

Pathways to Climate 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

Data from 25352 civil unrest events in Latin American countries, from Jan 2011 to March 2015, are analyzed to estimate the pathways from climate disaster to climate related protests. With all kinds of protest records in hand, we build a text classifier based on logistic regression model to screen the climate related protest events. This methodology has been approved to be the most efficient one compared with other four classical classifiers, with accuracy of 0.932 and F-measure of 0.602. Based on the classification results, we investigate all the climate motivated protests in Latin American countries and present a detailed analysis including climate protest frequency, spatial distribution, and temporal dependency. In particular, we identify climate protest spikes have a clear temporal dependency upon climate disasters, with different time span according to disaster types. Furthermore, we figure out the proportions of protest type for each country and discover protest causalities which illustrate the pathways from climate events to armed conflicts. We also find that, each country has its own characteristic climate protest pattern, which disclose the coherent relationship among different protest demands. This study provides quantity analysis that can help investigate the evolution pattern from climate change to civil unrest, also help understand climate change and its influence on society.

Climate change | Extreme weather | Civil unrest | Protest pattern

1. Introduction

Climate change, extreme weather, and the state of the environment directly impact the availability of food [1] [2], energy [3], and shelter [4]. As finite resources become scarce, the residual impacts on local economies can have disastrous and sometimes long-lasting effects on the fundamental livelihoods of inhabitants for decades [5]. 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 [6, 7]. In a smaller scale example, the environmental impact of lead contamination in the drinking water in the United States led to protests in 2016¹. As we later show as an example of extreme weather, tropical storm Manuel devastated the western coasts of Mexico leading to subsequent protests over resources at times as long as 17 months after the initial event.

The path from climate, extreme weather and environmental effects to civil unrest is causally complex [8, 9] and involves various combinations of climate change [10], 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 [11]. A commonly studied pathway is the effect of climate on food prices that then induces civil unrest. An example of this occurrence is the Arab Spring uprisings in 2011, and how weather effects food prices [12]. 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 [13]. The pathways of food prices to protest have also been studied in the global south [14], Africa, and Asia [15, 16]. However, even this path of climate effects on income level leading to conflict is not eminently clear [17].

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) [18] containing all of the protests that have occurred in Latin America from 2011-2015.

We address three foundational problems. First, we use machine learning to classify climate related protests. 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. Second, we use the textual description of protests to extract the climate protest category for protests in each country. For each major climate category, we adopt the knowledge graph approach to define linkage relationship between entities, and study the potentially related protest attributes. Third, we find that the massive climate protests show that certain protest types are shown to correlate with other protest types. Specifically, we find clues such as a lack of water is highly linked with power shortage. This is the first large-scale study of climate related

Significance Statement

Although much evidence has accrued on the strong causal associations between climate change and human wellbeing and security, important gaps remain in our understanding of the interaction between climate change and civil unrest. Improving the understanding of these dynamics as well as developing the mechanism to identify the climate related civil unrest events, is critical for developing interventions and adaptations to mitigate these risks. This study integrates a collection of disciplinary perspectives to identify climate change related protests in Latin America, and further trace back protest causalities, and discover the main 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.

The authors declare no conflict of interest.

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

¹<http://www.cnn.com/2016/01/11/health/toxic-tap-water-flint-michigan/>

protests to our knowledge.

2. Climate Motivated Protests

There were a total of 25352 recorded civil unrest events in Latin American countries from Jan 2011 to March 2015 that were included in our dataset. Using our climate protest classifier, we were able to separate out protests directly or indirectly resulting from a major climatic, severe weather, or environmental event. In the subsequent analysis, these three categories of event types are labeled with a common definition of “climate event”. Of the candidate civil unrest events, 991 (3.9%) events are classified as climate-motivated across all Latin American countries for that time period. In the subsequent sections, we conduct a multi-dimensional analysis of these protests to understand potential implications of the breadth of impact resulting from climate motivated protests.

