Stevens Institute of Technology: Summer Research Institute Program

"Social Media and Climatic Based Proxies for Disaster Resource Allocation: Mexican Case"

Andrea Garcia Tapia

Advisors:

Dante Gama Dessavre
Jose Ramirez-Marquez

August 5, 2015

STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

Castle Point on Hudson, Hoboken NJ 07030-5991
USA +1.201.216.5000

Abstract

Over the last decade, Mexico has experienced a sharp increase in the economic costs associated with climatic events, such as floods, tropical cyclones and heavy rain. Different risk transfer instruments for disaster mitigation have been used in Mexico. At Federal level the main instruments are the National Disaster Fund (FONDEN) and the Catastrophic Bond (CATBOND). While CATBOND has clear operational rules and triggers, the resource allocation of the National Disaster Fund (FONDEN) is not clear. This paper address the main determinants for FONDEN using a set of political, socioeconomic and meteorological data ……(SVM & PCA) …. for all 32 Mexican federal entities for the period 2002-2007.

Our results show that the main determinants for FONDEN resource allocation are : ……, these findings lend us an important insight into the mechanisms ……

Based on our results, the recommendation for local governments is to develop an institutional structure that focuses on enhancing …….

Keywords: Natural Disaster, Disaster National Funds

1 Introduction

Over the last decade, Mexico has experienced an increase in the economic costs associated with hydro meteorological disasters such as floods, hurricanes and droughts. This is attributed to the combination of an increasing population and expanding economic activities along Mexico's coastal areas and arid zones with the mismanagement of its urban growth.[1] The most expensive year regarding disaster relief in Mexican history occurred in 2010 driven by geological and hydro meteorological phenomenon's, with an economic loss equivalent to 0.8% of the country's GDP.[2] Due to the effects of climate change and the unplanned expansion of urban areas, damages and losses from hydro meteorological events will likely continue to grow. In preparation, Mexico must strengthen its existing mechanisms for generating economic development, including Disaster Risk Management and resilience to climate change as key components of its development process[3].

Disaster resilience refers to the *ability of a system and its component parts to anticipate, absorb, accommodate, or recover from the effects of a hazardous event in a timely and efficient manner, including through ensuring the preservation, restoration, or improvement of its essential basic structures and functions.* [4] Three of the main factors for achieving resilience are: preparedness, reaction time and adaptability. For each of these factors, tools that enable monitoring and event identification are extremely important[5]specially for public policy management.

++Include disaster risk transfer instruments++

2 Mexican Disaster Relief Scheme

++Federal & local existing risk transfer instruments++

Regarding Mexican Disaster Risk Management the Ministry of Interior (SEGOB) is responsible for manage the National Disaster Fund (FONDEN), the National Center for Preventing Disasters (CENAPRED) and the National System for Civil Protection (SINAPROC).[6] SINAPROC is in charge of the first response during a disaster, mainly attending affected population. FONDEN is a national fund for attending public spending regarding disaster recovery. Meanwhile CENAPRED does the post disaster damage evaluation and the research for preparedness and early warning alarm system. The figure 1 explains the organization chart and the hierarchy of Mexican Disaster Risk Management.

The Mexican Civil Protection Law defines a disaster as *"... a situation in which the population of one or more State entities suffers several damages from the impact of a natural or man–made disaster calamity, resulting in loss of life, infrastructure or environment, in a way that disrupts the social structure and disturbs the essential activities of society, affecting livelihoods."* When a disaster occurs, the local government (municipal and state level) attends the disaster through the local civil protection system and a state fund for disasters.

[1] Garcia A(2014) Desastres Naturales: ¿Destrucción Creativa?, ITAM

[2] CENAPRED(2010)

[3] Garcia A, Muñoz C (2014) The Effect of Natural Disasters on Mexico's Regional Economic Growth: Growing Disparity or Creative Destruction? LACEEP working paper

[4] IPCC, 2012, Glossary of terms. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change

[5] Ramirez-Marquez JE, Rocco CM. (2009) Stochastic network interdiction optimization via capacitated network reliability modeling and probabilistic solution discovery. Reliability Engineering and System Safety; 94(5):913–921

[6] This section heavily relies in the 2nd chapter "National context" Garcia, A (2014) Desastres Naturales: ¿Destrucción Creativa?, ITAM Batchelor thesis

