

Interactive comment on “Predicting the hurricane damage ratio of commercial buildings by claim payout from Hurricane Ike” by J. M. Kim et al.

Anonymous Referee #1

Received and published: 3 September 2013

The manuscript by Kim et al. describes a statistical analysis to uncover which indicators can be used to predict hurricane damage. The manuscript adds to the current stock of literature on the subject in that it combines various indicators found by other studies to be significantly related to hurricane damage into a single regression analysis (though this could be stressed better in the manuscript). The methodology used in the paper is quite straightforward. Overall, the manuscript has some potential, but definitely needs a bit more body before it can be considered publishing in NHESS.

Generally, my main points are: 1) The manuscript is very descriptive and at times actually reads like a (statistics) textbook (“the models is statistically significant because the calculated p-value is less than 0.05 . . . this means that there is a significant relationship between the dependent and independent variable”). Instead of describing which num-

C1062

bers come out of the analysis, the authors should address what this significant finding actually means (the statistical significance of parameter C implies that . . .). Without making this step the manuscript has little relevance (nor academic as the methodology is straightforward, nor social as the implications are not discussed). 2) The paper is low on references, and many references seem quite old (from the nineties) whilst this is a subject that has been researched quite a bit in the last decade. For instance, a reference in the manuscript from 1995 states that predicting hurricane damage is complicated, but a lot has happened in 18 years! 3) The data and methodology should be elaborated upon to make clear what the study actually looks at. The authors should elaborate on the ratio of insured vs. uninsured losses in their case study (as they use insurance data) and the relation between flood and wind damage in case of hurricane Ike and what is included in the data they use (is that only wind damage, only flood damage, or both?). Moreover, some choices are not motivated well (why the 500 random sample instead of taking all 4000 points? Linear regressions aren’t that computer intensive to run). 4) The figures and tables could be improved. I’ll give some suggestions later on.

Specific comments: p. 3450, l. 22: statement about rising of occurrence of natural disasters due to climate change and abnormal weather needs a reference (especially as it’s quite a strong statement, there is a lot of discussion on this subject)

p. 3450, l. 24: there are for sure more recent refs than the 1998 ref stated here. Take a look at the work of L.M. Bouwer.

p. 3450, l. 25: tsunamis are not meteorological disasters as stated here.

p. 3451, l. 1-3: again, this statement needs a reference

p. 3451, l. 7-8: establishing a model doesn’t reduce financial loss at all, it only reveals the potential loss. Actual measures are necessary to reduce the real loss

p. 3451, l. 12: more recent reference necessary as a lot has happened in 18 years

C1063

(there are various recent studies to hurricanes and damage in NYC for instance)

p. 3452, l. 2: change 'Initially' to 'First'

p. 3452, l. 8-11: basically says the same as the last sentence before.

p. 3452, l. 17: add 'At the time ...' before 'Hurricane Ike' as with Sandy in 2012 I think Ike moved down to fourth place now.

p. 3452, general: some words are necessary on how Ike caused damage. This can differ substantially between hurricane. Take for instance Andrew, which caused mainly storm damage, and Katrina, which mainly caused damage by flooding. Also the type of damage covered by TWA and how this relates to the damaging effect of Ike is crucial. Does the dataset used in this study only relate to storm losses? (as I would infer from line 25)

p. 3453, l. 5: add 'spatial' before 'distribution'

p. 3453, l. 7-8: why do you use claims from a duration of 3.5 years? And more strikingly: you use claims dating from a month before the actual storm hit?

p. 3453, l. 9-18: this whole paragraph just sums up the numbers found in the table, so adds nothing. Highlighting the important message from the table is sufficient. Also: too many significant numbers are used, i.e. 'overall claim payout of 450 M\$'

p. 3453, l.19-22: why go through the trouble of getting a random sample of 500 instead of using all your data? 4150 data points should be no problem for a linear regression. Also an example is given for a sample population of 5000, whilst the real population is 4150.

p. 3454, l. 14: what do you mean with 'wind analysis'?

p. 3454, l. 15-16: you're describing data, not an analysis.

p. 3454, l. 17: I fail to see how the swath map gives any indication of damage, extra

