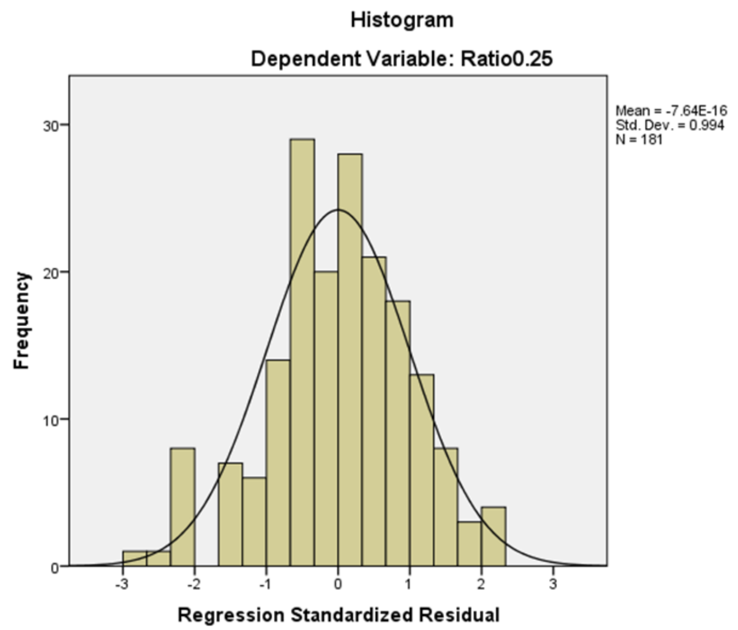

Figure 36. Histogram Plot of Standardize Residual for Payout Ratio Model – after Transformation

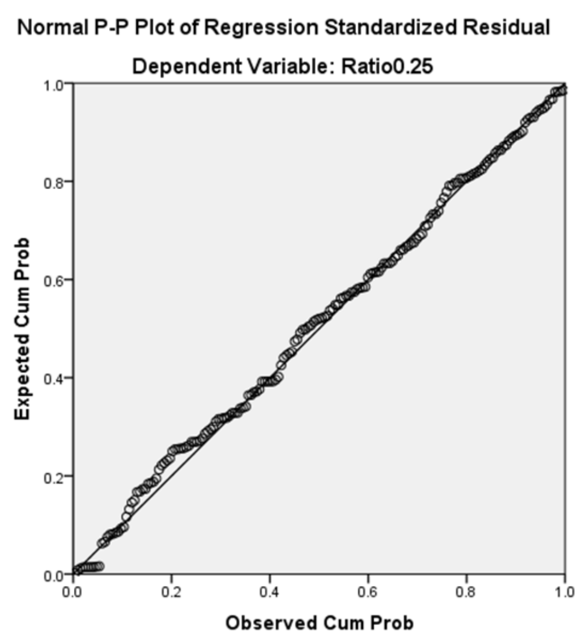

Figure 37. Normal Probability Plot of Standardized Residual Plots

Table 32. Test for Normality - after Transformation

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Studentized Residual	.050	181	.200*	.987	181	.094