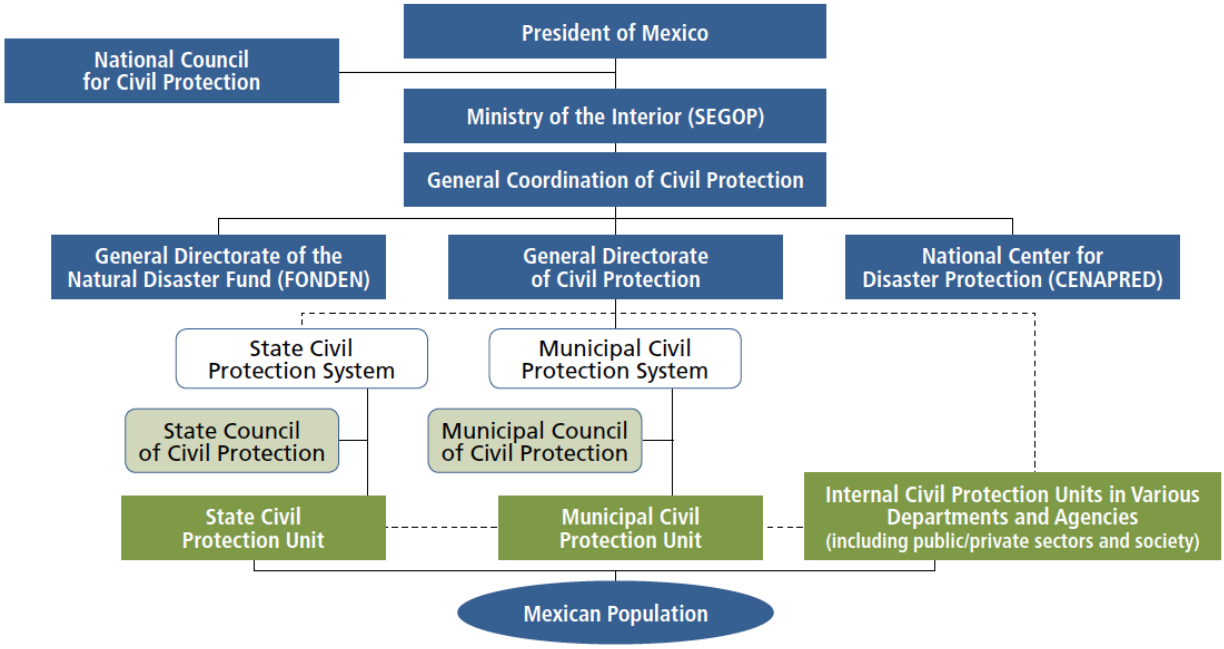

Figure 1: Flow Chart of the Civil Protection System in Mexico, Source: "FONDEN: Mexico´s Natural Disaster Fund, A Review", World Bank (2012)

If the disaster overwhelms local capacity, the state government request help from SEGOB. One of the main problems is there are no indicators nor measures for "local capacity".

When a disaster occurs the process of declaration is the following: First, SEGOB must issue a declaration of a natural disaster in order for FONDEN resources to be accessible by affected federal agencies or state governments. Once this declaration has been made, the federal agencies and/or state government(s) can apply for funding and the damage assessment process. In order to ensure efficiency and accuracy of the damage assessment process, a committee composed by CENAPRED, the National Water Commission (CONAGUA) and the National Forestry Commission (CONAFOR), make a report of the disaster. Based on the findings of the damage assessment, SEGOB reviews the related funding applications, determines the appropriate allocations, and requests the Ministry of Finance and Public Credit (SHCP) to convene the FONDEN Technical Committee to authorize the transfer of funds to a subaccount for the reconstruction program in the FONDEN Trust. From this subaccount, resources are transferred to the service providers implementing reconstruction works. FONDEN resources finance 100 percent of the reconstruction costs for federal assets and 50 percent of those for local assets the first time that the assets are affected by a disaster, afterwards the percentage of recovery declines and local government should buy insurance for the reconstructed assets[7]. Figure 2 explains the process to access and execute FONDEN resources for post-disaster reconstruction.

[7]GFDRR (2012)

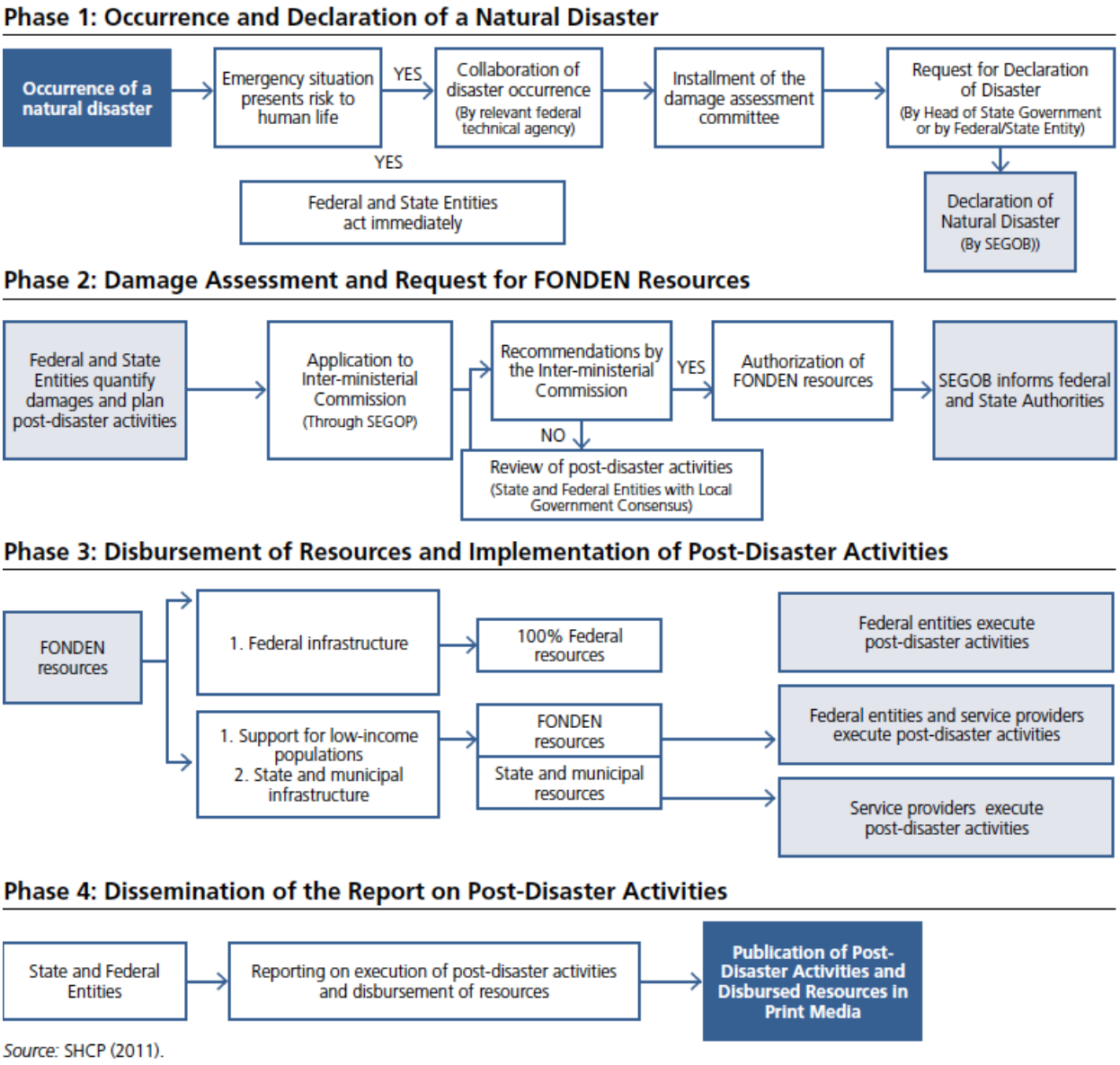

Figure 2: Process to access and execute FONDEN resources for post-disaster reconstruction, Source: "FONDEN: Mexico´s Natural Disaster Fund, A Review", World Bank (2012)

