

Pathways to Climate, Weather, and Environmentally Related Protest Events

Fang Jin^{a,1}, Wei Wang^a, Brian J. Goode^a, Chang-Tien Lu^a, and Naren Ramakrishnan^a

^aDiscovery Analytics Center (Dept. of Computer Science), Virginia Tech - NCR, Arlington, VA 22203

This manuscript was compiled on May 1, 2016

Abstract goes here.

Climate change | Extreme weather | Civil unrest | Evolution Pattern

1. 1. Introduction

Climate change, extreme weather, and the state of the environment directly impact the availability of food [1], energy [], and shelter []. As finite resources become scarce, the residual impacts on local economies can have disastrous and long-lasting effects on the fundamental livelihoods of inhabitants for decades []. In some cases, the resulting instability can severely detriment the ability of an established political system to maintain peace. The examples of this occurring are numerous. The extended drought in Syria in 2011 is cited as one of the principle causes of civil war [2, 3]. In a smaller scale example, the environmental impact of lead contamination in the drinking water in the United States led to protests in 2016. The extreme weather event, Hurricane Manuel, that devastated the western coasts of Mexico led to subsequent protests over resources at points as long as 1 year after the initial event.

Of course, the occurrence of either a shift in climate, extreme weather, or environmental catastrophe is not sufficient to guarantee that civil unrest is likely to follow. In general the causal mechanisms leading to civil unrest are very complex, and there is no easy way to determine a linear pathway to protest. However, to date, little quantitative analysis has been performed on the residual effects of changes resulting from climate, extreme weather, and the environment using a large volume of data. In this analysis, we focus on the breadth of the climate events by looking at events generated from a large Gold Standard Report (GSR) containing all of the protests that have occurred in Latin America from 2011-2013. By developing a logistic regression classifier, 25352 GSR civil unrest events were classified as either being climate or non-climate related using terms in the description of the event.

NEXT... WHAT DID WE FIND??? NOTE: THERE ARE A LOT OF FIGURES - these are observations - BUT WE NEED TO DRAW CONCLUSIONS. WHAT DO YOU LEARN FROM THESE FIGURES? HOW CAN YOU MAKE QUANTIFIABLE COMPARISONS? WHAT INTERESTING THINGS DO YOU FIND IN THE DATA. CURRENTLY, THERE IS SPATIAL, (TEMPORAL?), and "CAUSALITY". PLEASE TRY TO THINK ABOUT WAYS TO CHARACTERIZE THE DIFFERENCES/SIMILARITIES BETWEEN CLIMATE/NON-CLIMATE PROTESTS. ALSO, HOW WOULD YOU COMPARE ACROSS THE 3 COUNTRIES YOU'VE ANALYSED?

- First, we develop a logistic regression classifier, which can classify climate protests from non-climate protests automatically based on protest events description.

- Second, we analyze the climate protest spikes and disclose its relationship with climate disasters. For example, we plot the climate protests time series in Mexico and Brazil, and overlay with corresponding climate disasters, found for storm and hurricane events in Mexico, the protest time line last much longer. However, for drought events in Brazil, the protests being initiated more swift, also last shorter time.
- Thirdly, we figure out the proportion of protest causality. By studying some major climate disasters, we also discover each protest category's demands.
- Finally, We investigate the climate cocurrence. For instance, the water related protests are often accompanies with electricity shortage. while land protest are often isolated reasons.

2. 2. Related research

The path from climate, extreme weather and environmental effects to civil unrest is causally complex [4, 5] and involves various combinations of climate change [6], natural resources, human security, and social stability. In general, sensitivities to climate change, exposure to climate change, and the ability of a society to adapt are indicators of whether or not violence will erupt [7]. A commonly studied pathway is the effect of

Significance Statement

From 1995-2015, 90% of global disasters are caused by extreme weather events such as flood, storm, heat wave and so on. What is more, the extreme weather events are increasing continuously, resulting natural disasters happen more and more frequent. Records from the international disaster database EM-DAT shows, the average disasters per year from 2005 to 2014 is 14% higher than the years between 1995 and 2004, and is twice of years between 1985 and 1994. However, how the extreme weather events lead to armed conflicts is still under disclosed. This work presents a promising strategy to identify climate change related protests in South America, and further trace back protest causalities, and discover the mainly evolution patterns from climate change to civil unrest. This work can be extended into United States and other countries, to help decision makers understand the evolution path from extreme weather events to civil unrest, and benefit for society stabilization and human well-beings.

