Nat. Hazards Earth Syst. Sci. Discuss., 1, 3813–3855, 2013
www.nat-hazards-earth-syst-sci-discuss.net/1/3813/2013/
doi:10.5194/nhessd-1-3813-2013

Natural Hazards and Earth System Sciences
Discussions
Open Access

Predicting the Texas Windstorm Insurance Association claim payout of commercial buildings from Hurricane Ike

J. M. Kim[1], P. K. Woods[1], Y. J. Park[2], and K. Son[3]

[1]Texas A&M University, Department of Construction Science, College Station, TX, USA
[2]Korea Military Academy, Department of Construction Engineering and Environmental Sciences, Seoul, Republic of Korea
[3]University of Ulsan, School of Architecture, Ulsan, Republic of Korea

Received: 10 June 2013 – Accepted: 9 July 2013 – Published: 6 August 2013

Correspondence to: K. Son (sky9852111@ulsan.ac.kr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Following growing public awareness of the danger from hurricanes and tremendous demands for analysis of loss, many researchers have conducted studies to develop hurricane damage analysis methods. Although researchers have identified the significant indicators, there currently is no comprehensive research for identifying the relationship among the vulnerabilities, natural disasters, and economic losses associated with individual buildings. To address this lack of research, this study will identify vulnerabilities and hurricane indicators, develop metrics to measure the influence of economic losses from hurricanes, and visualize the spatial distribution of vulnerability to evaluate overall hurricane damage. This paper has utilized the Geographic Information System to facilitate collecting and managing data, and has combined vulnerability factors to assess the financial losses suffered by Texas coastal counties. A multiple linear regression method has been applied to develop hurricane economic damage predicting models. To reflect the pecuniary loss, insured loss payment was used as the dependent variable to predict the actual financial damage. Geographical vulnerability indicators, built environment vulnerability indicators, and hurricane indicators were all used as independent variables. Accordingly, the models and findings may possibly provide vital references for government agencies, emergency planners, and insurance companies hoping to predict hurricane damage.

1 Introduction

1.1 Escating demand for natural diaster damage predicion

Natural disasters in the US have been increasing because abnormal weather and climate change have stimulated severe weather events. Increased populations in seaside areas and cities have become vulnerable to widespread risks including danger from cyclones, hurricanes, deluges, and even tsunamis (Pielke Jr. and Landsea, 1998).

NHESSD

1, 3813–3855, 2013

Furthermore, this rapid increase in disaster events has caused unavoidable damage to property and infrastructure during the past five decades. In a brief evaluation, direct losses per year have exceeded \$7.6 billion in the US (Cutter and Emrich, 2005). This estimate does not cover indirect costs such as insurance compensation from the US government or indirect costs to companies and individuals. Moreover, Hurricane Andrew, in August 1992, created insured losses of \$150 million in a single event (Boissonnade and Ulrich, 1995). Not only has the US suffered significant losses, it also has spent a tremendous amount of money on restoration: \$150 billion between 2004 and 2005 alone (Pielke Jr. et al., 2008).

Although a number of communities have recognized the seriousness of the damage and will spend their budgets on mitigation plans, the core problem is how and where to invest their limited funds to prevent and prepare for natural disasters. Therefore, research in this area may help analyze the damage suffered and reduce future monetary loss. Although damage is inescapable, creating damage prediction models can provide a key solution for decreasing these losses.

Following a growing public awareness of the danger from disasters and the tremendous demand for damage prediction, many researchers have conducted studies to develop natural disaster damage prediction methods. Nevertheless, their research has not comprehensively identified the interrelationships among the vulnerabilities, natural disasters, and economic losses of commercial buildings. Consequently, this research will fill this gap in hurricane damage prediction using Hurricane Ike in Texas's coastal counties as a case study.

1.2 Hurricane Ike and Texas Windstorm Insurance Association

Hurricane Ike was a critical disaster which began on 1 September 2008 and ended on 14 September 2008; the storm struck the Bahamas, Cuba, and the Gulf Coast of the US (i.e. Florida, Louisiana, and Texas), in that order. The hurricane formed on the African coast as a tropical depression and became a hurricane when it traveled through the eastern Caribbean Sea. After that, the storm arrived at Cuba and the Bahamas as

NHESSD
1, 3813–3855, 2013
Predicting claim payout of buildings
J. M. Kim et al.

a Category 4 hurricane on the Saffir–Simpson Scale. By the time Ike hit the coastlines of Louisiana and Texas, it had become a Category 2 storm with a central pressure of 950 mb and a maximum wind speed of 95 knots (Berg, 2009). Due to its abnormally large size, Hurricane Ike impacted a wide area, accompanied by strong winds and heavy rainfall which created huge waves and extensive surges. This impact caused fatalities and substantial damage to properties along the hurricane's path (Kennedy et al., 2010). Particularly, the hurricane directly hit the Bolivar Peninsula and Galveston Island in Texas and devastated properties in those areas with severe storm surges and waves. The hurricane was recorded as the third costliest hurricane to strike the mainland of the US, following hurricanes Katrina and Andrew. In Arkansas, Louisiana, and Texas, the estimated total monetary loss was approximately $24.9 billion with twenty human casualties (Berg, 2009).

The Texas Windstorm Insurance Association (TWIA) was established in 1971 to shield insurance policy holders in Texas coastal counties from unexpected meteorological catastrophes. This association is made up of a group of windstorm insurance companies that cover direct loss of property, indirect loss of property or income, and casualties suffered in the Texas coastal counties. TWIA not only provides hurricane protection and training for agents and policy holders, but also receives insurance premiums and makes payments for acceptable claims.

1.3 Research objectives and methods

The objectives of this research are: (1) to identify the relationships among hurricane damage loss, vulnerability indicators, and hurricane indicators for commercial buildings, (2) to predict hurricane damage by vulnerability factors and hurricane indicators, based on insured loss payments for the Texas coastal counties, (3) to decide the magnitude and significance of the indicators, and (4) to create a methodical process using Geographical Information Systems (GIS) to assess other times and states in order to predict hurricane damage. These factors provide the framework necessary to identifying the spatial distribution of financial hurricane loss.

Figure 1 shows the outline of the data collection process used for this research. First, the TWIA claim payout properties were mapped within the study area using the ArcGIS address locator. Second, sample payouts were randomly selected. Third, geographical vulnerabilities, building environment vulnerabilities, and hurricane indicators were combined, respectively, with the TWIA claim payouts by joining them with the data obtained from ArcGIS by using the Join Data function. Finally, regression models were generated and analyzed.

After the creation of the data, a multiple linear regression method was applied to analyze the data, which resulted in two global equations that allowed for an understanding of the relationship between the dependent and independent variables. The global model assumes that the relationships are fixed and coherent throughout all of the data. This study identified the interrelationships among the vulnerability indicators and TWIA claim payouts using a statistical method. The statistical method order is listed below.

(1) Descriptive statistics: mean, max, min, median, and standard deviation.

(2) Scatter plots: to check the relationships among the dependent and independent variables.

(3) Correlation test: Pearson's and Spearman's Tests to check the relationships among the variables.

(4) Multi-collinearity analysis: to check the correlations among the variables.

(5) ANOVA test and linear regression: to check the significance of the regression model.

(6) Test of normality: to check the normality of the data.

(7) Test of homoscedasticity: to use residual plots to check the variance of errors.

NHESSD
1, 3813–3855, 2013

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

(8) Transformation: to use log transformation analysis, if required.

(9) A regression model.