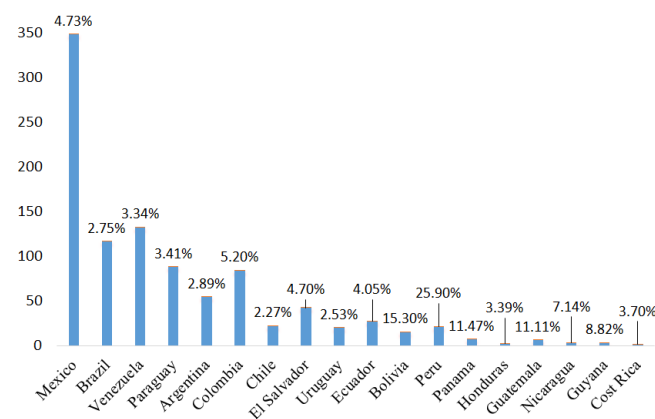


Fig. 1. Blue bar shows all the climate related protests for each country, yellow bar shows the climate protest percentage over its total protest, from Jan 2011 to March 2015.

Frequency Analysis by Country. The first analysis we conduct is a comparison of the representative number of protests within and across each country. The results of our classifier selection show the number of climate motivated protests and the percentage over all the protests in that country, as can be seen in Figure 1. The country with the most climate protests overall is Mexico, and Costa Rica has the least. As evidenced by the climate to non-climate protest ratio, the portion of protests related to climate remains fairly constant across countries with the exception of Peru. In this particular case, there were numerous protests centered on mining and its effect on the environment that dominate the overall protest landscape. As the number of total protests decrease, we see more variability in the ratio as expected. For these countries, which typically have smaller populations, the significance of a single type of protest has more of an impact on the measure than larger countries.

To show the effect of the population on the number of climate protests, we plot the result of a linear regression in Figure 2. The result of this shows an $R^2 = 0.70$, showing a slight linear relationship. However, the interesting part of this analysis lies in the residual errors. The set of countries including Mexico, Venezuela, Paraguay, and Colombia, all demonstrate the occurrence of more climate protests than

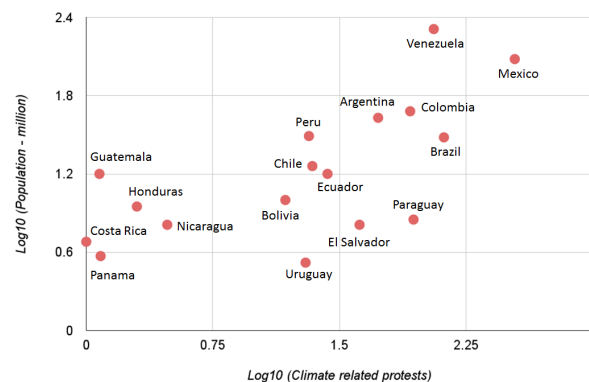


Fig. 2. Log10 (Climate protest events) and log10 (population - million) of each country. The two series have a Pearson correlation coefficient 0.70.

would be expected given the entire dataset. On the contrary, Brazil has fewer climate protests given the size of their population. There could be a number of reasons for these findings such as socio-political stability, environmental sensitivity, and the type of climate events. All of these are potential avenues for further causal or anecdotal studies. In the following, however, we choose Brazil, Mexico, and Venezuela for further analysis into overall trends of climate protests, and how these are shaped in the data recovered by the classifier.

Spatial Distribution of Climate Motivated Protests. In this manuscript we are defining the climate protest as being different from a regular civil-unrest event by a relation to an climate event. Next, we investigate if there is any fundamental difference in terms of where these protests occur in relation to protests in general. For this analysis we use Mexico, Brazil, and Venezuela which all have many protest events, and the percentage of those that are related to the climate are all at about 4%. The spatial distribution of events is shown in Figure 3. Both the total number of protests and those that are climate motivated are shown and represented by the size of the blue and red shaded circles, respectively.

In both Brazil and Venezuela, many of the protests appear at or near their coastal boundaries, and Mexico has more inland activity. However, we have already established a connection between population and protests. This is no different for the spatial distribution, where much of the population of Brazil and Venezuela is located in coastal regions. The protests in Brazil mainly center at two major cities Sao Paulo and Rio de Janeiro. In Mexico and Venezuela climate protests have a more uniform distribution across the cities. Therefore, there is no particularly strong evidence to suggest that certain regions of these countries are more prone to protest with respect to the climate than they would normally be willing to protest in general. In terms of the climate events defined in this study, effects of climate, the environment, and extreme weather are not regionally exclusive to certain populations. Through complex channels such as food supply, the effects of climate impact can ripple across spatial networks.