It is important to mention that federal disaster relief assistance is a grant and all the system depends in the declaration made by SEGOB. If the disaster is not declared then the state has no help from the federal level, meaning, it only has local assistance. If the disaster is declared, SEGOB has four types of declaration[8]

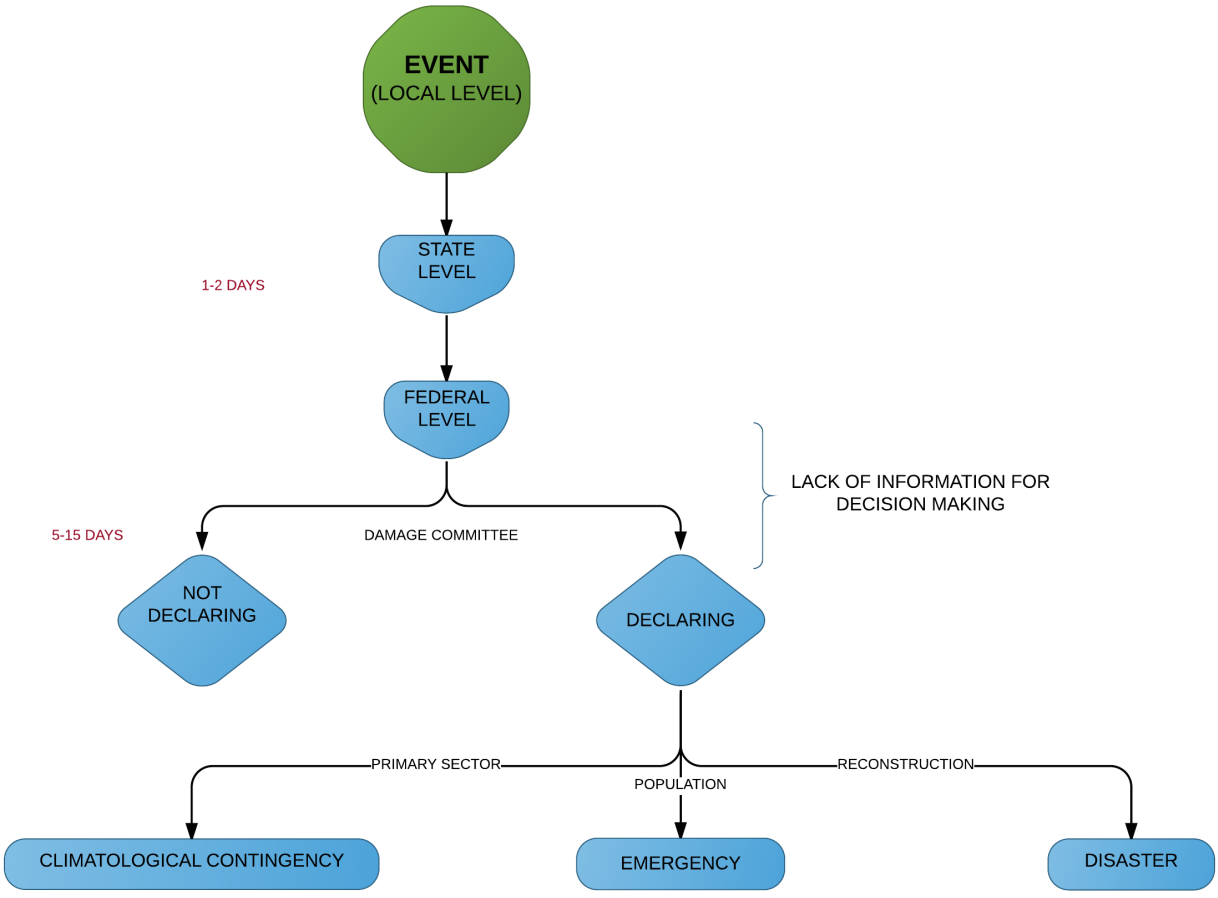

Figure 3: Disaster Declaration Process

[8]The type of aid depends on the type of declaration

3 Research Problem

The lack of information during and after a disaster is one of the main problems for public policy makers for disaster mitigation and even conflict prevention.[9][10] As explained above when Mexican government decide between declaring and dot declaring a disaster they only rely on the information given by the local governments and the meteorological data (precipitation, wind and temperature)