2 Data collection and management

2.1 Dependent variable

This study considered as observational units only improved commercial buildings that had insured claim payouts from the TWIA in Texas coastal counties from Hurricane Ike. As shown in Fig. 2, Hurricane Ike, a Category 2 hurricane on the Saffir–Simpson Scale, struck the Texas coastal counties on 13 September 2008. The financial damages suffered by Texas coastal counties are shown in Fig. 3.

Table 1, Figs. 4 and 5 show the total amount of claim payouts and the number of claim payouts collected from the TWIA for commercial property damage from Hurricane Ike from 17 August 2008 to 22 February 2012.

The total claim payout was \$450 518 330 and the total number of claims was 4150. Galveston County received the most damage from Hurricane Ike in terms of both dollar amount of damage (\$255 333 818; 56.68 %) and the number of claims (1807; 43.54 %). Other damaged counties included: Jefferson County (1218 claims totaling \$104 249 917); Brazoria County (597 claims totaling \$46 922 396); Chambers County (470 claims totaling \$39 755 609); Harris County (45 claims totaling \$4 126 821); Matagorda County (9 claims totaling \$36 981); Liberty County (2 claims totaling \$67 501); and Nueces County (2 claims totaling \$5,287).

In this study, 500 of the total damage reports (4150) were randomly selected as samples. The sample size needed to be larger than 370, which is determined when the size of a population is 5000 with a 95 % confidence level and a ±5 % precision level (Israel, 1992).

2.2 Explanatory variables

2.2.1 Geographical vulnerability and indicators

Geographical vulnerability is defined as a substantial exposure to peril (Cutter, 1996). Since vulnerability is an essential feature of natural disasters, it can be explained by biophysical risks such as elevation and other geographical impacts (Cutter et al., 2003). In general, geographical features differ depending on the location, and the level and amount of exposure to natural hazards can also be diverse. For instance, the Federal Emergency Management Agency (FEMA) created the FEMA Q3 Flood Data study in an effort to understand the risks of hurricanes and floods. FEMA designated flood zones based on the level of flood risk (Howard and Scott, 2005). The zones show the potential risk of flood in each defined area. As shown in Table 2, there are three types of flood zones. Zone A is an area anticipated to have a 1 %, or larger chance to flood in any given year. Zone X500 is an area anticipated to have a 0.2–1 % chance to flood in any given year. Zone X is an area anticipated to have a 0.2 % or smaller chance to flood in any given year. Although floods can occur anywhere, flood prone areas exist. Based on historical flood data, geographical vulnerability presents flood prone areas.

The National Weather Service created a five point scale to represent the hurricane surge zone in an effort to help clarify the dangers of hurricanes in coastal areas. As shown in Table 3, the categories created are based on sustained wind speed and surge height. Each scaled area is predicted to be influenced by a defined category called the Hurricane Surge Zone. This scale not only presents hurricane risks in scaled areas, but also compares the geographical vulnerability of each area.

The distance from a building to the water also plays a key role in defining geographical vulnerability. Highfield et al. (2010) measured the distance from a building to the water to assess the damage to Galveston Island and Bolivar Peninsula caused by Hurricane Ike. They found that the damage increased as the distance from the water decreased (Highfield et al., 2010). These findings indicated that areas closer to water have more geographical vulnerability than areas further from water.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Accordingly, geographical vulnerability indicators should be considered in hurricane damage prediction. FEMA Flood Zones, Hurricane Surge Zones, and distance from water should all be integrated into the hurricane damage prediction model as geographical vulnerability indicators.

2.2.2 Built environment vulnerability and indicators

Natural disasters have a tremendous impact on both people and property, and the level of exposure to the disaster determines the magnitude of the damage. Therefore, insurers must estimate the vulnerability of an insured built environment to measure the likelihood of economic loss (Khanduri and Morrow, 2003). On a large scale, for instance, water-related infrastructure systems such as dams, seawalls, and dikes are constructed in flood and hurricane-prone areas, and play a prominent role in preventing damage from natural disasters (Brody et al., 2008). On a smaller scale, the building features of each building such as building age, building floor area, and appraised value of the building are important components of natural exposure (Chock, 2005; Dehring and Halek, 2006; Highfield et al., 2010; Khanduri and Morrow, 2003). Highfield et al. (2010) used building age to assess the damage to Galveston Island and Bolivar Peninsula from Hurricane Ike. They found that the damage increased as the building age increased (Highfield et al., 2010). Dehring and Halek (2006) used building floor area to assess the residential property damage from Hurricane Charley in Lee County. These researchers revealed that the damage increased as the building floor area increased (Dehring and Halek, 2006). The research implies that the building's features decide the level of vulnerability, because each building can be classified by combining the characteristics of the buildings to determine the amount of damage and exposure (Chock, 2005).

Consequently, quantifying built environment vulnerabilities are important for assessing the damage caused by natural disasters; built environment vulnerability indicators (e.g. building age, building floor area, and appraised value of the building) should be included in the hurricane damage prediction model.

2.3 Hurricane assessment and indicators

Every year, hurricanes impact large areas and frequently affect both people and property. Numerous parameters of hurricanes can act as key factors contributing to the amount of damage sustained, such as frequency, magnitude, and others. For example, wind parameters play a key role in hurricane damage and cause related disasters such as floods, hurricane surges, and landslides.

The Hurricane Research Division (HRD) of the National Oceanic and Atmospheric Administration (NOAA) created the HRD real-time hurricane wind analysis system (H*Wind) to make an integrated hurricane observation system. The HRD collects measured wind data from meteorological observing stations every four to six hours during hurricanes and integrates the data into a wind field which contains information such as maximum sustained wind speeds, duration and direction of maximum sustained wind speeds, and wind direction steadiness (Dunion et al., 2003; Powell and Houston, 1998; Powell et al., 2010). This wind analysis utilizes the information gathered by measuring a hurricane's intensity, and thus improves upon earlier hurricane wind analyses. H*Wind analyses include gridded data, image data, and Geographical Information System (GIS) shape files. Researchers can use the H*Wind analyses to assess both wind and storm surges. Additionally, the swath map can be useful for hurricane damage assessments (Dunion et al., 2003; Powell and Houston, 1998; Powell et al., 1998).

The map also includes gridded data, image data, and Geographical Information System (GIS) shape files. As shown in Fig. 6, the swath map consists of grids. Each grid has location information (i.e. longitude and latitude) and wind measurements (i.e. maximum sustained wind speeds, duration and direction of maximum sustained wind speeds, and wind direction steadiness). Using the location information and the wind measurements, researchers should be able to plot the wind database based on their interest time, area, and particular hurricane, and be able to study the relationship between the hurricane's damage and wind (Burton, 2010; Powell et al., 1998).

NHESSD
1, 3813–3855, 2013

The side of a hurricane also plays an important role in measuring damage. In the Northern Hemisphere, areas located on the right side of a hurricane track usually sustain more damage than the left side of a hurricane track (Keim et al., 2007; Noel et al., 1995). The difference occurs because of the differences in wind intensity and direction on either side, due to the interaction of the two opposing actions of a hurricane (i.e. forward movement and counterclockwise rotation). As a result of the interaction, the areas located on the right side of the hurricane always face stronger and more extensive winds, and therefore becomes prone to a greater level of hurricane damage. Hence, the right side of the hurricane track is significantly more exposed to damage than the left side of the hurricane track.

As a consequence, hurricane indicators should be considered in damage predictions, and H*Wind analyses and the side of the hurricane track an area falls on should also be integrated into the hurricane damage prediction model as hurricane indicators.