Temporal Dependency on Climate Events. The temporal dependency of climate protest occurrences is analyzed for each country. As with the spatial domain, the effects of climate events are non-local in time in some cases. The ground truth

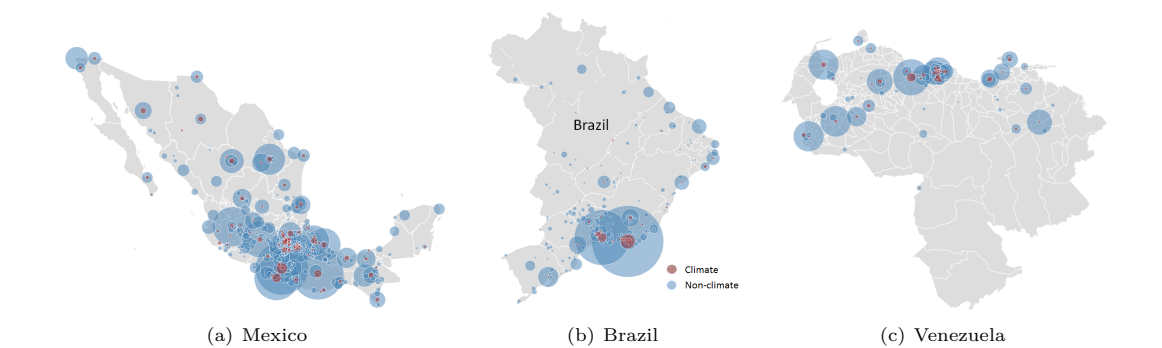


Fig. 3. Climate and non-climate protests from Jan 2011 to March, 2015. Red circle represents climate related protest events, and blue circle represents non-climate related protests.

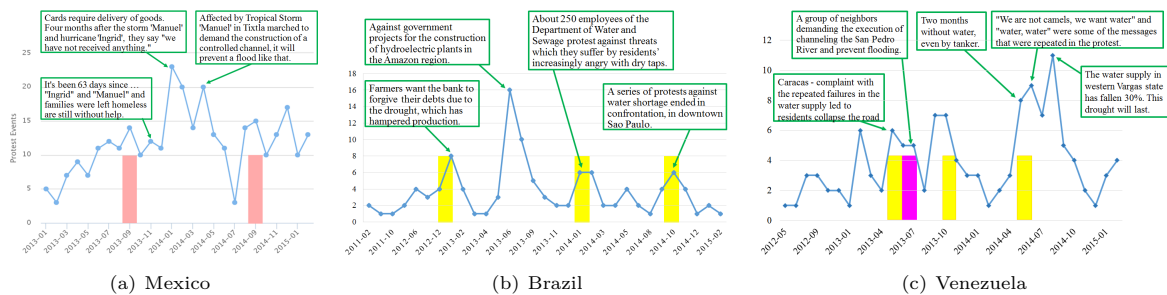


Fig. 4. (a) Mexico climate disasters and climate protests. The blue time series shows the climate related protest events, and light red vertical lines show two storm disasters in Mexico, storm Manuel in September 17, 2013 and hurricane Odile in September 15, 2014 respectively. (b) Brazil climate disasters and climate protests. The blue time series shows the climate related protest events, and yellow vertical lines show three drought 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 rose vertical line show local area flood disaster, and yellow vertical lines drought disasters.

for the events was established for extreme weather only, as the event itself is more local in time than climate and environmental changes. This data is available by combining the following sources: International Disaster Database EMDAT², World Disasters Timeline³ and European Commission's Humanitarian Aid and Civil Protection department (ECHO)⁴. The official climate disaster report for each country is shown with climate related protests in Figure 4.

Mexico climate disasters Figure 4(a) shows the Mexican extreme weather events and protests where the blue time series represents the climate protests events, and the two red bars shows the occurrence of two storms. The first storm is the combined tropical storm Manuel (category 1) and hurricane Ingrid in September 17, 2013. The track maps can be seen in Figure 5(a). Tropical storm Manuel crossed the west coast of Mexico and resulted in more than 23,000 people fleeing their homes due to heavy rains spawned by what had been Hurricane Ingrid. Of those displaced 9,000 went to emergency shelters. In terms of infrastructure, at least 20 highways and 12 bridges had been damaged⁵. After the storm, related protests and other civil unrest events broke out and lasted for more than 17 months because the government's response had been inadequate. The storm related protests reached a climax in January

2014, and second climax in April 2014. 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"⁶ Four months after the storm Manuel and the effects of Hurricane Ingrid, they say "we have not received anything". On April 7, protest descriptions said "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 last protest event we have on record from the climate protest classifier occurred 17 months after the original event. This demonstrates that the residual capacity of these events to impact the livelihoods of people is not guaranteed to be local in time. As we show, the range of impact can extend even beyond the occurrence of other storms.