Please provide details of author contributions here.

Please declare any conflict of interest here.

¹To whom correspondence should be addressed. E-mail: jfang8@vt.edu

climate on food prices which then induces civil unrest. An examples of this occurrence is the Arab Spring uprisings in 2011, and how weather effects food prices [8]. The pathway to civil unrest is also not limited to a local region, where one study shows the Chinese drought effecting the supply wheat causing prices to rise in the Egyptian break market leading to protest [9]. The pathways of food prices to protest have also been studied in the global south [10], Africa, and Asia [11, 12]. However, even this path of climate effects on income level leading to conflict is not eminently clear [13].

This second paragraph will discuss work more related to the specific conclusions that we draw... we don't have these yet. Or do we? Anyway cite Hsiang here...

Protest work w/ EMBERS and related programs. Tools available for use. Brian to add in using text/refs similar to his other protest papers.

3. Climate Protests Overview

We study 25352 GSR civil unrest events across Latin American countries from July 2011 to March 2015. We build a climate change protest classifier, aim to identify how many civil unrest events are caused by climate change or extreme weathers.

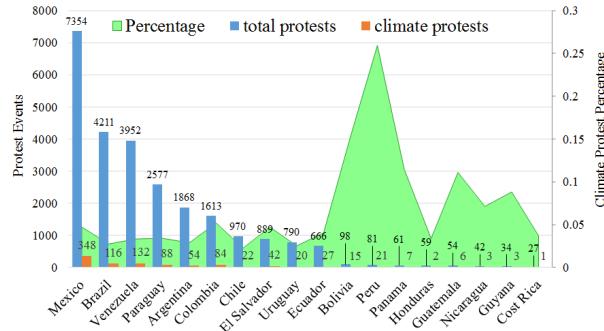


Fig. 1. Blue bar shows all the GSR protest events, yellow bar shows climate related protest events, green area shows the climate protest percentage over all the Latin American countries, from July 2012 to March 2015.

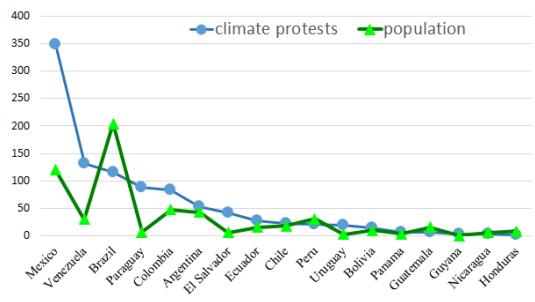


Fig. 2. Climate protest events and population (million) of each country. The two series have a Pearson correlation coefficient 0.64.

3.1 Climate protest distribution. Among all the 25352 GSR protest events, 991 events are classified into climate-related events. In other words, in Latin American countries, the climate related protests accounts for 3.91%. As shown in Figure 1, Blue bar shows all the GSR protest events, from the highest amount of Mexico 7354 to lowest country Cost Rica 27; yellow bar shows climate related protest events; the

blue area describes the climate related protests percentage over all the GSR protest events. Figure 3(a) plots the climate related protests events numbers, we can see Mexico has the most climate protests events, as high as 348. And Figure 3(b) gives us a straight view of climate protest percentage, of which, Peru climate protest accounts as high as 25.9%, Bolivia climate protest percentage reaches as high as 15.3%, Panama 11.4% and Guatemala 11.1% respectively.