2.4 Regression model

In this study, two statistical models were generated to predict the hurricane damage caused by Hurricane Ike in Texas coastal counties for commercial buildings. The goal of this model is to predict the insured claim payout. The dependent variable, the Texas Windstorm Insurance Association (TWIA) claim payout ($), can be predicted by the independent variables, as shown in Eq. (1)

$$\begin{aligned} \mathrm{PDL} = {} & \beta_0 + \beta_1 \cdot \ \mathrm{Wind\_Speed} \ + \beta_2 \cdot \ \mathrm{Side\_Right} \ + \beta_3 \cdot \ \mathrm{Age} \ + \beta_4 \cdot \ \mathrm{Area} \\ & + \beta_5 \cdot \ \mathrm{Imp\_Value} \ + \beta_6 \cdot \ \mathrm{FEMA\_Zones} \ + \beta_7 \cdot \ \mathrm{Surge\_Zones} \\ & + \beta_9 \cdot \ \mathrm{Dist\_Shore}. \end{aligned} \tag{1}$$

2.5 Data management

This study utilized GIS to combine, manage, and create spatial information for a statistical examination. As a computerized database management system, GIS facilitates

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

spatial data to store, capture, control, make, analyze, and present geographically referenced data (Bill, 1994). Generally, spatial data presents the figure and position of the data by layers using raster data, digitally imaged grid data, and vector data, based on polygons, points, and lines, respectively (Hellawell et al., 2001). The primary benefit of using this application is in creating a new layer of data by using various useful functions such as merge, clip, union, intersection, join, buffer, overlay, and dissolve. Particularly, this research produced a new layer of data by using the overlay function to combine diverse sorts of obtained data from the related organizations, based on their locations.

Figure 7 presents an outline of the GIS process. This research utilized ArcGIS tools to combine both a dependent variable and independent variables. After the GIS process, data collection was completed as shown in Table 8. The process described below explains the GIS process:

(1) The TWIA claim payout properties were mapped in the study area using the ArcGIS address locator.

(2) Geographical vulnerability indicators, building environment vulnerability indicators, and hurricane indicators were joined with the TWIA claim payouts by joining the data of with ArcGIS.

(3) The data was completed for the regression models.

2.6 Descriptive analysis

Descriptive statistics present important properties such as number of samples, mean, median, standard deviation, quartiles, skewness, and kurtosis. Table 4 numerically shows the descriptive statistics of the dependents and independent variables used in this study. The mean and median present the central tendency of the data. The standard deviations measure the spread of the samples. The quartiles show the dispersion of data, and the skewness and kurtosis describe the distribution shape. In accordance with the skewness, the distribution of the PDL are excessively skewed to the right. The values, 2.61, higher than 0, indicate that the distribution is positively skewed (i.e. that the left of the tail is shorter than the right side of the tail, and the data distribution is left sided). According to the kurtosis, the distribution of the PDL is leptokurtic, which indicates higher and sharper peaks than a normal distribution. The values, 9.41, higher than 3, mean that the data is not normally distributed.

3 Correlation between claim payout and variables

Table 5 shows the summary of the correlation results with the TWIA claim payouts and continuous variables. A Pearson Correlation analysis was used for testing the continuous variables. Each result represents the relationship between two variables. The building age has only an insignificant relationship with the claim payout. On the other hand, other variables (i.e. max. sustained wind speed, building floor area, appraised value of the building, and distance from the property centroid to the shoreline) have significant relationships with the claim payout. The sign of the coefficients determine whether the relationship is positive or negative, and the coefficients indicate the amount of the linear relationship with a range of +1 to −1.

Table 6 displays the summary of the correlation results with the TWIA claim payout and the ordinal variables. Spearman's rho Correlation analysis was used to test the ordinal variables. Each result represents the relationship between two variables. The right side of the hurricane track has only an insignificant relationship with the claim payout, while the FEMA flood zones and hurricane surge zones each have significant relationships with the claim payout. The sign of the coefficients determines whether the relationship is positive or negative, and the coefficients indicate the amount of the linear relationship with a range of +1 to −1.

NHESSD
1, 3813–3855, 2013

4 Diagnostics for residuals and transformation

The Kolmogorov–Smirnov value was adopted to test for the normality of the residuals. In Table 7, the p value of 0.000 is smaller than 0.05, which implies that the residuals are not normally distributed. Moreover, in Fig. 8a and b, the standardized residuals histogram and the Q–Q plot also verify that the initial model's residuals are not normally distributed.

The residual plot tested whether the residuals have the constant variance to check for homoscedasticity, as shown in Fig. 9. The fan-shaped residuals plot determined that the residuals have demonstrated a trend (i.e. that there is no dispersion based on the regression line). This means that the residuals' variance is not constant. In conclusion, these results, the residuals analyses, and the test all prove that the dependent variable needed a transformation.

The TWIA claim payout was transformed by a natural log. The transformed dependent variable is as follows:

$$\text{Transformed PDL} = \log(\text{TWIA claim payout (\$)}).$$

After the log transformation of the dependent variable, the Kolmogorov–Smirnov value shows that the transformed model's residuals are normally distributed because the P value of 0.200 is higher than 0.05, as seen in Table 7.

Moreover, the standardized residuals histogram and the Q–Q plot also confirm that the transformed model's residuals are normally distributed, as shown in Fig. 10. The residual plot checks the homoscedasticity, as shown in Fig. 11. The residuals are randomly spread without any systematic patterns. This represents that the residuals' variance is constant.

The backward elimination method was used to find the best-fit regression model. Table 9 includes a summary of the transformed TWIA claim payout regression model. The model is statistically significant because the P value of 0.000 is less than 0.05, which represents that independent variables and the dependent variable have a significant

NHESSD
1, 3813–3855, 2013

linear relationship. Also, the null hypothesis which states that there is no linear relationship between the independent variables and the dependent variable can be rejected. Thus, the regression model is allowed to predict the transformed dependent variable.

The adjusted R-square of 0.401 indicates that the transformed dependent variable can be explained with 40.1 % of variability by the significant variables (i.e. max. sustained wind speed, the right side of the hurricane track, building age, building floor area, appraised value of building, hurricane surge zones, and distance from the property centroid to shoreline). On the other hand, this study disregards the rest of the variability of 59.9 %. The remainder could be explained by some unidentified variables.

Table 10 illustrates a summary of the coefficients for the transformed TWIA claim payout regression model. The seven significant predictors include: (1) max. sustained wind speed, (2) the right side of the hurricane track, (3) building age, (4) building floor area, (5) appraised value of the building, (6) hurricane surge zone, and (7) distance from the property centroid to the shoreline; each were identified as able to predict the transformed claim payout. The FEMA flood zones, however, were eliminated because the *P* value was higher than 0.10. The Variance Inflation Factor (VIF) ranged from 1.130 to 2.208. These values verify that the individual predictors have no multicollinearity, which means that the predictors are not correlated with each other.

The beta coefficients, also called the standardized coefficients, were used to determine which independent variables have a significant influence on the claim payout; they ranged from 0 to 1, reflecting when the variables have different units. Following the amount of the coefficients, the rank was listed in sequence: (1) the appraised value of building, (2) building age, (3) hurricane surge zone, (4) building floor area, (5) right side of the hurricane track, (6) maximum sustained wind speed, and (7) distance from property centroid to the shoreline.