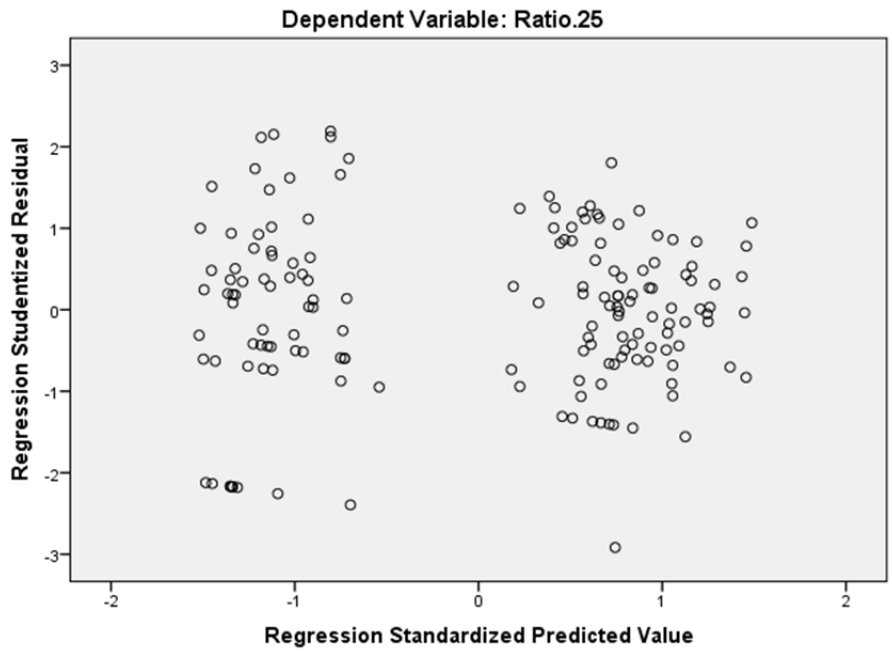

Figure 38. Studentized Residual of Payout Ratio Model – after Transformation

Table 33. Payout Ratio Model - Using Robust Standard Errors

Model	Coefficients	Robust Std. Error	t	Sig.	R Square
(Constant)	.812	.053	15.26	.000	
No. of Building	-.099	.045	-2.22	.028	0.1212
$Coast^{0.25}$	-.018	.004	-4.576	.000	

Payout Ratio Model – Building Age Included

For parcels where there is only one property, another model with building age is also conducted. After performing backward elimination, we came up with 6 models (Table 33). Five out of six models are significant at the level of 0.05 (Table 34). Based on the Adjusted R Square, the best regression model is Model 5. However, in Model 5, the VIF of Galveston Island and distance from coastline are 7.618 and 7.545. Although these VIF are not larger than 10, there is a potential problem of collinearity. Based on the R Square and VIF, Model 6 is selected. Model 6 has two independent variables, distance from the coastline and building age. However, the Adjusted R Square for this model is 0.057 (Table 33), which is pretty small.

One again, it is essential to check model assumptions before we proceed to the further analysis. The model does not satisfy the normality condition as shown in the Histogram and Normal P-P Plot of regression standardized residual (Figure 39, 40). Furthermore, the residual plot does not seem to satisfy the constant variance assumption (Figure 41). Therefore, transformation is needed for this model.

Table 34. Payout Ratio Model Summary – Building Age Included

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.282[a]	.079	.037	.19774
2	.281[b]	.079	.043	.19710
3	.281[c]	.079	.049	.19649
4	.279[d]	.078	.054	.19597
5	.274[e]	.075	.057	.19561
6	.261[f]	.068	.056	.19574

a. Predictors: (Constant), Building Age, Floodplain X500, Building Area, Distance Coastline, Distance Track, Floodplain A, Island

b. Predictors: (Constant), Building Age, Floodplain X500, Distance Coastline, Distance Track, Floodplain A, Island

c. Predictors: (Constant), Building Age, Floodplain X500, Distance Coastline, Distance Track, Island

d. Predictors: (Constant), Building Age, Distance Coastline, Distance Track, Island

e. Predictors: (Constant), Building Age, Distance Coastline, Island

f. Predictors: (Constant), Building Age, Distance Coastline

Table 35. Payout Ratio Model ANOVA – Building Age Included

Model		Sum of Squares	df	Mean Square	F	Sig.
	Regression	.509	7	.073	1.858	.080[b]
1	Residual	5.904	151	.039		
	Total	6.413	158			
	Regression	.508	6	.085	2.179	.048[c]
2	Residual	5.905	152	.039		
	Total	6.413	158			
	Regression	.506	5	.101	2.620	.026[d]
3	Residual	5.907	153	.039		
	Total	6.413	158			
	Regression	.499	4	.125	3.246	.014[e]
4	Residual	5.914	154	.038		
	Total	6.413	158			
	Regression	.482	3	.161	4.201	.007[f]
5	Residual	5.931	155	.038		
	Total	6.413	158			
	Regression	.436	2	.218	5.689	.004[g]
6	Residual	5.977	156	.038		
	Total	6.413	158			