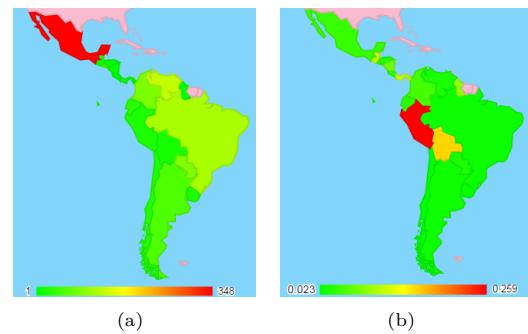


Fig. 3. (a) Climate related protests events numbers; (b) Climate related protests percentage in Latin American countries, from July 2012 to March 2015.

3.2 Spatial distribution. We pull out climate and non-climate events location, and show their distribution on the map, as shown in Figure 9. Generally, the climate protest distribution is closely correlated with population density, the higher population density area, the more protest events, regardless of climate or non-climate.

3.3 Protest time interval deviation. Figure 5 shows Mexico climate and non-climate protests time interval distribution. We found the non-climate protests have very high percentage with small time interval, less than three days. The climate protests, on the contrary, time interval is longer than non-climate protests.

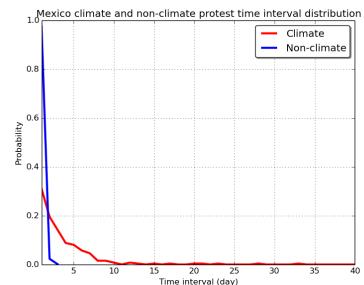


Fig. 5. Mexico climate and non-climate protests time interval deviation. Blue curve shows the non-climate protest time intervals percentage, and red curve shows climate protest time interval percentage.

4. Climate protests causality

4.1 Word cloud. Of the climate related protests, we are interested in what are the protesters demanding. To have a birds view of climate protests, we extract all the climate protest descriptions and plot the word cloud, as shows in Figure 6.

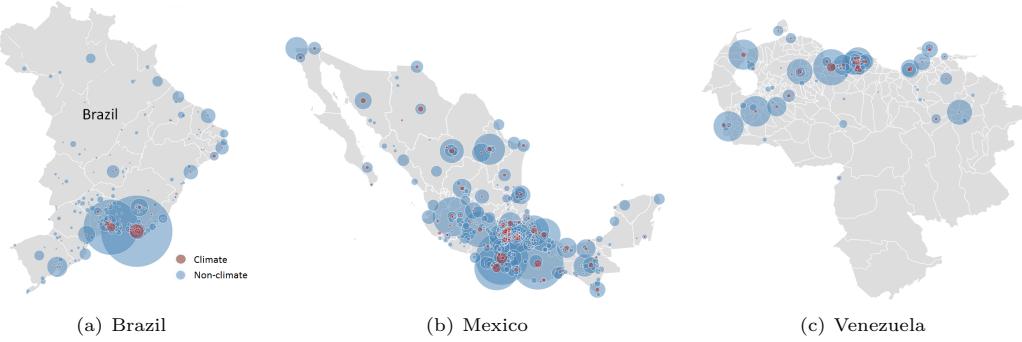


Fig. 4. Climate and non-climate protests from July 2012 to March, 2015. Red circle represents climate related protest events, and blue circle represents non-climate related protests.



Fig. 6. Word cloud of all the climate related protests, from GSR descriptions.

We can see words like ‘water’, ‘storm’, ‘mining’, ‘rain’, ‘construction’, ‘power’, ‘heat’, ‘gas’, ‘environment’, ‘electricity’, and other weather, environment related keywords are dominant, which gives us a general idea of what are protesters demanding.