Based on the unstandardized coefficients, a multiple linear regression model was created with seven predictors to predict the transformed claim payout, as shown in Eqs. (2) and (3). The model can explain a 40.9 % variability of the transformed

NHESSD
1, 3813–3855, 2013
Predicting claim payout of buildings
J. M. Kim et al.

dependent variable

$$\log(\text{Predicted TWIA claim payout (\$)}) = 2.973 + (\text{Wind\_Speed} \times 0.019) + (\text{Side\_Right} \times 0.1) + (\text{Age} \times 0.007) + (\text{Area} \times 2.522 \times 10^{-4}) + (\text{Imp\_Value} \times 1.526 \times 10^{-6}) + (\text{Surge\_Zones} \times -0.111) + (\text{Dist\_Shore} \times -5.254 \times 10^{-6}). \quad (2)$$

$$\text{Predicted TWIA claim payout (\$)} = \exp(2.973 + (\text{Wind\_Speed} \times 0.019) + (\text{Side\_Right} \times 0.1) + (\text{Age} \times 0.007) + (\text{Area} \times 2.522 \times 10^{-4}) + (\text{Imp\_Value} \times 1.526 \times 10^{-6}) + (\text{Surge\_Zones} \times -0.111) + (\text{Dist\_Shore} \times -5.254 \times 10^{-6})). \quad (3)$$

Based on Eq. (2), the interpretation of the unstandardized coefficients in the regression model are as follows:

(1) β_1 is 0.019 which implies that if the maximum sustained wind speed increases by $1\,\mathrm{m\,s^{-1}}$, the log transformed claim payout increases by 1.9 %.

(2) β_2 is 0.100 which implies that if a building is located on the right side of the hurricane track, the log transformed claim payout increases by 10 %.

(3) β_3 is 0.007 which implies that if the building age increases by 1, the log transformed claim payout increases by 0.7 %.

(4) β_4 is 2.522×10^{-4} which implies that if the building floor area increases by $1\,\mathrm{m^2}$, the log transformed claim payout increases by 0.025 %.

(5) β_5 is 1.526×10^{-6} which implies that if the appraised value of the building increases by \$1, the log transformed claim payout increases by 0.00015 %.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◂ ▸
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

(6) β_6 is −0.111 which implies that if the hurricane surge zone number increases by 1, the log transformed claim payout decreases by −11.1 %.

(7) β_7 is -5.254×10^{-6} which implies that if the distance from the property centroid to the shoreline increases by 1, the log transformed claim payout decreases by −0.0005254 %.

5 Transformed model and validity

In this study, the backward elimination method was utilized to find the best-fit multiple linear regression model and to identify the significant predicators. In The TWIA claim payout regression, seven indicators were seen to be significant as predicators of the transformed dependent variable. The range of the Variance Inflation Factor (VIF), from 1.130 to 2.208, also confirms that the individual predictors have no multicollinearity, which verifies that the predictors are not correlated with each other. The model's adjusted R^2 of 0.401 indicates that the transformed dependent variable can be explained with 40.1 % of variability by the significant independent variables. Figure 12 shows a scatter plot of the actual log-transformed TWIA claim payout versus the predicted log TWIA claim payout.

6 Summary and conclusions

With growing public awareness of hurricane danger and with tremendous demands for damage analysis, many researchers have conducted studies to develop hurricane damage prediction methods. However, to date there has been no comprehensive research directed towards identifying the relationships among vulnerabilities, hurricanes, and the economic loss of individual commercial buildings. To fill this gap, this research has identified vulnerability indicators and hurricane indicators, developed metrics to

NHESSD
1, 3813–3855, 2013

measure the influence of economic losses from hurricanes, and visualized the spatial distribution of vulnerability to evaluate overall hurricane damage.

In this research, TWIA claim payouts from Hurricane Ike were used as the dependent variable to predict the actual financial damage and to decide the magnitude and significance of the indicators. Geographical vulnerability indicators, built environment vulnerability indicators, and hurricane indicators were used as independent variables.

The models and findings produced in this study could provide vital references for government agencies, emergency planners, and insurance companies seeking to predict hurricane damage. This research may help analyze damage and reduce financial loss. Moreover, this study defines hurricane-prone areas and the distribution of hurricane losses in an effort to reduce the perceived risks for residents who live in hurricane vulnerable areas.

This study considered improved commercial buildings in Texas coastal counties that had received insured claim payouts from the Texas Windstorm Insurance Association (TWIA) resulting from Hurricane Ike. The observational unit ranged from 17 August 2008 to 22 February 2012.

According to the claim payout records, the total claim payout was \$450 518 330 and the total number of claims was 4150. Galveston County received the most damage from Hurricane Ike in both the dollar amount of damage (\$255 333 818; 56.68 %) and the number of claims (1807; 43.54 %). Therefore, we recognized from the distribution of the damages that Galveston county is the most hurricane-prone area in the Texas coastal counties.

The model is statistically significant because the P value of 0.000 is less than 0.05, which means that the independent variables could predict the TWIA claim payout. The adjusted R^2 of 0.401 represents that the 40.1 % of variability in the transformed dependent variable can be explained by the significant variables. Checking the P values reveal seven significant variables: maximum sustained wind speed, the right side of the hurricane track, building age, building floor area, appraised value of the building, hurricane surge zone, and distance from the property centroid to the shoreline. In this phase, the FEMA flood zones were rejected due to the high *P* value. Following the values of the coefficients, the significant variables also measured the magnitude of the dependent variable. Therefore, the claim payout can be measured by using the prediction model, as follows:

In the prediction model,

(1) The maximum sustained wind speed has a positive relationship with the TWIA claim payout, which means that if the maximum sustained wind speed increases, the claim payout increases. This result supports the results of the previous studies that wind speed is a significant indicator of hurricane damages and is useful for predicting hurricane damages (Burton 2010; Dunion et al., 2003; Powell and Houston 1998; Powell et al., 1995, 1998).

(2) The right side of the hurricane track has a positive relationship with the TWIA claim payout, which means that if a building is located on the right side of the hurricane track, the claim payout increases. This result reinforces former studies that a building located on the right side of the hurricane track usually has more damage than one on the left side of the hurricane track, in the Northern Hemisphere (Keim et al., 2007; Noel et al., 1995), and confirms that the variable is a critical indicator for hurricane damage prediction.

(3) Building age has a positive relationship with the TWIA claim payout, which means that if the building age increases, the claim payout also increases. This result proves that the building age is a significant variable for predicting hurricane damage (Highfield et al., 2010).

(4) Building floor area has a positive relationship with the TWIA claim payout, which means that if the building floor area increases, the claim payout increases. This result corroborates previous studies which conclude that this variable is one of the indicators for measuring hurricane damage (Dehring and Halek, 2006).

NHESSD
1, 3813–3855, 2013

(5) Appraised value of the building has a positive relationship with the TWIA claim payout, which means that if this value increases, the claim payout also increases. This result confirms that the appraised value of the building is a significant indictor in assessing the damage from hurricanes.

(6) Hurricane surge zone has a negative relationship with the TWIA claim payout, which means that if the hurricane surge zone number increases, the claim payout decreases. This result verifies that the hurricane surge zone is a useful indicator for predicting hurricane damage.

(7) Distance from the property centroid to the shoreline has a negative relationship with the TWIA claim payout, which means that if the distance increases, the claim payout decreases. This result confirms that the distance is related to the damage and is a significant variable for predicting hurricane damage (Highfield et al., 2010).

7 Future research

The adjusted R^2 value of the claim payout is 0.401, which means that the rest of the variability could be explained by some unidentified variables. Consequently, it would be valuable to come up with prospective indicators and make additions to find the best-fit regression model.