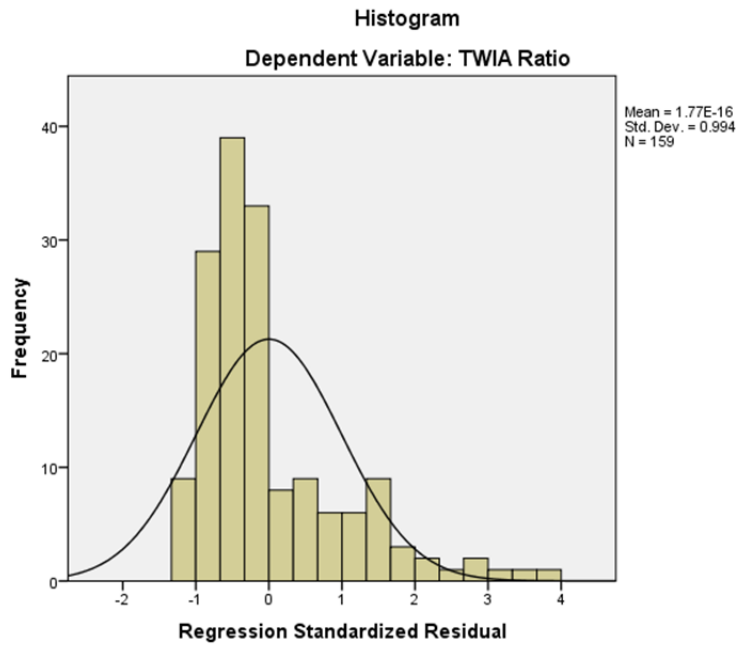

Figure 39. Histogram Plot of Standardize Residual for Payout Ratio Model – Building Age Included

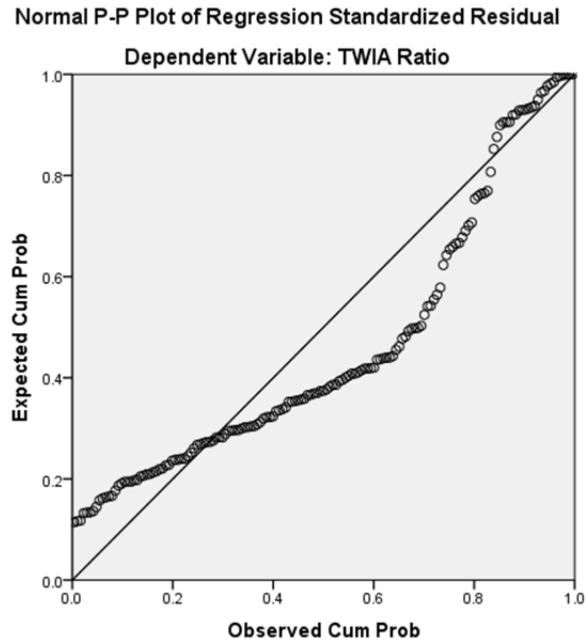

Figure 40. Normal Probability Plot of Standardized Residual Plots – Building Age Included