4.2 Climate protests proportion. However, what is the specific protest reasons, how is the proportion of each protest category? We build a classifier to categorize the climate protest types, which generally falls into nine categories. As can be seen in Figure 7, the most dominant two categories are environment concern and lack of water. After that, the third commonest reason is about power, blackout. Also, extreme weathers like storm, hurricane, drought also accounts a considerable portion. The interesting thing is, each country has its own distinguishing protest features. In Mexico, the most notable protest reasons are lack of water, environment concern, storm and hurricane. In Venezuela, apart from lack of water, environment problem, the dominant reasons are about blackout and energy issue. In Peru, more than half of climate protests are demanding mining project, which is related with environment concern. While in Argentina, 35% events protest against blackout issue.

4.3 Climate protests causality analysis. The causality analysis is based on three steps, as described following.

a. Text enrichment. Messages with textual content (Tweets, Newsfeeds, Blog postings, etc.) are subjected to shallow linguistic processing prior to analysis. Applying BASIS technologies' Rosette Language Processing (RLP) tools, the language of the text is identified, the natural language content is tokenized and lemmatized and the named entities identified and classi-

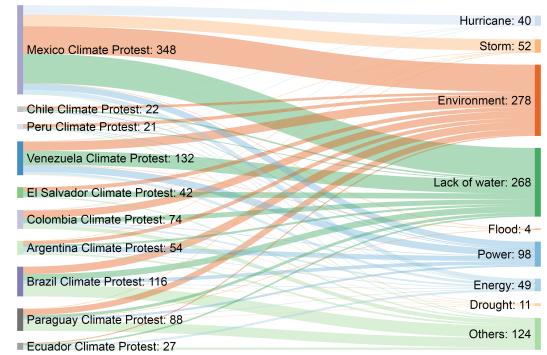


Fig. 7. Climate protest causality diagram. Left bar shows ten countries' climate protest numbers, and right bar shows nine climate event categories which cause climate protests.

fied. Finally, messages are geocoded with a specification of the location (city, state, country), being talked about in the message.

b. Construct knowledge graph In the scope of this study, an entity network is a graph $G(E, R)$ where entities $E = e_1, \dots, e_n$ can be linked to one another through relationships $R = r_1, \dots, r_n$ defined by conceptual interactions, and thus called a knowledge graph. knowledge graphs are used to represent its logic based on semantic networks.

How to represent knowledge and extract entities and their relationship from text? The general extraction patterns rule is Noun Phrase - Verb Phrase - Noun Phrase. We can also employ multilingual rules for phrase patterns learnt using association mining based on collocations. Using an open source, GraphDb (Neo4j) to store and query knowledge triples. Neo4j is a very robust, scalable solution for Knowledge representation, which provides an expressive query language : CYPHER, and a framework for implementing graph theoretic search algorithms. The general scheme is Subject Predicate Object. Subject is a concept, which can be person, organization, location, event, noun phrase; Predicate can be either a verb action or (predefined) relation type; Object can be either a concept or (some) value.

With all the enriched articles as input, since their named entities and verb phrases being identified and classified, the system can automatically extract the structural knowledge and load them into Graph database. It worths to note the entities may evolve along space and time.