This study only considered improved commercial buildings in Texas coastal counties. The results and findings would likely be different with residential properties. Future studies will need to include residential properties to strengthen the results and findings. In addition, the hurricane damages considered were only those resulting from Hurricane Ike. Therefore, it would be worthwhile to study various other categories of hurricanes in the future.

NHESSD
1, 3813–3855, 2013

Moreover, using the developed methodology and indicators in this study, it should be possible to predict hurricane damage for other hurricane-prone areas such as Florida, South Carolina, North Carolina, Alabama, and Louisiana.

Acknowledgements. This work was supported by the 2013 Research Fund of University of Ulsan.

References

Berg, R.: Tropical Cyclone Report: Hurricane Ike (al092008), 1–14 September 2008, National Hurricane Center, Miami, FL,, 2009.

Bill, R.: Multimedia GIS–Definition, Requirements and Applications, The 1994 European GIS Yearbook, Blackwell, Oxford, 1994.

Boissonnade, A. and Ulrich, P.: How to best use engineering risk analysis models and geographic information systems to assess financial risk from hurricanes, in: Casualty Actuarial Society Discussion Paper Program, Casualty Actuarial Society Discussion Paper Program, Arlington, VA, 179–206, 1995.

Brody, S. D., Zahran, S., Highfield, W. E., Grover, H., and Vedlitz, A.: Identifying the impact of the built environment on flood damage in Texas, Disasters, 32, 1–18, doi:10.1111/j.1467-7717.2007.01024.x, 2008.

Burton, C. G.: Social vulnerability and hurricane impact modeling, Natural Hazards Review, 11, 58–68, doi:10.1061/(ASCE)1527-6988(2010)11:2(58), 2010.

Chock, G. Y. K.: Modeling of hurricane damage for Hawaii residential construction, J. Wind Eng. Ind. Aerod., 93, 603–622, 2005.

Cutter, S. L.: Vulnerability to environmental hazards, Prog. Hum. Geog., 20, 529–539, 20, doi:10.1177/030913259602000407, 1996.

Cutter, S. L. and Emrich, C. T.: Are Natural Hazards and Disaster Losses in the US Increasing?, National Emergency Training Center, Emmitsburg, MD, 2005.

Cutter, S. L., Boruff, B. J., and Shirley, W. L.: Social vulnerability to environmental hazards, Soc. Sci. Quart., 84, 242–261, 2003.

Dehring, C. A. and Halek, M.: Do Coastal Building Codes Mitigate Hurricane Damage to Residential Property?, Property, Land Use & Real Estate Law eJournal 09/2006, 40 pp., doi:10.2139/ssrn.928009, 2006.

Dunion, J. P., Landsea, C. W., Houston, S. H., and Powell, M. D.: A reanalysis of the surface winds for Hurricane Donna of 1960, Mon. Weather Rev., 131, 1992–2011, 2003.
Hellawell, E., Kemp, A., and Nancarrow, D.: A GIS raster technique to optimise contaminated soil removal, Eng. Geol., 60, 107–116, 2001.
Highfield, W. E., Peacock, W. G., and Van Zandt, S.: Determinants and Characteristics of Damage in Single-Family Island Households from Hurricane Ike 1, The Association of Collegiate Schools of Planning Conference, Minneapolis, Minnesota, 2010.
Howard, J. and Scott, D. M.: Risk Management Method, Google Patents, US 20050177529 A1, 2005.
Israel, G. D.: Determining Sample Size, Florida Cooperative Extension Service, University of Florida, Gainesville, Florida, 1992
Keim, B. D., Muller, R. A., and Stone, G. W.: Spatiotemporal patterns and return periods of tropical storm and hurricane strikes from Texas to Maine, J. Climate, 20, 3498–3509, 2007.
Kennedy, A., Rogers, S., Sallenger, A., Gravois, U., Zachry, B., Dosa, M., and Zarama, F.: Building destruction from waves and surge on the Bolivar Peninsula during Hurricane Ike, J. Waterw. Port C-ASCE, 137, 132–141, 2010.
Khanduri, A. and Morrow, G.: Vulnerability of buildings to windstorms and insurance loss estimation, J. Wind Eng. Ind. Aerod., 91, 455–467, 2003.
Noel, J. M., Maxwell, A., Platt, W. J., and Pace, L.: Effects of Hurricane Andrew on cypress (*Taxodium distichum* var. nutans) in south Florida, J. Coastal Res., 21, 184–196, 1995.
Pielke Jr., R. A. and Landsea, C. W.: Normalized hurricane damages in the United States: 1925–95, Weather Forecast., 13, 621–631, 1998.
Pielke Jr., R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., and Musulin, R.: Normalized hurricane damage in the United States: 1900–2005, Natural Hazards Review, 9, 29–42, 2008
Powell, M. D. and Houston, S. H.: Surface wind fields of 1995 hurricanes Erin, Opal, Luis, Marilyn, and Roxanne at landfall, Mon. Weather Rev., 126, 1259–1273, 1998.
Powell, M. D., Houston, S. H., Amat, L. R., and Morisseau-Leroy, N.: The HRD real-time hurricane wind analysis system, J. Wind Eng. Ind. Aerod., 77, 53–64, 1998.
Powell, M. D., Murillo, S., Dodge, P., Uhlhorn, E., Gamache, J., Cardone, V., Cox, A., Otero, S., Carrasco, N., and Annane, B.: Reconstruction of Hurricane Katrina's wind fields for storm surge and wave hindcasting, Ocean Eng., 37, 26–36, 2010.

Table 1. TWIA claim payout records per County from Hurricane Ike.

County	Total claim payouts ($)	No. of claim payouts
Galveston	255 333 818	1807
Jefferson	104 249 917	1218
Brazoria	46 922 396	597
Chambers	39 755 609	470
Harris	4 126 821	45
Matagorda	36 981	9
Liberty	67 501	2
Nueces	5287	2
SUM	450 518 330	4150

NHESSD
1, 3813–3855, 2013

Table 2. Definition of FEMA flood zone.

Zone	Explanation
A	Areas have a 1 %, or larger, chance to flood on any given year.
X500	Areas have a 0.2–1 % chance to flood on any given year.
X	Areas have a 0.2 %, or smaller, chance to flood on any given year.

Source: http://www.fema.gov/

NHESSD
1, 3813–3855, 2013

Table 3. Definition of Hurricane surge zone.

Hurricane surge zone	Wind speed (mph)	Surge height (ft)
5	74 ~ 95	4 ~ 5
4	96 ~ 110	6 ~ 8
3	111 ~ 129	9 ~ 12
2	130 ~ 156	13 ~ 18
1	> 157	> 18

NHESSD

1, 3813–3855, 2013

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

Table 4. Descriptive statistics.

		PDL ($10 000)	Max. sustained wind speed ($\mathrm{m\,s^{-1}}$)	Right side of the hurricane track	Building age	Building floor area ($100\,\mathrm{m^2}$)	Appraised value of building ($10 000)	FEMA flood zones	Hurricane surge zones	Distance from shoreline (1000 m)
N		500	500	500	500	500	500	500	500	500
Mean		1.18	36.17	–	34.32	3.64	15.03	–	–	4.49
Median		0.77	36.00	–	35.00	2.81	11.85	–	–	0.88
Std. deviation		1.22	2.11	–	18.00	2.68	11.72	–	–	6.64
Percentiles	25	0.41	34.84	0.00	23.00	1.90	7.23	1.00	3.00	0.37
	50	0.77	36.00	0.00	35.00	2.81	11.85	2.00	3.00	0.88
	75	1.50	36.74	1.00	47.00	4.55	18.82	3.00	3.75	6.03
Skewness		2.61	0.23	1.13	0.45	1.83	1.83	−0.07	−0.05	1.64
Kurtosis		9.41	0.76	−0.72	1.32	3.89	3.99	−1.58	0.04	1.49

Table 5. Results of Pearson correlation analysis.