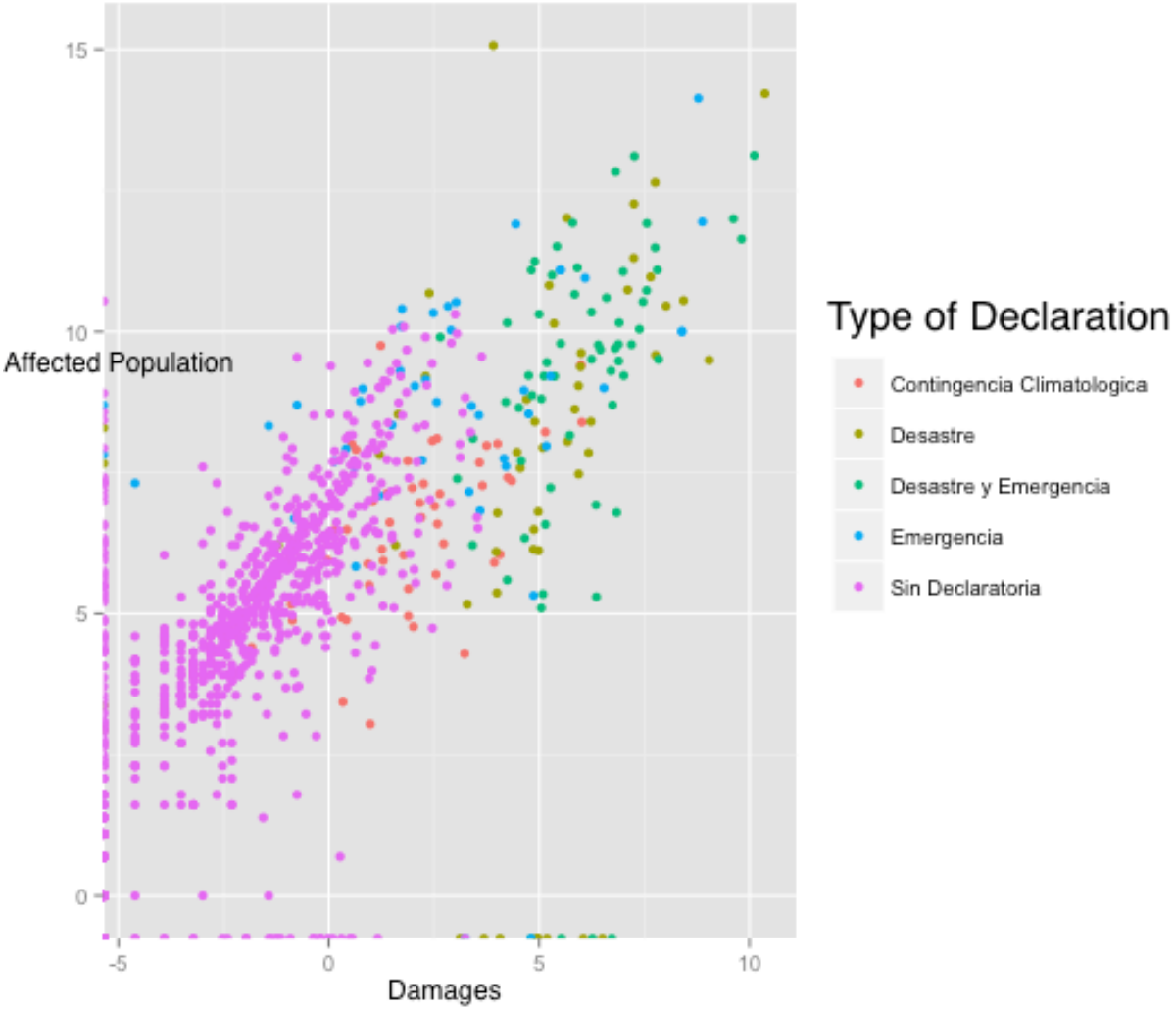

Figure 4: Damages vs Affected Population

[9]Meier, Patrick. (2014). Crisis Mapping in Areas of Limited Statehood. In Information and Communication Technologies in Areas of Limited Statehood, ed. Steven Livingston and Gregor Walter-Drop. Oxford University Press

[10]Meier, Patrick. (2014). Human Computation for Disaster Response. In Handbook of Human Computation, ed. Pietro Michelucci et. al, Springer

4 Literature Review

5 Knowledge Gaps

6 Research Objectives

This project analyzes the main variables determining the Mexican process of disaster declaration looking forward a predictor for helping Mexican government to make a transparent process regarding disasters driven by hydro meteorological phenomena's. The objective of this project is to develop and apply methods to assess the suitability of using news flows and precipitation data to characterize disaster damages in Mexico looking forward to resource allocation improvement.

The extension of the project depends of the data needed and gathered. As mentioned above we will work towards an open, real time visualization platform for coordinated disaster mitigation for decision-making.

7 Data

The period of study is 2002-2013. The following section describes the data groups needed for this project:

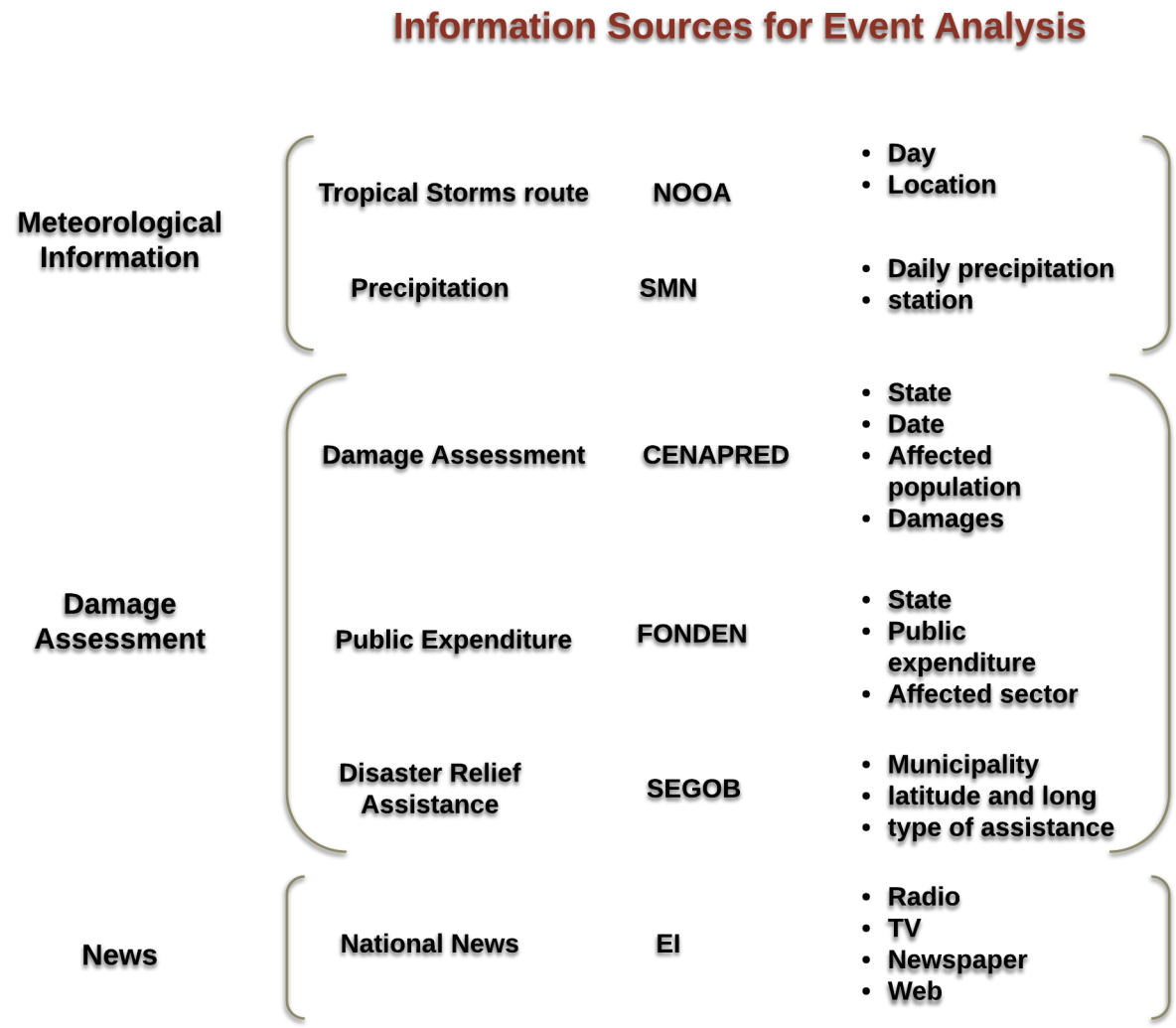

Figure 5: Sources

- **News data**: The Company "Eficiencia Informativa" EI gathers information from electronic, newspaper, radio and TV news. The idea is to scrap all the news related to climatic events during the period of study. The company gave an acces for using their information till the endo of the year.

Figure 6: Eficiencia Informativa System

They have different formats for downloading the information. The best is CSV

ID	Fecha	Hora	Medio	Tipo de Medio	Autor	Genero	Título	Resumen	Transcripción	Tema	Programa	Duración	Sección	Página	Precio	URL del Testigo
4100 9670	2013-09-16	18:58:04	Sin Embargo.com	Periodismo Digital	Redacción	Nota Informativa	Aerolíneas buscan "puente aéreo" para sacar a turistas de Acapulco; autoridades instalan 10 refugios	http://www.sinembargo.mx/16-09-2013/756281 Las aerolíneas Interjet y Aeroméxico informaron que apoyarán a los turistas que quedaron varados en el Puerto de Acapulco, en el estado de Guerrero, a causa de los daños en la Autopista del Sol que ocasionó la tormenta tropical ?Manuel?. Por medio de su cuenta de Twitter Interjet indicó a sus usuarios que sino no pudieron viajar este lunes podrán hacerlo el día de mañana sin cargos adicionales. Indicó que también se podrán hacer cambios de ruta y fecha sin que se cobren cargos por las operaciones. Sin embargo, precisó que si existe una diferencia entre las tarifas deberá ser cubierta la diferencia. Aeroméxico comunicó que realizaráa vuelos extraordinarios de Acapulco a México en cuanto sea reabierto el Aeropuerto Internacional de Acapulco. ?Aeroméxico también ha dispuesto una tarifa especial en estos vuelos, de mil 599 pesos viaje sencillo, para apoyar a todos los pasajeros que por algún motivo busquen viajar durante esta contingencia?, informó la aerolínea en un boletín de prensa. Indicó que ?en las próximas horas se darán a conocer los horarios de los vuelos a través de nuestra página web (www.Aeroméxico.com) y recordamos que, la página oficial de facebook, cuentas de twitter @Aeroméxico_com y @am_escucha, así como el call center 51-33-40-00 desde la Ciudad de México, se encuentran a disponibles para atender a los pasajeros?, agregó. Las fuertes lluvias que han originado la tormenta y depresión tropical ?Manuel? en el Pacífico e ?Ingrid? en el Golfo, respectivamente, han golpeado principalmente a los estados de Guerrero en una costa y Veracruz en la otra. Desde el fin de semana ambos fenómenos meteorológicos han generado lluvias y afectaciones en las dos costas del país, con un saldo de al menos 34 muertos y 1 millón 200 mil afectados. Tan sólo en Guerrero, Luis	http://www.sinembargo.mx/16-09-2013/756281 Las aerolíneas Interjet y Aeroméxico informaron que apoyarán a los turistas que quedaron varados en el Puerto de Acapulco, en el estado de Guerrero, a causa de los daños en la Autopista del Sol que ocasionó la tormenta tropical ?Manuel?. Por medio de su cuenta de Twitter Interjet indicó a sus usuarios que sino no pudieron viajar este lunes podrán hacerlo el día de mañana sin cargos adicionales. Indicó que también se podrán hacer cambios de ruta y fecha sin que se cobren cargos por las operaciones. Sin embargo, precisó que si existe una diferencia entre las tarifas deberá ser cubierta la diferencia. Aeroméxico comunicó que realizaráa vuelos extraordinarios de Acapulco a México en cuanto sea reabierto el Aeropuerto Internacional de Acapulco. ?Aeroméxico también ha dispuesto una tarifa especial en estos vuelos, de mil 599 pesos viaje sencillo, para apoyar a todos los pasajeros que por algún motivo busquen viajar durante esta contingencia?, informó la aerolínea en un boletín de prensa. Indicó que ?en las próximas horas se darán a conocer los horarios de los vuelos a través de nuestra página web (www.Aeroméxico.com) y recordamos que, la página oficial de facebook, cuentas de twitter @Aeroméxico_com y	-	"Default 0:00 - > 0:00	00:00:00			$ 1.00	http://data4.efinf.com/reader/display?id=338a9c19f6f487dbb51b55114235c7a8c<link