In Figure 4(a), the second red bar shows hurricane Odile. It is a category 3 storm that occurred in 2014, and the track of the storm's path is shown in Figure 5(b). Despite hurricane Odile being a more intense storm, there were not many protests related to the event. Comparing the storm's paths in Figure 5, Tropical Storm Manuel hit Mexico's mainland, which caused more destruction. Hurricane Odile 2014 had less of an impact on the Mexican mainland, even though it crossed the state of Baja California. However, this is the second smallest Mexican state by population. This can explain why storm 2013 lead to

²<http://www.emdat.be/database>

³<http://www.mapreport.com/>

⁴<http://ec.europa.eu/echo/>

⁵<https://weather.com/storms/hurricane/news/tropical-storm-manuel-hurricane-ingrid-hit-mexico-opposite-coasts-20130916>

⁶Quotes are translated from the native language of the country.

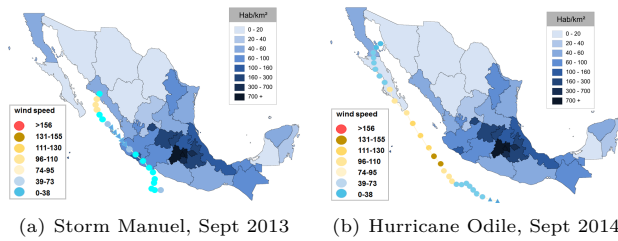


Fig. 5. The map shows population density of all Mexico's 32 states. The track shows tropical storm Manuel of 2013 and hurricane Odile of 2014, points in different color represent the wind speeds.

tremendous protests, while hurricane 2014 does not.

Brazil climate disasters Figure 4(b) shows the relationship between protests classified by our algorithm and actual extreme weather events in Brazil. The three yellow bars show three separate drought events in Brazil, which resulted in drought related protests almost immediately. The drought in February 2012 hampered production, which caused farmers to protest. The heat wave in February in 2014, and drought in October 2014 resulted in water shortages, causing civil unrest. The biggest spike in June 2013 described protests against government's projects for the construction of hydroelectric plants in the Amazon region⁷ and is more of an environmental impact type of event. In general, for these events we see predominantly local relationships in time between the protest and the preceding event. For Brazil in particular, the extreme weather event matches fairly well with the onset of drought.

Venezuela climate disasters In Figure 4(c), the climate motivated events are shown in relation to relevant extreme weather events for Venezuela. The pink bar represents sudden onslaught of rain in June 2013 that caused a heightened risk of flooding and landslides in the densely populated communities on the outskirts of Caracas. It triggered a small portion of protests to prevent flooding. The yellow bars denote drought disasters. The drought in May 2014 triggered rationing of tap water in the capital, Caracas, where residents formed lines lasting hours to fill jugs of water⁸. This drought disaster lasted so long that related protests reached a climax in September 2014. Unlike Brazil, the data in Venezuela on droughts proved tough to ascribe to a particular drought event. They occur rather frequently and there is a substantial amount of overlap in the residual protest events that it was difficult to distinguish to which it was referring.

3. Analysis of Protest Descriptions

As stated previously, we are not blind to realizing that the causes of climate motivated protests are in general complex. In the following, we analyze the descriptions of the protest events in order to gain insight into the general pathways by which protests within our corpus have occurred. Shown in Figure 6 is a weighted Sankey diagram showing the bipartite graph of the most common keywords in the descriptions of protests from each country. Apparently, many of the protests

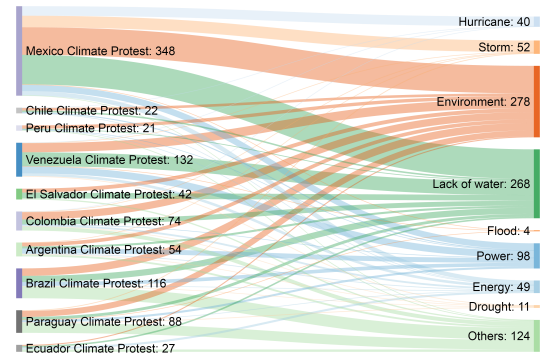


Fig. 6. Climate motivated protest keywords diagram. Countries on the left are matched with keywords appearing in the description of the protest event on the right.

identified by the classifier in one way or another have something to do with lack of water followed by climate effects in general. Other prominent keywords include mentions of power and energy issues. Each country also exhibits its own protest keyword categories. In Mexico, the most notable protest keywords involve lack of water, environmental concern, storm and hurricane. In Venezuela, apart from lack of water and environment problems, the dominant keywords are blackout and energy issues. In Peru, more than half of climate protests are about a mining project, which is an environment concern. While in Argentina, 35% events protest against blackout issues. We expand on these observations in the following where we analyze several dimensions of the keywords to extract details about pathways to protest.