		PDL ($)	Wind_Speed ($m\,s^{-1}$)	Age	Area (m^2)	Imp_Value ($)	Dist_Shore (m)
PDL ($)	Pearson Correlation	1	0.203*	0.029	0.400*	0.364*	−0.190*
	Sig. (2-tailed)		0.000	0.512	0.000	0.000	0.000
Wind_Speed ($m\,s^{-1}$)	Pearson Correlation	0.203*	1	0.040	−0.057	0.007	−0.183*
	Sig. (2-tailed)	0.000		0.375	0.199	0.879	0.000
Age	Pearson Correlation	0.029	0.040	1	−0.123*	−0.383*	−0.062
	Sig. (2-tailed)	0.512	0.375		0.006	0.000	0.167
Area (m^2)	Pearson Correlation	0.400*	−0.057	−0.123*	1	0.572*	0.044
	Sig. (2-tailed)	0.000	0.199	0.006		0.000	0.322
Imp_Value ($)	Pearson Correlation	0.364*	0.007	−0.383*	0.572*	1	0.006
	Sig. (2-tailed)	0.000	0.879	0.000	0.000		0.899
Dist_Shore (m)	Pearson Correlation	−0.190*	−0.183*	−0.062	0.044	0.006	1
	Sig. (2-tailed)	0.000	0.000	0.167	0.322	0.899	

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

NHESSD
1, 3813–3855, 2013
Predicting claim payout of buildings
J. M. Kim et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀◀ ▶▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Table 6. Results of Spearman's correlation analysis.

		PDL ($)	FEMA_Zones	Surge_Zones	Side_Right
PDL ($)	Spearman's rho Correlation	1.000	0.186*	−0.321*	−0.011
	Sig. (2-tailed)	–	0.000	0.000	0.803
FEMA_Zones	Spearman's rho Correlation	0.186*	1.000	−0.521*	−0.243*
	Sig. (2-tailed)	0.000	–	0.000	0.000
Surge_Zones	Spearman's rho Correlation	−0.321*	−0.521*	1.000	0.071
	Sig. (2-tailed)	0.000	0.000	–	0.114
Side_Right	Spearman's rho Correlation	−0.011	−0.243*	0.071	1.000
	Sig. (2-tailed)	0.803	0.000	0.114	–

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

NHESSD
1, 3813–3855, 2013

Table 7. Test of normality for initial TWIA claim payout regression model.

	Kolmogorov–Smirnov[a]			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PDL	0.190	500	0.000	0.734	500	0.000

[a] Lilliefors Significance Correction.

NHESSD
1, 3813–3855, 2013

Table 8. Test of normality for transformed TWIA claim payout regression model.

	Kolmogorov–Smirnov[a]			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Log_PDL	0.028	500	0.200[b]	0.993	500	0.029

[a] Lilliefors Significance Correction.
[b] This is a lower bound of the true significance.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

Table 9. Summary of transformed TWIA claim payout regression model.

Model	Sum of squares	df	Mean square	F	Sig.	R^2	Adj-R^2
Regression	32.628	7	4.661	48.721	0.000	0.409	0.401
Residual	47.071	492	0.096				
Total	79.699	499					

Predictors: (Constant), Dist_Shore, Imp_Value, Wind_Speed, Age, Side_Right, Area, Surge_Zones
Dependent Variable: Log_PDL

NHESSD
1, 3813–3855, 2013

Table 10. Coefficients of transformed TWIA claim payout regression model.

	Model	β	Std. error	Beta	Sig.	VIF
Constant		2.973	0.269		0.000	
Hurricane Indicators						
Max. sustained wind speed	0.019	0.007	0.099	0.007	1.130	
Right side of the hurricane track	0.100	0.039	0.109	0.011	1.506	
Built Environment Vulnerability Indicators						
Building age	0.007	0.001	0.317	0.000	1.246	
Building floor area	0.000	0.000	0.169	0.000	1.537	
Appraised value of building	1.526×10^{-6}	0.000	0.448	0.000	1.808	
Geographical Vulnerability Indicators						
Hurricane surge zones	−0.111	0.017	−0.295	0.000	1.741	
Distance from shoreline	-5.254×10^{-6}	0.000	−0.087	0.090	2.208	

NHESSD
1, 3813–3855, 2013

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

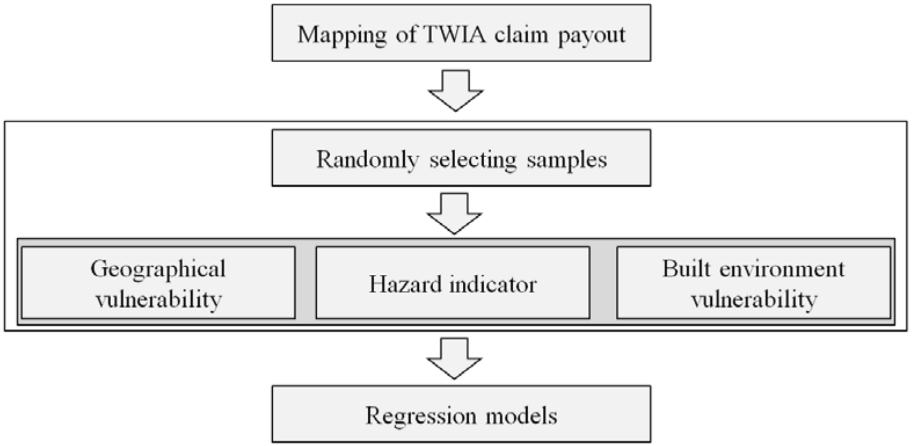


Fig. 1. Data collection process.