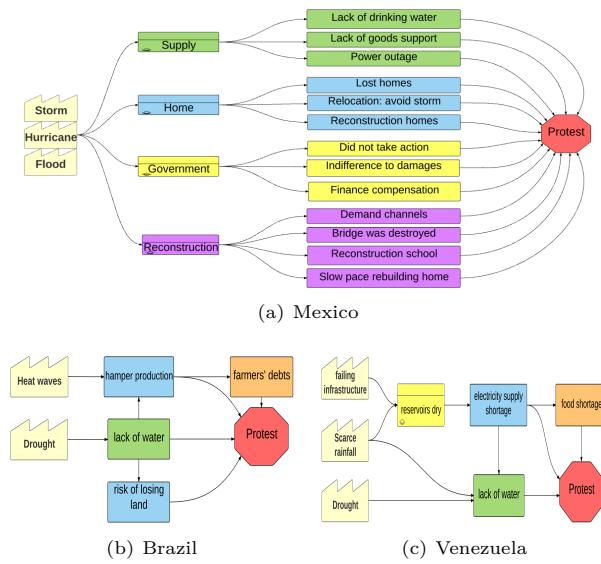


Fig. 8. Climate protest causality diagram

c. Query knowledge graph to identify causality. With a complete knowledge database, we are going to pose domain specific questions, for example, who are the main players in the article? what are the reasons for the protest? To such questions like that, we are seeking causal (base) relations between concepts. Specifically, we are interested in entities to be storm, hurricane, heat, draught, etc., climate related keywords. By matching the object or subject with climate related keywords, and predicate to be causality relationship like result of, cause by, lead to, blamed, accused of, demanding, we can locate and further identify the event-specific descriptors, get the relations between concepts (players and protest keywords) or relations of base type - temporal, change state, explicit facts.

match (a)-[relations]-(b) where a.name = '**climate.**', and a.relations = '**cause.**' return a, relations, b. We are interested in relations like 'cause', 'demanding', 'against', 'request', and so on.

For some severe and dominant climate events, such as storm, hurricane, flood, and drought events, we use the knowledge graph to identify their evolution pattern, thus from climate disasters, how does it evolve into armed conflicts? As shown in Figure 8, we illustrate storm caused protests demands in Mexico, blackout caused protest demands in Venezuela, and drought caused protest demands in Brazil.

5. Climate disasters and climate protests. We are also wondering, what is the interaction between climate change and civil unrest, how the climate change events evolves into social movements, what is the time span?

Mexico climate disasters Let us see an example of Mexico climate disasters and protests, as shown in Figure 9(a). One major climate disaster occurred in September 17, 2013, tropical storms Manuel and Ingrid hit Mexico, more than 23,000 people fled their homes in the state due to heavy rains spawned by what had been Hurricane Ingrid, and 9,000 went to emergency shelters, at least 20 highways and 12 bridges had been damaged.

After the storm Manuel, related protests and other civil unrest events break out dozens of times, and lasts for more

than two years because the government's response had been desperately inadequate. The related protest reached climax in January 2014, and second climax in April 2014. We can see on November 19, 2013, there was report saying 'it's been 63 days since the onslaught of 'Ingrid' and 'Manuel' and families were left homeless are still without help'. On January 22, 2014, news reporting 'Cards require delivery of goods. Four months after the storm 'Manuel' and the effects of Hurricane 'Ingrid', they say 'we have not received anything'. On April 7, protest description saying 'Affected by Tropical Storm 'Manuel' in the municipal head of Tixtla marched to demand the construction of a controlled channel, it will prevent a flood like that caused the overflow from the Black Lagoon in September 2013'. The related climate protests descriptions are shown in Figure 10.



Fig. 10. Word cloud of Mexico storm Manuel, Sept 13, 2013.

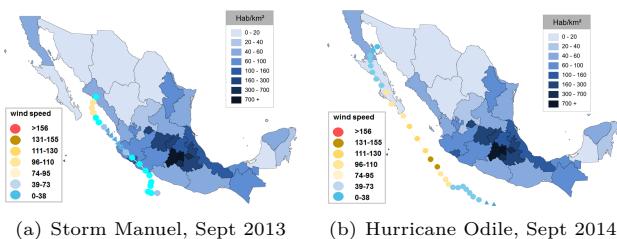


Fig. 11. Track map of Tropical Storm Manuel of the 2013 and Hurricane Odile of the 2014 Pacific hurricane season. The points show the location of the storm at 6 hour intervals. The colour represents the storm's maximum sustained wind speeds as classified in the Saffir Simpson hurricane wind scale, and the shape of the data points represent the nature of the storm. The map shows population density of all Mexico's 32 states.

The interesting thing is, after hurricane Odile 2014, there was not many related protests. Why of the two storms in Mexico, storm Manuel 2013 caused a series of climate protests, but storm Odile 2014 did not cause considerable protests? Storm 2014 belongs to category 3, much severe than storm 2013 (which belongs to category 1). However, from Figure 11, we can see storm Manuel 2013 is very close with Mexico's residence area, while hurricane Odile 2014 is far from Mexico coast, even though it cross some part of Mexico, but the population density there is very low. This can explain why storm 2013 lead to tremendous protests, while hurricane 2014 does not.