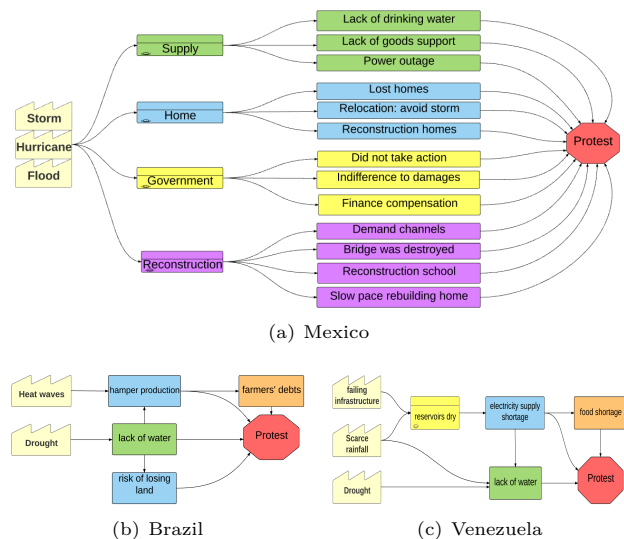


Fig. 7. Climate protest causality diagram

Pathways to Climate Motivated Protest. For some severe and dominant climate events, such as storm, hurricane, flood, and drought events, we employ the knowledge graph to represent the link relationships between entities. By matching the object or subject with climate related keywords, and sentence predicate to a causal keyword like “result of”, “cause by”, “lead to”, “blamed”, “accused of”, “demanding”, “against”, or “request”, we can locate and further identify the related occurrences of

⁷ <http://www.bloomberg.com/news/articles/2013-06-05/protests-over-brazil-hydropower-leads-to-delays-and-boosts-costs>

⁸ <http://www.breitbart.com/national-security/2014/05/31/severe-scarcity-prompts-venezuelan-government-to-ration-water/>

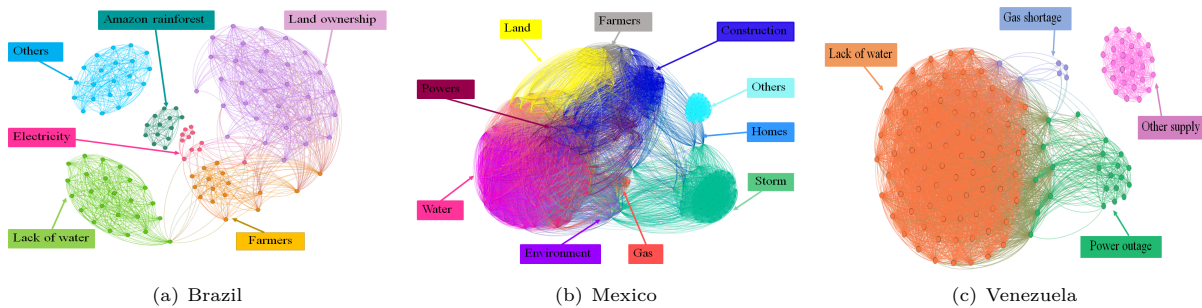


Fig. 8. Climate protest clustering results

most protest descriptors.

Figure 7(a) shows how the tropical storm caused protest demands in Mexico, which generally fall into four categories: supply, home, government and reconstruction. In the supply related protests, many co-occurring events are: lack of drinking water, lack of goods support, and power outage. The second category is about homes, they protest either because of lost homes, requests to relocate to avoid storm, or requests to reconstruct homes. Another protest type targets the government. They fight because government did not take action, blame government's indifference to damages, or request finance compensation to the damages. In the reconstruction category, residents demand reconstruct channels to avoid more storms, request to reconstruct bridges, roads, schools, or are unsatisfied with the slow pace of rebuilding homes. Figure 7(b) describes the co-occurring events of Brazil climate related protests. One possibility is that heat waves hampered production, which led to the farmers' protests over losing land, and the other possibility is that droughts led to residents having a lack of drinking water. Thus, both pathways lead to protest. In Venezuela, the protests are more centered around hydroelectricity, as shown in Figure 8(c). Scarce rainfall, drought and failing infrastructure makes water shortages and blackouts a common occurrence in Venezuela. The electricity shortage deteriorates with water shortages, leading to food shortage, and ultimately reduces food quality.