Figure 7: Example of the DB

- **Disaster Database**:
  - *Disaster Evaluation*: The National Center for Natural Disaster Prevention (CENAPRED) is in charge of the disaster damage and loss evaluation for the affected states. They gave us an access to the National Risk Atlas ANR and the evaluation dataset. Since Ingrid and Manuel Storms affected 19 of the 31 mexican states the government could not make a comprehensive damage

evaluation. They were only able to do 5 damage evaluations (Durango, Guerrero, Nuevo Leon, Sianloa and Veracruz). The access to the ANR is not working so good since an older versión of explorer is needed. We are trying to access the platform in other system. [11]

– *Public Spending*: The Ministry of Interior(SEGOB) manages the Disaster Relief Fond (FONDEN) through the National Civil Protection Service (SINAPROC)[12]. The public spending records are available online Recursos Autorizados. Ingrid y Manuel disasters have a lag of 3 years of public spending. Each year has diferent format and information. We are working on the data cleaning. For following up the disasters, SEGOB opend a platform to inform the reconstruction actions and the money expeded Presidencia. The data in this platform does not match with the data obtained in FONDEN web site in spite of being the same govenrment entity. We already send a mail asking for this inconsistencies.

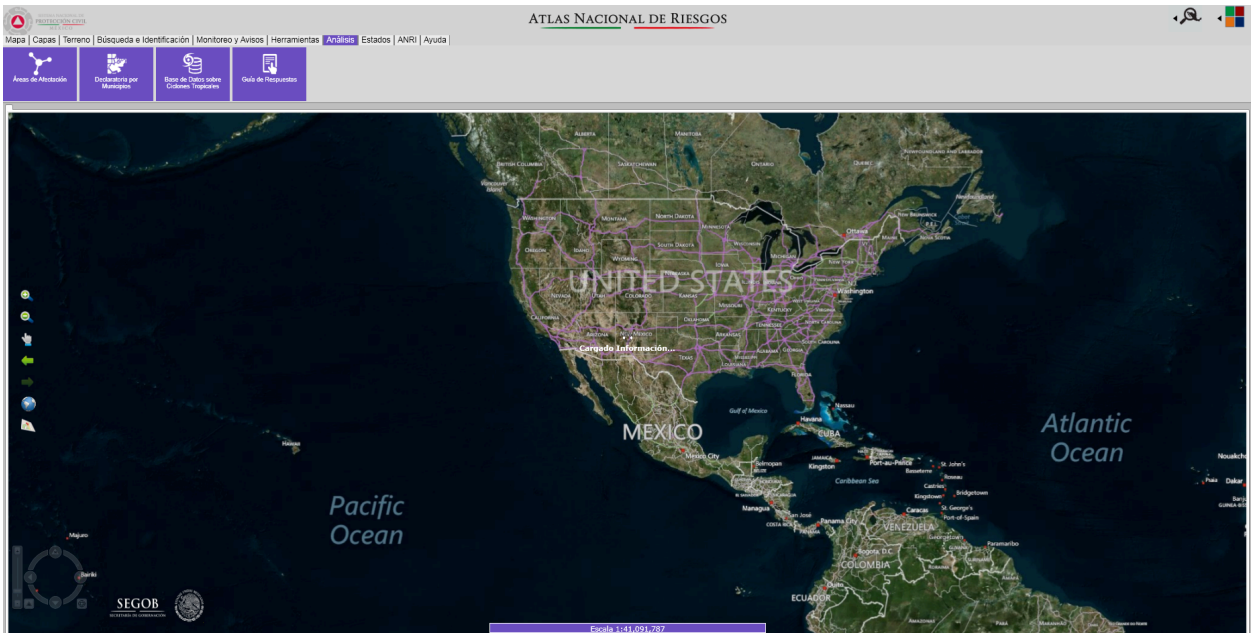

Figure 8: National Risk Atlas

- **Meteorological Information**: The precipitation data could help us as a proxy for damage metrics. The data could be obtained by NASA precipitation grid or by the National Meteorological System.

8 Methodology

++Intro clasificadores++

To clasify the types of delcarations a Support Vector Machine and a Random Forest will be used...

[11] if the acces is not working next week we will procede to ask for each data base

[12] For more information about FONDEN system see Garcia, A. (2014) "Desastres Naturales: Destrucción Creativa?",Chapter2, BA thesis ITAM

8.1 News mining

In the news dataset we have the transcripts of the news on TV, radio and newspapers. Radio transcripts are smaller than the TV or newspaper transcripts so for the text mining processing we need a methodology that allow us to search for the relevant (more frequent words) weighting the extension of the transcript and using stop words.

The idea is to do text mining for processing this database and follow the coverage of the disasters in the news.