Brazil climate disasters Figure 9(b) shows Brazil climate disasters and climate protests relationship. The three red bar shows three drought events in Brazil, which caused drought related protests immediately. The drought in February 2012 hampered production, which arouse farmers protest. The heat wave in February in 2014, and drought in October 2014 results in water shortage, thus a series of protest abrupt in Brazil. The biggest spike in June 2013 are protests against govern-

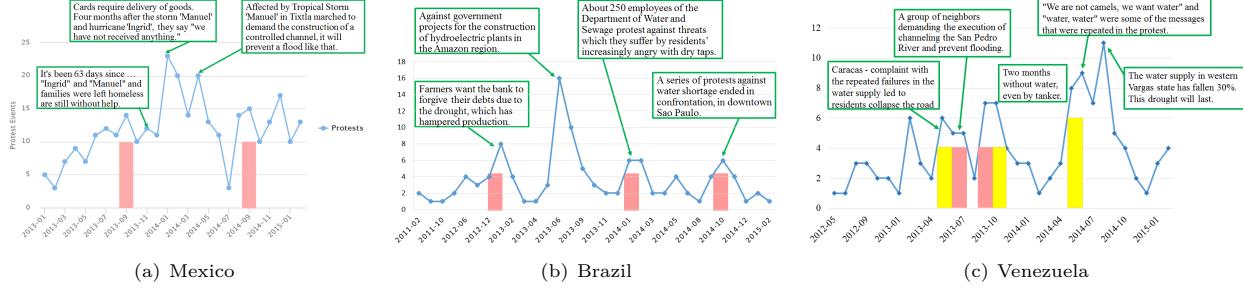


Fig. 9. (a) Mexico climate disasters and climate protests. The blue time series shows the climate related protest events, and light red vertical lines show two climate disasters in Mexico, storm Manuel in September 17, 2013 and hurricane Odile in September 15, 2014 respectively. (b) Brazil climate disasters and climate protests. The blue time series shows the climate related protest events, and light red vertical lines show three climate disasters in Brazil, drought in Feb 2012, Heat wave in Feb 2014, and drought in Oct 2014, respectively. (c) Venezuela climate disasters and climate protests. The blue time series shows the climate related protest events, and light red vertical lines show flood disasters, and yellow vertical lines drought disasters.

ment's projects for the construction of hydroelectric plants in the Amazon region, which can be ascribed into environment category.

Figure 9(c) shows Venezuela climate disasters and climate protests relationship. In June 2013, sudden torrential rains, a phenomenon associated with climate change, cause a heightened risk of flooding and landslides in the densely populated communities on the outskirts of Caracas. In May 2014, a drought has triggered rationing of tap water in the capital, Caracas, where residents must form lines that last hours to fill jugs of water for drinking, showering, and other needs.

6. Climate protest pattern According to the protest content, we cluster each country's protests type. Based on protest descriptions, we calculate two descriptions text similarity, and assign weight between the two protest IDs. Specially, we pay special attention to the protest themes or protest demands, if two descriptions have the same protest demanding, they will have high weight, otherwise, if their protest demanding are different, their connection weight tends to be 0. In this way, we build a weighted undirected network $G(V, E, W)$, with each protest as node V , and their connection as edge E , their weight as W . If the weight between two nodes is 0, their will be no edge. We employ Louvain method [14] to split the network into several clusters.

For Venezuela's climate protests, as can be seen in Figure 12(c), the yellow cluster which represents lack of water protests accounts for 55.8%, the green cluster which denotes power outage accounts for 22.1%, and the blue cluster which stands for gas shortage accounts for 5%, the purple cluster shows the rest climate protest portion, which include food shortage, medicine shortage, water tank robbery behavior, etc..