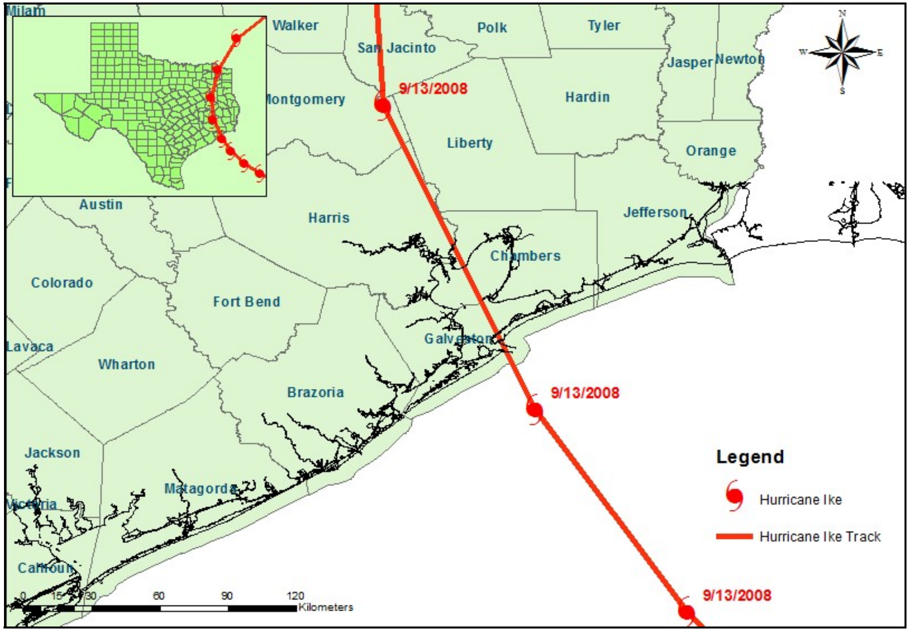


Fig. 2. Hurricane Ike.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

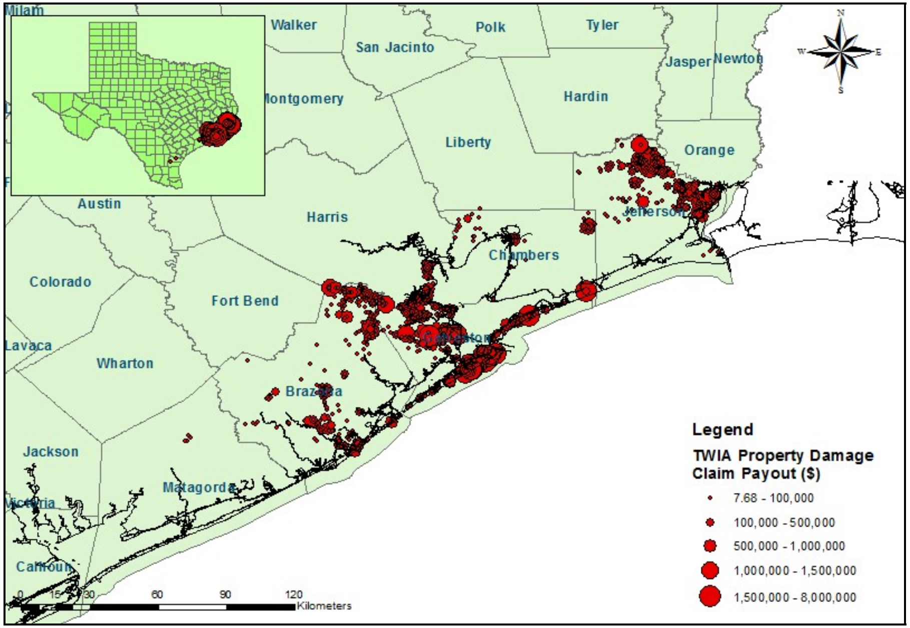


Fig. 3. Distribution of TWIA property claim payouts.

NHESSD
1, 3813–3855, 2013

J. M. Kim et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

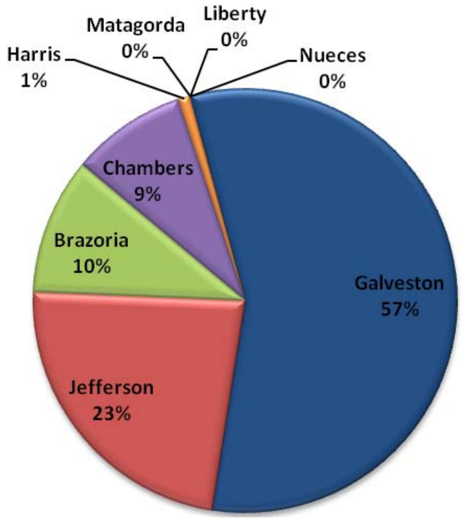


Fig. 4. Distribution of total claim payout amounts ($) per County from Hurricane Ike.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

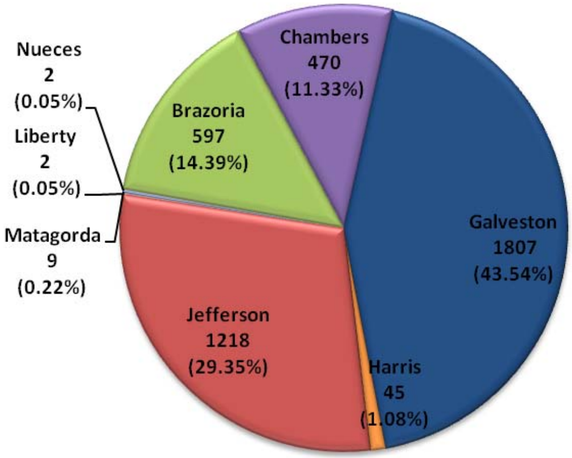


Fig. 5. Distribution of total claim payout amounts per County from Hurricane Ike.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

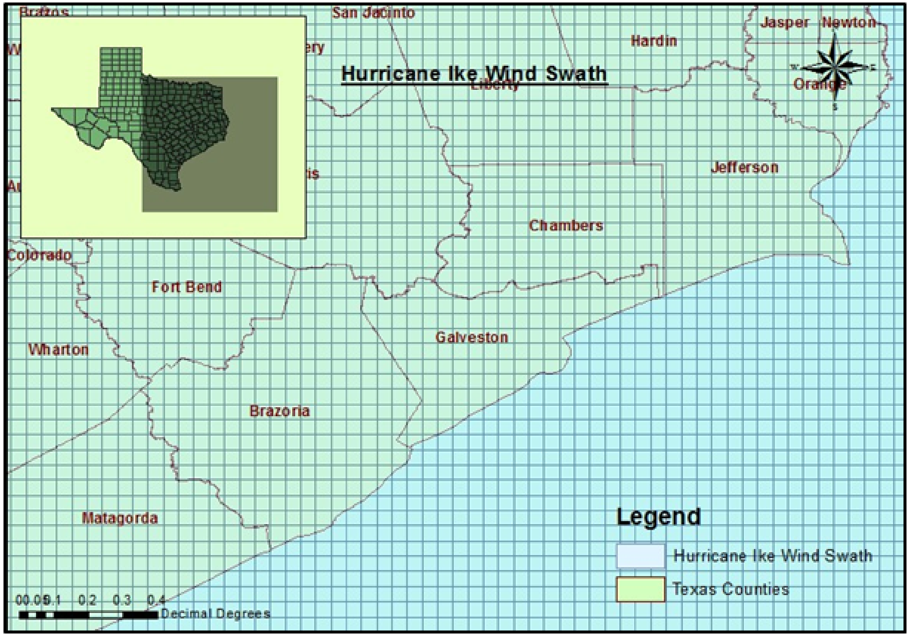


Fig. 6. Hurricane Ike of H*wind swath for Texas coastal counties.