C1064

data (on assets) and a model is necessary for that. Same goes for line 23

p. 3453, l. 26-27: on the left and right side of the hurricane, doesn't this also relate on the direction the system moves? I can understand this distinction in a system that moves from east to west, but what if a systems moves from south to north?

p. 3455, l.4: robust and variable winds? Those characteristics sound quite conflicting to me.

p. 3455, l. 4-7: it would be nicer to let the analysis show if this is indeed the case! That's exactly what you can use correlation and regression analysis for.

p. 3455, l. 11: to assess the probability of loss the insurer would also have to know about the frequency of events, when evaluating the insured built environment one can only say something about the a 'probable' or 'possible' loss (not the probability)

p. 3455, l. 12-13: the vulnerability is not influenced by the magnitude of the loss, but the magnitude of the loss is determined by the vulnerability.

p. 3455, end: what data is used for the area and age of buildings? Is this the year of building, or the year of last substantial renovation? Also: why not try more indicators and see if they maybe affect hurricane damage as well? (like amount of floors, foundation type, building material, etc.)

p. 3456, l. 27: I'd consider rephrasing those final sentences saying '... should be considered when prediction hurricane damage' into 'In this study we thus consider ... for predicting hurricane damage'. (also goes for earlier instances)

p. 3457, l. 6: unconditional financial damage?

p. 3457, 6-9: explain why this ratio is chosen as indicator, and what it means (the percentage of property damaged?)

p. 3457, l.10: I don't see components 5, 6 and 7 in the equation, though they are

C1065

explained.

p. 3457, l. 19: what is the difference between X-zone and unregistered?

p. 3457, l. 22: I would call this section differently: Results.

p. 3458, l. 19-20: this sentence adds nothing. Discuss your results here: 'all coefficients have positive values, indicating ...' (same goes for final sentence of 4.2)

p. 3459, l. 23-26: this part looks straight from a textbook and says nothing about your results. Instead of this part discuss what your significant relationship actually means.

p. 3460, l. 15-16: why 'scaled from 0 to 1', standardization already allows to compare variables with different units.

p. 3460, l. 24: $\$=\$$ should be $\$/\$$? (or just remove it)

p. 3460, l. 2 vs. l. 22: there seems to be an inconsistency in the text. First 33.7% of the variability is explained, and later 34.3%. In any case, I would not use decimals, two significant numbers is enough.

p. 3460, l. 6-9: other factors were found significant as well (see Table 4 and 5), like wind speed. But because of strong autocorrelation between variables (wind speed and distance to shoreline in this case), one of the two is enough for the model. Please explain such details in your results.

p. 3461, section 4.4: can be removed (is same as 4.3)

p. 3461, l. 17: I see no figure of a map with the spatial distribution of losses (which could easily be done with the formula that is calculated), only of the payouts (which is input for the study, not so much a result).

p. 3461, l. 20-21: the ratio ($\$/\$$) should be unit-less I'd say, so the regression does not say anything about real pecuniary losses, but about the share of total possible loss.

p. 3462, l. 10-14: remove this part, focus on the findings of your analysis. How could

C1066

the TWA use your results? I.e. what use is (spatial explicit) hurricane prediction for the TWA? p. 3462, l. 15-16: avoid this kind of textbook sentences

Tables and Figures:

A table with the independent variables used in the study, their sources (of data), unit, and studies which have studied them before would give a good overview for the reader.

Table 3: the median and 50 percentile are given, which are exactly the same, so one can be removed from the table.

Table 4, caption: add: '... for continuous variables used in regression'. Add mark for significance (possibly even for different levels).

Table 5, caption: should be Results of Spearman correlation analysis. And I would add '... for ordinal variables used in regression'.

I would merge Table 6 and 7 to save space.

Table 9: show would be nicer to show all coefficients of the regression model, so the reader can see for him/herself which ones are significant.

Figure 1: shows not only data collection, but also the analysis (i.e. regression model is not data collection)

Figure 3: the legend seems off as it shows standard deviation. Should be something related to wind speed I presume

Figure 4: can be removed, also doesn't match the caption at all

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 1, 3449, 2013.

C1067