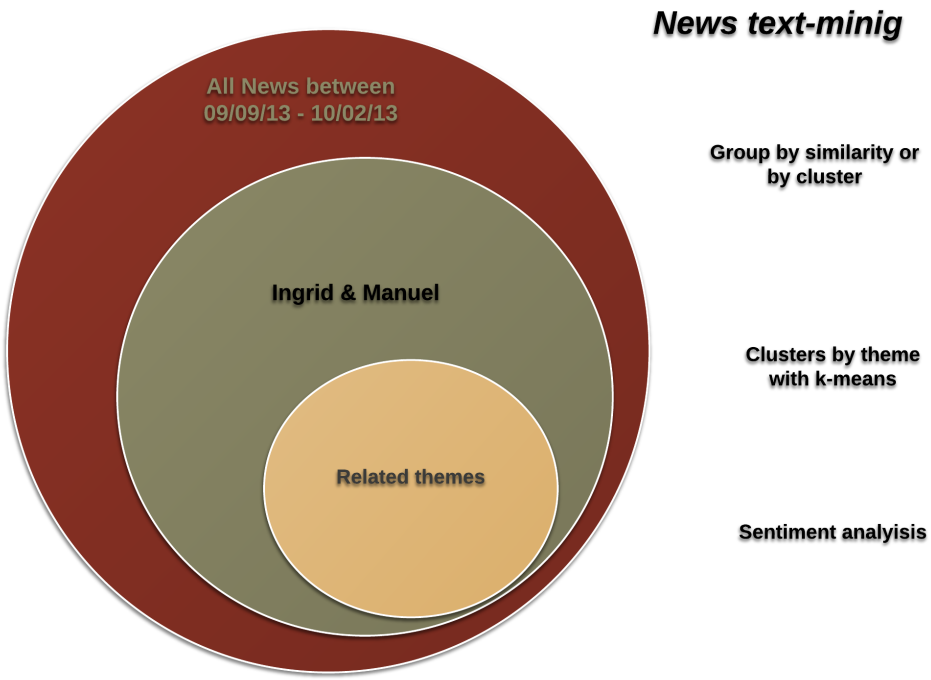

Figure 9: Example news procesing

++ Como se mencionó anteriormente , no basta con filtrar las noticias según la aparición de las palabras clave, ni tampoco un conteo simple. se podría tomar en cuenta los conteos, pero usando técnicas para disminuir la importancia de la longitud de los documentos y de la frecuencia de aparición de las palabras comunes.

Denotamos la frecuencia del término w (puede ser una palabra o la raíz de una palabra obtenida con *stemming*) en el documento d como $tf_{w,d}$. De manera natural, denotamos por tf_d al vector de todas las frecuencias

de los términos conocidos del documento d. Supongamos que queremos comparar los documentos q y d. Un primer enfoque que podríamos utilizar sería usar la distancia coseno, que toma en cuenta únicamente la frecuencia relativa de las palabras dentro de los documentos:

$$d_{coseno}(q,d) = \frac{tf_q \cdot tf_d}{\|tf_q\|\|tf_d\|}$$

El problema con lo anterior es que las palabras comunes podrían tomar un papel protagónico y en nuestro caso compartir por ejemplo artículos o preposiciones no es muy interesante. Para mitigar esto utilizaremos la frecuencia inversa en documentos idf_w, definida para una palabra w. Si N es el número total de documentos en la colección y df_w es el número de documentos de la colección que contienen a w, entonces definimos la frecuencia inversa de documentos como sigue:

$$idf_w = log(\frac{N}{df_w})$$

Y entonces, en lugar de describir a d por medio de tf_d, lo describimos por medio de c_d, donde

$$c_{d,w} = idf_w \times tf_{d,w}$$

Y así, calculamos la distancia entre dos documentos como la distancia coseno entre sus vectores característicos:

$$d(q,d) = \frac{c_q \cdot c_d}{\|c_q\|\|c_d\|}$$

La idf_w es el logaritmo del inverso de la probabilidad de que el término w aparezca en un documento elegido al azar. El efecto que esto tiene es que las palabras comunes casi no tendrán ningún efecto en la distancia, puesto que su probabilidad es cercana a 1 y por lo tanto su idf es cercana a cero. Por el contrario, las palabras raras tendrán una probabilidad cercana a 0, por lo que su idf será grande y contribuirán mucho. La razón detrás de hacer esto es que queremos que discriminen las palabras especiales o específicas a un contexto y no las genéricas. El esquema expuesto está escrito con mayor detalle en el libro Mining of Massive Datasets, con la pequeña diferencia de que ahí además normalizan tf_w por la máxima frecuencia obtenida por el término en la colección de documentos.

Indicador de Noticias

Plan A

Procesar todas las noticias en el periodo de estudio y sacar los temas relevantes en las fechas de desastres y ver si tuvo impacto en las noticias

Plan B Porcesar solo las noticias relacionadas a palabras clave relacionadas a eventos climáticos

++Pedir asesoría al profesor Luis Felipe Gonzalez+++

++

9 Results

10 Conclusion

11 Further Investigation

12 Bibliography

1. Caragea et al, (2011). Classifying text messages for the Haiti earthquake. Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM2011).
2. CENAPRED (2010) Serie de Impacto Socioeconomico www.atlasnacionalderiesgo.gob.mx
3. Chew C & Eysenbach G (2010) Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. PLoS ONE 5(11):e14118.
4. Christakis, Nicholas A. Y James H. Fowler (2010). "Social Network Sensors for Early Detection of Contagious Outbreaks", http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0012948
5. Chunara R, Andrews JR, & Brownstein JS (2012) Social and News Media Enable Estimation of Epidemiological Patterns Early in the 2010 Haitian Cholera Outbreak. The American Journal of Tropical Medicine and Hygiene 86(1):39-45.
6. CONAGUA(2014), "Reporte del Clima en México 2013", Coordinación General del Servicio Meteorológico Nacional. 06/11/15, http://www.conagua.gob.mx
7. Earle PS, Bowden DC, & Guy M (2012) Twitter earthquake detection: earthquake monitoring in a social world. Annals of Geophysics 54(6):708-715.
8. FONDEN, Recursos Autorizados por Declaratoria de Desastre. 06/12/15, http://www.proteccioncivil.gob.mx/es/ProteccionCivil/Recursos_Autorizados_por_Declaratoria_de_Desastre
9. Freeman M (2011) Fire, wind and water: Social networks in natural disasters. Journal of Cases on Information Technology (JCIT) 13(2): 69-79.
10. Gao, H, Barbier, G. and Goolsby R. Harnessing the crowdsourcing power of social media for disaster relief (2011). DTIC Document.
11. Garcia A (2014) ¿Desastres naturales: destrucción creativa? Bachelor Thesis ITAM
12. Garcia A, Muñoz C (2014) The Effect of Natural Disasters on Mexico's Regional Economic Growth: Growing Disparity or Creative Destruction? Working Paper Series, WP68 Latin American and Caribbean Environmental Economics Program (LACEEP), available at : http://www.laceep.org/index.php?option=com_k2&view=item&id=226:the-effect-of-natural-disasters-on-mexico%E2%80%99s-regional-economic-growth-growing-disparity-or-creative-destruction?-andrea-garc&Itemid=87
13. Guan X & Chen C (2014) Using social media data to understand and assess disasters. Natural Hazards:1-14.
14. Guy M, Earle P, Ostrum C, Gruchalla K, & Horvath S (2010) Integration and dissemination of citizen reported and seismically derived earthquake information via social network technologies. Advances in intelligent data analysis IX, (Springer), pp 42-53.

