Jennifer Clark
28 April 2009
Final Project

Environmental Statistics

Disaster Risk Models

Natural disasters are an unavoidable part of the environment from which no location can escape as was seen from Hurricane Katrina in Louisiana to the vast tsunamis that racked Southeast Asia and the massive floods in Indonesia. In each of these natural disasters there was substantial loss in human life as well as social, economic, and environmental assets. Had there been a better infrastructure for dealing with such disasters then perhaps much of the losses associated with these disasters could have been reduced. Many international agencies such as the United Nations have begun looking into building such an effective framework using disaster risk management. However, such an infrastructure is costly for a nation, although implementation is less costly than paying for cleanup and rebuilding after a disaster has struck, not to mention the many lives that could be saved if effective measures were in place. The *Hyogo Framework for Action 2005-2015* was published in January 2005 as a collaborative effort by many governments and international agencies to define a set of risk management priorities. However, in order to reduce exposure and vulnerability to natural hazards, high risk countries and locations for natural disasters must be identified. Building such disaster risk models is still in its early stages with two of the first such global models being the Disaster Risk Index (DRI) which was developed by the United Nations Environment Program (UNEP) Division of Early Warning and Assessment Global Resource Information Database project under contract with the United Nations Development Program, and the Hoptspots project which was a collaboration between Columbia University and the World Bank. Many major organizations involved in disaster risk management regularly use the DRI and Hotspots results for country prioritization (Mosquera-Machado, 2009). Although some of the data used to build these models is the same, the methodologies behind them are quite diverse. In a paper by Mosquera-Machado and Dilley these two models are standardized to be on the same scale and compared to assess how well they identify high risk areas and to determine how future disaster risk models can be improved.

The DRI was created with the overarching objective of demonstrating the ways in which development contributes to human vulnerability and risk. Due to the limited availability of usable geophysical and hydrometeorological data to model each hazard's comparative extent and potential severity, the natural hazards that are identified by the DRI are earthquakes, cyclones, floods, and droughts. These hazards tend to be the most dominant in association with records of loss of life accounting 94.43% of those killed by natural disasters. Separate models for each of these hazards were fit using similar methods. The data for this model comes mostly from the International Emergency Disasters Database (EM-DAT) and various other international databases with data collected from 1980 to 2000. A simple measure of relative vulnerability (the ratio of mortality versus the exposed population) was used with disaster mortality being calculated as a product of hazard, population exposed, and vulnerability variables (often depending on the socio-political-economical context of the population) listed in Table T.2 in the Appendix. There were two underlying hypotheses on which the statistical analysis for the DRI was based, risk can be understood in terms of the number of victims of past hazardous events, and the equation of risk follows a multiplicative model where $K = C \times (PhExp)^{\alpha} \times V_1^{\alpha_1} \times V_2^{\alpha_2} \times \ldots \times V_p^{\alpha_p}$ where k is the number of persons killed by a certain type of hazard, C is the multiplicative constant, $PhExp$ is the physical exposure, V_i are the socio-economic parameters identified in Table T.2, α_i is the exponent of V_i, which can be negative. A log transformation was used on the multiplicative model in order to generate a linear relationship between logarithmic sets of values so that a linear regression could be used, $ln(K) = ln(C) + \alpha(PhExp) + \alpha_1 ln(V_1) + \alpha_2 ln(V_2) + \ldots + \alpha_p ln(V_p)$. Physical exposure for the DRI model was calculated two different ways, one way was by multiplying the hazard frequency, which was calculated for different strengths of event, by the population living in each exposed area $PhExp = \sum F_i \times Pop_i$, where $PhExp$ is the total physical exposure of a country (the sum of all physical exposure in the country), F_i is the annual frequency of a specific magnitude event in one spatial unit, and Pop_i is the total population living in the spatial unit. When data on the annual frequency of a specific magnitude event (such as earthquakes) was not available then physical exposure was found by taking the exposed population and dividing by the number of years when an even had taken place, $PhExp = \sum \frac{Pop_i}{Y_n}$ where Y_n is the length of time in years. The socio-economic covariates were averaged over the 21 year period before doing a log transformation. Covariates that were expressed as a percentage were converted into odds for the model, $V_i' = \frac{V_i}{1-V_i}$. In addition to these models an overall model summing all calculated deaths was created for a multi-hazard vulnerability index. Results from these regressions are given in the Appendix.