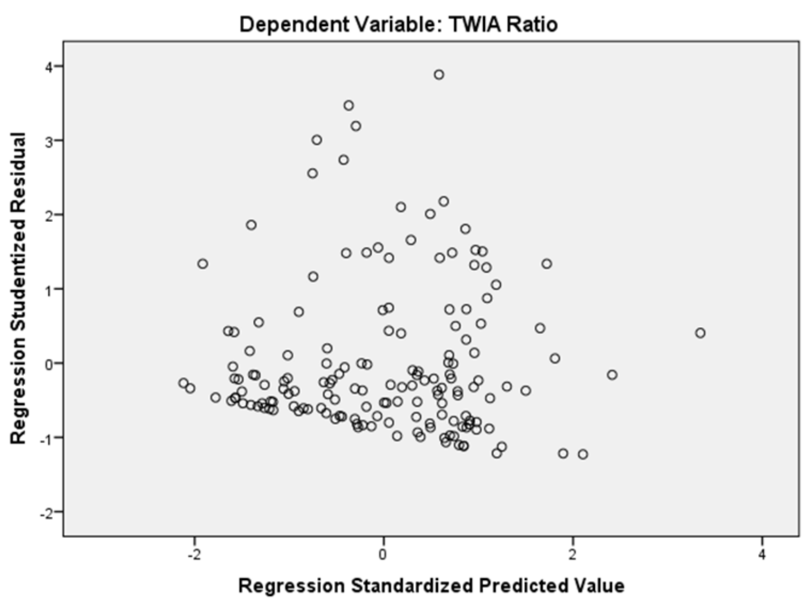

Figure 41. Studentized Residual of Payout Model - Building Age Included

Transformation

Because the original regression model residuals are not normally distributed, Box-Cox transformation is used for this data set. Form a grid of values of λ (-2, -1.75, -1.5, -1.25, -1.0, -0.75, -0.5, -0.25, 0, 0.25, 0.50, 0.75, 1.0, 1.25, 1.5, 1.75, 2, log), the transformation of λ=0.25 receives the best result when taking the normality and MS(Residual) into consideration.

The dependent variable TWIA payout ratio is transformed using λ=0.25 transformation (Equation 12). Because of the non-constant variance, the independent variable, distance from coastline and building age also transformed (Equation 13, 14).

$$\textbf{Transformed (TWIA payout ratio)} = \textbf{(TWIA payout ratio)}^{0.25} \qquad \textbf{(12)}$$

$$\textbf{Transformed (Distance from coastline)} = \textbf{(Distance from coastline)}^{0.25} \qquad \textbf{(13)}$$

$$\textbf{Transformed (Building age)} = \textbf{(Building age)}^{0.25} \qquad \textbf{(14)}$$

Final Payout Ratio Model – Building Age Included

After transformation, we have a new model with an Adjusted R Square of 0.139 (Table 35), which means that 13.9 percent of variability in transformed TWIA payout ratio could be explained by this model. Moreover, the model is significant at the level of 0.01 (Table 36). As shown in Table 37, VIF for the distance from the coastline and building age is around 1. Therefore, there is no collinearity problem in this model.

Regarding the assumptions of regression, the new histogram plot of standardize residual and Normal Probability Plot of Standardized Residual indicate that the transformed regression satisfy the normality condition (Figure 42, 43). The Kolmogorov-Smirnov value also confirms that the residuals are normally distributed because the p value is larger than 0.5 (Table 38). After performing BP test, there is still a violation of constant variance. Therefore, the OLS regression models are run using the robust standard errors to offset the heteroskedastic error. Results for this new model are shown in Table 39. R Square is 0.15.

In the new model using robust standard errors, the coefficients of (Distance from coastline)$^{0.25}$ and (Age)$^{0.25}$ are -0.018 and 0.078 (Table 39). The transformed TWIA payout ratio model could be established as shown in the Equation 16. The TWIA payout ratio could be predicted by distance from coastline and building age using the Equation 16.

$(\text{TWIA Payout Ratio})^{0.25} = 0.538 - 0.018 * (\text{Distance from coastline})^{0.25} + 0.078 * (\text{Building age})^{0.25}$ (15)

$\text{TWIA Payout Ratio} = [0.538 - 0.018 * (\text{Distance from coastline})^{0.25} + 0.078 * (\text{Building age})^{0.25}]^{(1/0.25)}$ (16)

The coefficients of independent variables mean:

- As the distance from the coastline increase, the TWIA payout ratio decreases.
- As the building age increases, the TWIA payout ratio increases.

Table 36. Final Payout Ratio Model Summary – Building Age Included

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.387[a]	.150	.139	.20814

a. Predictors: (Constant), $Coast^{0.25}$, $Age^{0.25}$
b. Dependent Variable: $Ratio^{0.25}$

Table 37. Final Payout Ratio Model ANOVA - Building Age Included

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.191	2	.595	13.743	.000[b]
Residual	6.759	156	.043		
Total	7.949	158			

Table 38. Coefficients of Payout Ratio Model – Building Age Included

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	.538	.085		6.333	.000		
$Age^{0.25}$	.078	.031	.186	2.498	.014	.983	1.017
$Coast^{0.25}$	-.018	.004	-.316	-4.243	.000	.983	1.017