NHESSD
1, 3813–3855, 2013

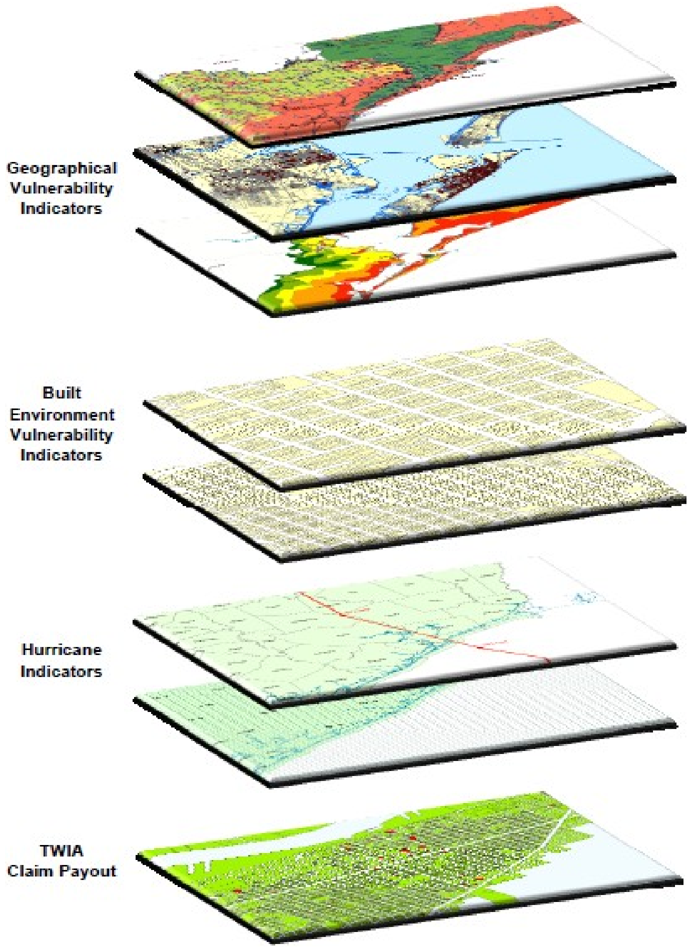


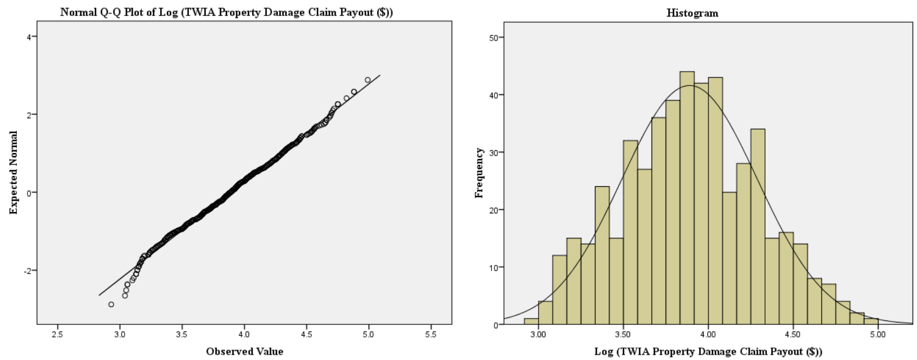
Fig. 7. GIS Process.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion



(a) (b)

Fig. 8. Q–Q plot **(a)** and histogram **(b)** of residuals for the initial regression model.

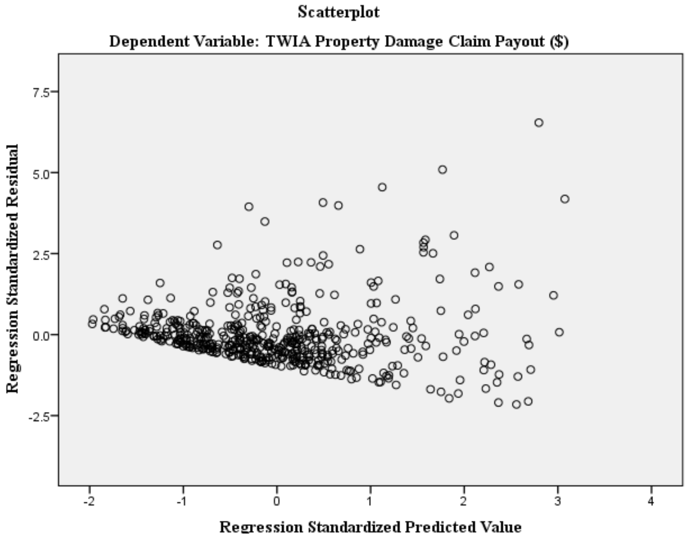


Fig. 9. Residuals plot for TWIA claim payout regression model.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

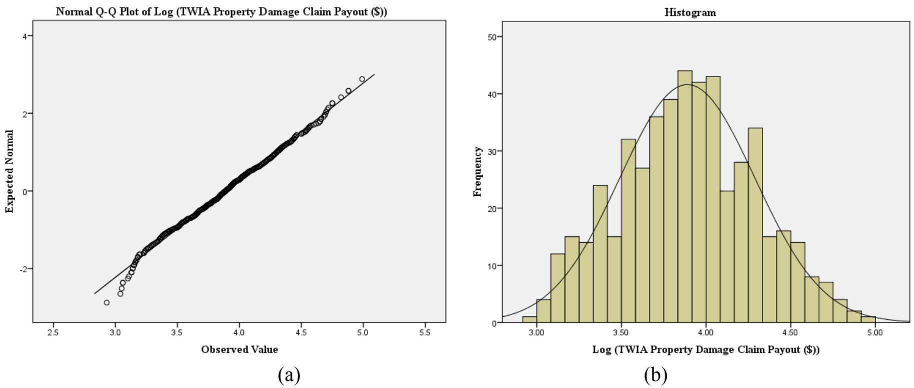


Fig. 10. Q–Q plot **(a)** and histogram **(b)** of residuals for transformed regression model.

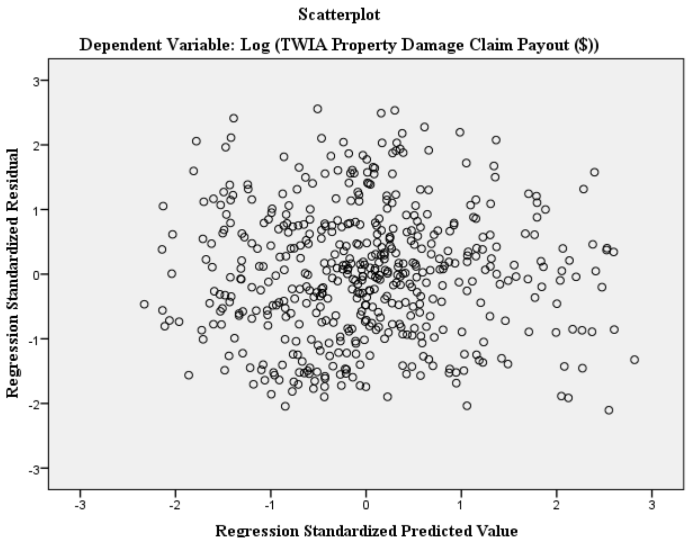


Fig. 11. Residuals plot for transformed TWIA claim payout regression model.

NHESSD
1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.

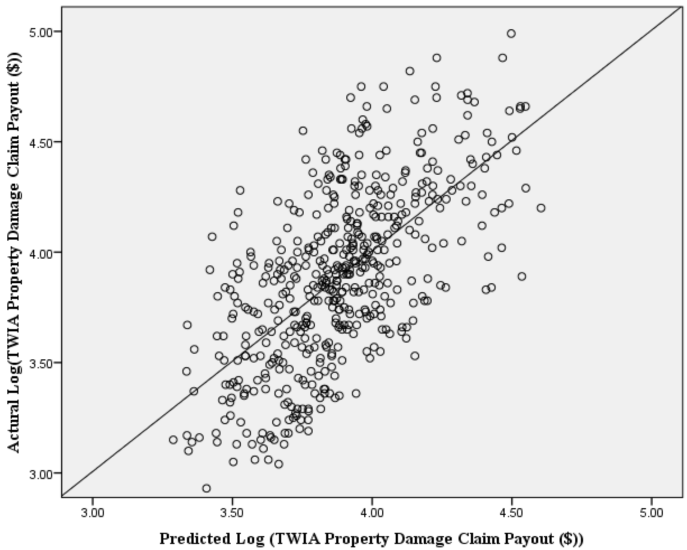


Fig. 12. Actual vs. predicted log TWIA claim payout ($).

NHESSD

1, 3813–3855, 2013

Predicting claim payout of buildings

J. M. Kim et al.