15. Henry D, Ramirez-Marquez JR. (2012) Generic metrics and quantitative approaches for system resilience as a function of time. Reliability Engineering and System Safety; 99(1):114–122.
16. Herfort B, de Albuquerque JP, Schelhorn S-J, & Zipf A (2013) Does the spatiotemporal distribution of tweets match the spatiotemporal distribution of flood phenomena? A study about the River Elbe Flood in June 2013.
17. Hughes AL & Palen L (2009) Twitter adoption and use in mass convergence and emergency events. International Journal of Emergency Management 6(3):248 260.
18. Imran M, Elbassuoni SM, Castillo C, Diaz F, & Meier P (2013) Extracting information nuggets from disaster-related messages in social media. Proc. Of ISCRAM, Baden-Baden, Germany.
19. Kryvasheyeu Y, Chen H, Moro E, Van Hentenryck P, & Cebrian M (2015) Performance of Social Network Sensors during Hurricane Sandy. PLoS ONE 10(2):e0117288.
20. Kumar S, Hu X, & Liu H (2014) A behavior analytics approach to identifying tweets from crisis regions. Proceedings of the 25th ACM conference on Hypertext and social media, (ACM), pp 255-260.
21. Li J & Rao H (2010) Twitter as a rapid response news service: An exploration in the context of the 2008 China Earthquake. The Electronic Journal of Information Systems in Developing Countries 42(4):1-22.
22. Meier, P. (2014) Crisis Mapping in Areas of Limited Statehood. In Information and Communication Technologies in Areas of Limited Statehood, ed. Steven Livingston and Gregor Walter-Drop. Oxford University Press
23. Meier, P. (2014) Human Computation for Disaster Response. In Handbook of Human Computation, ed. Pietro Michelucci et. al, Springer.
24. Meier, P. (2014) Using Advanced Computing to Verify User-Generated Content on Social Media, in The Verification Handbook, eds. Craig Silverman and Rina Tsubaki.
25. Myers SA, Sharma A, Gupta P, & Lin J (2014) Information network or social network?: the structure of the twitter follow graph. Proceedings of the companion publication of the 23rd international conference on World wide web companion, (International World Wide Web Conferences Steering Committee), pp 493-498.
26. National Hurricane Center, NOAA. 06/11/15, http://www.nhc.noaa.gov/aboutsshws.php
27. Power R, Robinson B, Colton J, & Cameron M (2014) Emergency Situation Awareness: Twitter Case Studies. Information Systems for Crisis Response and Management in Mediterranean Countries, Lecture Notes in Business Information Processing, eds Hanachi C, Bénaben F, & Charoy F (Springer International Publishing), Vol 196, pp 218-231.
28. Preis T, Moat HS, Bishop SR, Treleaven P, & Stanley HE (2013) Quantifying the Digital Traces of Hurricane Sandy on Flickr. Sci. Rep. 3:3141.
29. Protección Civil. 06/15/15, http://www.proteccioncivil.gob.mx/es/ProteccionCivil/Apoyos_otorgados_2013
30. Ramirez-Marquez JE, Rocco CM. (2009) Stochastic network interdiction optimization via capacitated network reliability modeling and probabilistic solution discovery. Reliability Engineering and System Safety; 94(5):913–921.
31. Sakaki T, Okazaki M, & Matsuo Y (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. Proceedings of the 19th international conference on World wide web: 851-860.
32. Servicio Meteorológico Nacional (SMN). 06/10/15, http://smn.cna.gob.mx/index.php?option=com_content&view=article&layout=edit&id=268

33. Skinner J (2013) Natural disasters and Twitter: Thinking from both sides of the tweet. First Monday; Volume 18, Number 9 - 2 September 2013.
34. Takahashi, Tetsuro, Shuya Abe y Nobuyuki Igata (2011). "Can Twitter Be an Alternative of Real-World Sensors?" Human-Computer Interaction. Towards Mobile and Intelligent
35. Topsy Labs (2012), Topsy website. Available: http://www.topsylabs.com, Accessed 2012 November 20
36. Tropical Cyclone Report. Hurricane Sandy (AL182012) 22 – 29 October 2012 (2013). National Hurricane Center.
37. Van Hentenryck P (2013) Computational disaster management. Proceedings of the Twenty- Third international joint conference on Artificial Intelligence, (AAAI Press), pp 12-18.
38. Vieweg S, Hughes AL, Starbird K, & Palen L (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (ACM), pp 1079-1088.
39. Wang D, Lin Y-R, & Bagrow JP (2012) Social Networks in Emergency Response. Encyclopedia of Social Network Analysis and Mining).
40. Wang Q & Taylor JE (2014) Quantifying Human Mobility Perturbation and Resilience in Hurricane Sandy. PLOS ONE 9(11):e112608.
41. Watts D, Cebrian M, & Elliot M (2013) Dynamics of Social Media. Public Response to Alerts and Warnings Using Social Media: Report of a Workshop on Current Knowledge and Research Gaps, (The National Academies Press), pp 22-33.
42. World Bank (2012) FONDEN: Mexico's Natural Disaster Fund – A Review.
43. World Bank (2013) World Development Report 2014: Risk and Opportunity—Managing Risk for Development. Washington, DC: World Bank