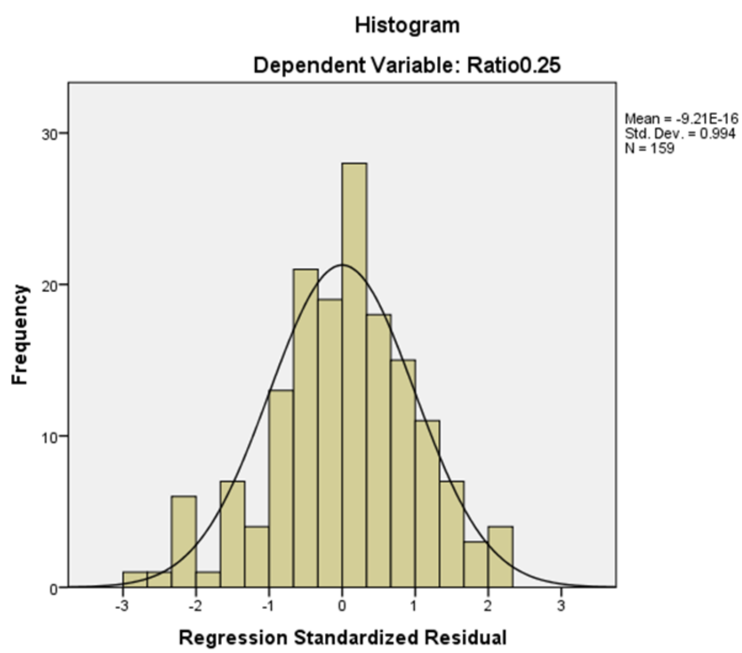

Figure 42. Histogram Plot of Standardize Residual for Payout Ratio Model – Building Age Included

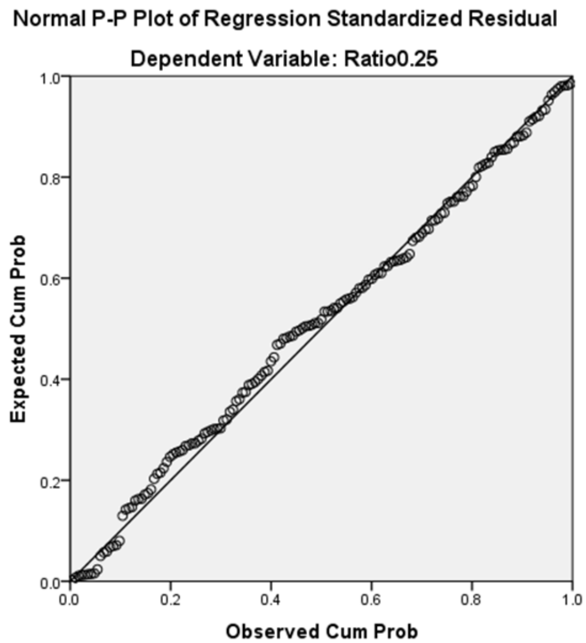

Figure 43. Normal Probability Plot of Standardized Residual Plots - Building Age Included

Table 39. Test for Normality - Building Age Included

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Studentized Residual	.059	159	.200*	.985	159	.074

Table 40. Payout Ratio Model w/ Age - Using Robust Standard Errors

Model	Coefficients	Robust Std. Error	t	Sig.	R Square
(Constant)	.537	.075	7.2	.000	
$Age^{0.25}$	.078	.027	2.86	.005	0.1495
$Coast^{0.25}$	-.018	.004	-4.08	.000	

CHAPTER V

CONCLUSIONS

As hurricanes cause enormous loss to life and property worldwide, there have been significant improvements in predicting, tracking, and warning the public of hurricane events in recent decades. Nevertheless, relatively little progress has been made in the ability to predict hurricane-induced damage. Predicting the damage caused by hurricane and figuring out what factors are responsible for the damage are important. Multiple linear regression models are utilized to predict a hurricane – induced TWIA payout and payout ratio with independent variables that could affect the hurricane intensity, including weather, geographic, and building variables.

Results and Interpretation

Correlation Results

The Person's Correlation results show the correlation between dependent and independent variables.