7. Climate protests in Twitter We are also interested in climate events influence on social media, such as Twitter. Using keywords list we are able to filter tweets, then cluster tweets into different partitions based on similarity among tweets using distance function, taking tweets content, geolocation and other features into consideration. Each partition includes similar tweets stand for a specific event. As shown in Figure 13(a), there are three distinct extreme weather event type in Mexico in different locations, the word cloud shows discussion on Twitter as per that event.

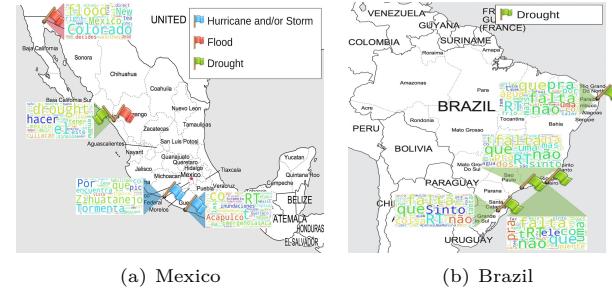


Fig. 13. Climate protest events in Mexico, Sept 2013 and Brazil, May 2012. Different flag represents different climate disasters. The adjacent world cloud shows Twitter discussion as per that event.

Table 1. Classification methods comparison.

	Precision	Recall	F_measure	Accuracy
Majority assign	0.1274	0.1289	0.1258	0.8136
KNN	0.1906	0.4913	0.2723	0.7154
Naive Bayes	0.2432	0.8779	0.3798	0.6777
Weighted SVM	0.6543	0.5565	0.5966	0.9218
Logistic Regression	0.7513	0.5102	0.6018	0.9322

Discussion

Using human analysts, MITRE organizes a gold standard report (GSR) of protests by surveying newspapers for reportings of civil unrest. From GSR events, we devised a climate protest classifier to identify the climate related protest events automatically.

The classifier is built based on logistic regression model. With input of GSR descriptions, we aim to train a classifier which can label a protest description as climate related or not. The GSR includes many potential important features, such as status, description, crowd size, headline, event Type, event Date, location, date, population etc. The description feature is brief description of the events, which plays a dominant role in the entire dataset. In order to adopt logistic regression on GSR dataset, we need to vectorize text data in the dataset. First of all, we construct a word corpus which includes every word x_i shown in the training dataset (includ-

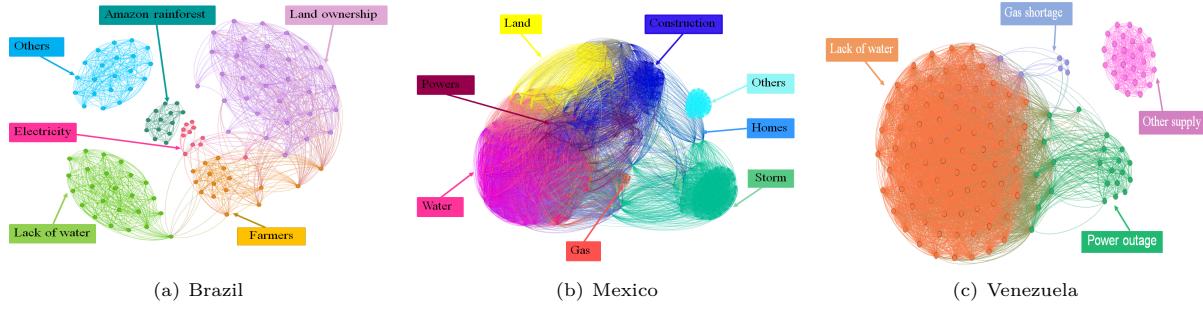


Fig. 12. Climate protest clustering results

ing non-words). We accept non-words because most coinages come from Internet and some of them might be important for the events. As we accept non-words, the corpus might be large than our corpus vocabulary. The word corpus is composed with $[x_1, x_2, \dots, x_i, \dots, x_N]$. Second, take each GSR description as a vector, we assign values to each vector, if x_i appeared in GSR record, the corresponding value will be assigned as 1, otherwise 0. In this way, every GSR record being converted to a corresponding vector based on the corpus. Third, set climate protest as $Y = 1$, non-climate protest as $Y = 0$, the weight for each term x_i as k_i , then $Y_j = \sum_{i=1}^N k_i x_i$. By training process, we calculate the weight k_i for each term x_i . The last step is test. Given a new GSR description, the probability of classification is:

$$P(Y = 0|X) = \frac{1}{1 + \exp(\sum_{i=1}^N k_i x_i)}$$

$$P(Y = 1|X) = \frac{\exp(\sum_{i=1}^N k_i x_i)}{1 + \exp(\sum_{i=1}^N k_i x_i)}$$

We manually labelled 1700 GSR protest records as climate or non-climate protests. Using 70% dataset as training, and the rest 30% as test. To ensure we have a trustworthy classification results, we evaluate the performance carefully by cross evaluation. The evaluation criteria are precision (positive predictive value), recall (true positive rate), F-measure (a measure that combines precision and recall) and accuracy

(the proportion of true results both true positives and true negatives among the total number of cases examined). We compare with four well-known classification methods: majority assign, K-nearest neighbor, Naive Bayes, and weighted support vector machine (SVM). Since the climate events account for a small portion of all the events, which make it an unbalanced classification problem, so we change the traditional support vector machine into weighted SVM, by adding more importance to the climate protest events (we set the class weight to be 100). From Table 1, we prove our logistic regression method outperforms other methods uniformly.

Conclusion We build a climate protest classifier based on the logistic linear regression model, which achieves the best performance comparing with other classical classifiers. We analysis the climate related protest both temporally and geographically. Further more, we investigate the climate protest causalities, and show that the shortage of water and electricity tops the climate protests. Finally, we also demonstrate the climate protests patterns.

ACKNOWLEDGMENTS. Supported by the Intelligence Advanced Research Projects Activity (IARPA) via DoI/NBC contract number D12PC000337, the US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the US Government.

1. Antwi-Agyei P, Fraser ED, Dougill AJ, Stringer LC, Simelton E (2012) Mapping the vulnerability of crop production to drought in ghana using rainfall, yield and socioeconomic data. *Applied Geography* 32(2):324 – 334.
2. Gleick PH (2014) Water, drought, climate change, and conflict in syria. *Weather, Climate, and Society* 6(3):331–340.
3. Kelley CP, Mohtadi S, Cane MA, Seager R, Kushnir Y (2015) Climate change in the fertile crescent and implications of the recent syrian drought. *Proceedings of the National Academy of Sciences* 112(11):3241–3246.
4. Hsiang SM, Meng KC, Cane MA (2011) Civil conflicts are associated with the global climate. *Nature* 476(7361):438–441.
5. Scheffran J, Brzoska M, Kominek J, Link PM, Schilling J (2012) Climate change and violent conflict. *Science* 336(6083):869–871.
6. Burke M, Hsiang SM, Miguel E (2014) Climate and conflict, (National Bureau of Economic Research), Technical report.
7. IPCC (2007) *Climate Change 2007: Climate Change Impacts, Adaptation and Vulnerability*. (Cambridge University Press).
8. Johnstone S, Mazo J (2011) Global warming and the arab spring. *Survival* 53(2):11–17.
9. Sternberg T (2012) Chinese drought, bread and the arab spring. *Applied Geography* 34:519 – 524.
10. Bush R (2010) Food riots: Poverty, power and protest1. *Journal of Agrarian Change* 10(1):119–129.
11. Wischnath G, Buahug H (2014) On climate variability and civil war in asia. *Climatic Change* 122:709–721.
12. Hendrix C, Haggard S, Magaloni B (2009) 1 grievance and opportunity: Food prices, political regime, and protest.
13. Sarsens H (2011) Rainfall and conflict.
14. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008(10):P10008.