The Hotspots project was created with a similar goal as the DRI to provide a relative representation of disaster risk, but there are many notable differences. It was created after the DRI and calculates the relative risks for six natural hazards, as opposed to the DRI's four, earthquakes, volcanoes, landslides, floods, drought, and cyclones. Also, rather than calculating a national risk for each country, Hotspots uses a global gridding system of 2.5' x 2.5' (25 sq kilometers) latitude-longitude grid cells known as the Gridded Population of the World (GPW). Most countries had population data which the GPW transformed into a grid of spherical quadrilaterals where each cell contains an estimate of the population and population density on land. Cells with a population density less than 5 persons per square kilometer (ie, cells < 105 people) and without significant agriculture were not included as the total casualties and agricultural losses would be quite small and these low density cells could potentially dominate the results if included. Population estimates for 1990, 1995, and 2000 for 375,000 sub-national administrative units were used (Dilley et al). Much of the global hazard data for this was collected from the same sources as the DRI over a similar period of time from 1981 to 2000; sources can be seen in Table 2 in the Appendix. The actual risk assessment was calculated first by finding the total global losses from 1981-2000 (ie, number of fatalities, economic losses, etc.) by hazard h, M_h, and the total population estimated to live in the area affected by the hazard in 2000, P_h. From this a simple mortality rate for this hazard can be computed as $r_h = \frac{M_h}{P_h}$ which, assuming that 1981-2000 is representative of a typical 20-year period for that hazard in that area, is the estimate of the proportion of persons killed in the exposed area during a 20-year period. A location-specific expected mortality was computed for each grid cell i that falls into the hazard zone for h, $M_{hi} = r_h \times P_i$. A kind of stratification by the various combinations, denoted by j, of region and country-wealth status (see Tables 3 and 4 in the Appendix for these combinations) was applied to get estimated mortality as $M_{hij} = r_{hj} \times P_i$. A weight, $W_h i$, was applied to each of the six hazards to reflect the number of times the hazard has hit the region and the degree of severity with which it was hit, $M'_{hij} = r_{hj} \times W_{hi} \times P_i$. However, the results from these accumulated mortalities for each hazard cannot be summed to get a multi-hazard disaster risk hotspot index as measurements for different hazards are on different scales and summing could inadvertently cause one hazard to dominate if it happens to be on a bigger scale. The estimated mortalities for each hazard were therefore standardized so that the total mortality for each region summed to the total that was found in the EM-DAT dataset used for this analysis which results in a mortality for hazard h in cell i with region/wealth-class combinations j as $M^*_{hij} = M'_{hij} \times \frac{M_{hj}}{\sum_{i=1}^{n} M'_{hij}}$ with n being the number of grid cells in the hazard zone for h. The mortality-weighted multi-hazard disaster risk hotspot index was then calculated as $Y^*_i = \sum_{h=1}^{6} M^*_{hij}$. These measures were then converted into an index ranging from 1 to 10 to classify those grid cells that were used to make the index into deciles in order to provide a relative representation of disaster risk.

Although both the DRI and the Hotspots project provide estimates of disaster risk, they do so in very divergent ways. In order to do a comparison of the two indexes Mosquera-Machado et al had to first standardize them so they were measuring the same hazards across the same scale. Since the Hotspots index included volcanoes and landslides which were not included in the DRI, these hazards had to be removed and the unit of measure converted from grid cells to countries across the globe. A table containing the top 25 countries that were identified as being at risk by each disaster risk model was made for each of the hazards as well as a table with those countries identified as having the highest overall disaster risk assessment by each model. The range of overlap between these tables within hazard went from 7 out of 25 with earthquakes, to 15 out of 25 for cyclones and drought (floods had 8). The multi-hazard disaster-related mortality risk tables only overlapped by 6 out of the 25 countries that were identified by the two models. Spearman's rank correlation was also used to evaluate the degree of correspondence among the country rank orders. These showed varying degrees of correlation with the highest being 0.56 for cyclones and 0.41 for earthquakes among the individual hazards, and only 0.31 for the multi-hazard models. However, the low correspondence between these disaster risk models should not be seen as a failure in part of one or both of the models, but rather how the quality and availability of data along with methodological choices can substantially affect the results. While the results from these two models are not highly correlated, they are consistent with the input datasets and methods used to generate these models. The quality and availability of global disaster data has improved considerably over time and will probably continue to do so which will improve future disaster risk models. Since the results for disaster risk seem quite sensitive to the methodological choices for the model building, multiple methods for assessing particular types of risk could prove useful here. Future disaster risk models may want to consider using a Bayesian decision theory framework with a standard loss function to account for disaster risk in a given area.

References

Dilley M, Chen RS, Deichmann U, Lerner-Lam AL, Arnold M, Agwe J, Buys P, Kyekstad O, Lyon B, Yetman G (2005) Natural disaster hotspots: a global risk analysis. International Bank for Reconstruction and Development/The World Bank and Columbia University, Washington, CD, 132 pp

Mosquera-Machado S, Dilley M (2009) A comparison of selected global disaster risk assessment results. Natural Hazards, Springer Netherlands Vol 48 Number 3, 439 pp.

United Nations Development Programme-UNDP (2004) Reducing disaster risk: a challenge for development. United Nations Development Programme, Bureau for Crisis Prevention and Recovery, New York, 137 pp. www.undp.org/cpr/disred/documents/publications/rdr/english/rdr_english.pdf