In terms of TWIA payout, distance from the coastline, distance from the landfall center, and building area, are significantly correlated to the TWIA payout at the 0.01 level.

Distance from the coastline and distance from the landfall center have negative relations with the TWIA payout. Building area has a positive relation with the TWIA payout. Moreover, the improvement value is correlated to the TWIA payout at the 0.05 level. As the improvement value increases, the TWIA payout increases.

For TWIA payout ratio, distance from the coastline is correlated to the TWIA payout ratio at the level of 0.01 and distance from the landfall center is correlated to the TWIA payout ratio at the 0.05 level. These two variables have negative relations to the TWIA payout ratio. As the distance from the coastline or landfall center increases, the TWIA payout ratio decreases.

Regression Results

For the regression model of TWIA payout, we have a model with an Adjusted R Square of 0.264, which means 26.4 percent of variability in transformed TWIA payout could be explained by this model. There are two independent variables in this model and they are significant at the level of 0.01. Two independent variables are distance from coastline and building area. The model is as follows.

TWIA Payout = $[9.421 - 0.348 * (\text{Distance from coastline})^{0.25} + 0.979 * (\text{Building Area})^{0.25}]^{(1/0.25)}$

For the regression model of TWIA payout ratio, we have a model with a R Square of 0.121, which means that 12.1 percent of variability in transformed TWIA payout ratio could be explained by this model. There are two independent variables in this model and they are significant at the level of 0.01. Two independent variables are distance from coastline and No. of building. The model is as follows.

TWIA Payout Ratio= [0.812- 0.018 * (Distance from coastline) 0.25 – 0.099 * (No. of Building)] $^{(1/0.25)}$

Additionally, another model with building age is also established for the TWIA payout ratio when there is only one property per parcel. We have a model with a R Square of 0.150, which means that 15 percent of variability in transformed TWIA payout ratio could be explained by this model. There are two independent variables in this model and they are significant at the level of 0.01. These two independent variables are distance from coastline and building age. The model is as follows.

TWIA Payout Ratio = [0.538 - 0.018 * (Distance from coastline) $^{0.25}$ **+ 0.078 * (Building age)** $^{0.25}$**]** $^{(1/0.25)}$

Recommendations

For the regression models shown above, the R Squares are relatively small. This phenomenon indicates that there are other variables that are responsible for the TWIA payout or TWIA payout ratio. For instance, storm surge and building structure might be significant factors causing building damage. Therefore, future studies should take these factors into consideration.

Moreover, the study uses the multiple linear regression method for the model. Future studies could use different models such as non-linear regression model or neural networks to obtain better models.

Finally, the study uses the Hurricane Ike and Galveston County as the study subject. It is the limitation of this study. However, the study process is highly replicable. Other hurricane and hurricane areas could be used to study the hurricane-induced damage in the future.

REFERENCES

Boswell, M. R., Deyle, R. E., Smith, R. A., & Baker, E. J. (1999). A Quantitative Method for Estimating Probable Public Costs of Hurricanes. *Environment Management*, 23(3), 359-372.

Burton, C. G. (2010). Social vulnerability and hurricane impact modeling. *Natural Hazards Review*, *11*(2), 58-68.

Friedman, D. G. (1984). Natural hazard risk assessment for an insurance program. *The Geneva Papers on Risk and Insurance, 9*(30), 57-128.

Highfield, W. E., Peacock, W. G., & Van Zandt, S. (2010). Determinants & Characteristics of Damage in Single-Family Island Households from Hurricane Ike. *Association of Collegiate Schools of Planning Conference, Minneapolis, MN*

Landsea, C. W., Pielke Jr, R. A., Mestas-Nunez, A. M., & Knaff, J. A. (1999). Atlantic basin hurricanes: Indices of climatic changes. *Climatic Change*, *42*(1), 89-129.

Zane, D. F., Bayleyegn, T. M., Haywood, T. L., Wiltz-Beckham, D., Guidry, H., Sanchez, C., & Wolkin, A. F. (2010). Community Assessment for Public Health Emergency Response following Hurricane Ike—Texas, 25-30 September 2008. *Prehospital and Disaster Medicine*, 25(06), 503-510.