Climate protest patterns The above analysis shows events that commonly co-occur with protests; however, we intend to further illuminate the correlations surrounding protest activity. In the following we take a graph theoretic approach, where we treat each protest event as a node and connect two nodes with a weight based on their protest descriptions and text similarity. Specially, we pay attention to the protest themes or protest demands. If two descriptions have the same protest demand, their weight will be very high. Otherwise, their connection weight tends to be 0. In this way, we build a weighted undirected network and employ the Louvain method [19] to split the network into several clusters.

We show in Figure 8 the climate protest clustering results that provides the protest proportion and coherent correlations among different protest types. Figure 8(a) illustrates Brazil's climate protest pattern. We see that the largest protest cluster is about land ownership which accounts for 26.7% of the protests. The second cluster is lack of water and takes up 20.7%. The farmers cluster occupies 13.8%. We note that land and farmer clusters are closely correlated, and lack of water

is also closely binded to farmers as well. Amazon rainforest is another striking protest which is responsible for 11.3%. Figure 8(b) shows the protest patterns of Mexico, which has the most climate protest events and most complex patterns. We can see the red cluster which denotes lack of water is the most dominant protest, accounting for 20.5%. The green cluster represents the tropical storm, and is the second largest protest type, taking up 19.0%. The dark blue cluster construction accounts for 17%, and the yellow cluster land is responsible for 11.8%. We find that water related protests are intertwined with environment protests and power protests. Land protests are closely related with farmers, while construction clusters are coherent with homes (2.6%). Figure 8(c) gives the overview clustering results of Venezuela climate protests. We see the yellow cluster representing lack of water protests takes up the largest portion, as high as 55.8%, and the green cluster denoting power outages accounts for the second part at 22.1%. The blue cluster, which stands for gas shortage accounts for 5%. The purple cluster shows the rest climate protest portion, which includes food shortage, medicine shortage, water tank robbery behavior, etc. The lack of water protest is intertwined with power outage protest, which corresponds to the fact that lack of water and power shortage is everyday life in Venezuela.

Discussion

Climate changes, extreme weather and environmental catastrophes can all exert a devastating amount of harm to people around the world. To better understand how these events lead to protest behaviors, we show different pathways to protest following severe events in Latin America from 2011 to 2015. Our analysis differs from those previously published in that we consider the breadth of climate protests over a wide spatial and temporal domain. This is accomplished by identifying climate related protests using a logistic regression classifier acting over keyword vectors of protests descriptions in our protest GSR dataset. We found this approach achieved an F-score of 0.60 and accuracy of 0.93, which was the best performing of other common binary classifiers. The results of the classifier indicate a number of broad properties about climate related protests.

From our analysis, we found different climate disasters may cause related protests with different time span. For example, the tropical storm Manuel in Mexico was followed by a wave of climate related protests lasting as long as 17 months. In Venezuela, the protests caused by one drought always overlap with the other drought. This paper discloses protest causalities in Latin American countries, and illustrates the pathways from

climate disasters to climate protests. This paper also identifies the climate related protest patterns, and discovers the coherent relationships among different protests.

Materials and Methods

Dataset: GSR The GSR is a gold standard report of protests organized by MITRE [18], using human analysts, to survey newspapers for reportings of civil unrest. The GSR includes many features such as protest location, event date, protest type, status, crowd size, headline, date, population, protest description, first reported links, etc. The description feature is a brief description of the protest. It contains information like who, where, why and when surrounding protests.

Climate protest classifier **a. Logistic regression classifier.** The classifier is built based on logistic regression. To reduce the computational complexity, we only apply the GSR description text as input to the logistic regression classifier. To do this we first constructed a bag of words from the training dataset descriptions by deleting meaningless stop-words, like "the", "a/an", "at", and etc. The bag of words is composed of M words denoted as $[w_1, w_2, \dots, w_M]$. Each GSR description X is considered as a vector of length M . If word w_i occurs in its description, then $X(i)$ will be assign with 1; otherwise 0. Furthermore, each protest in the training dataset is assigned $Y = 1$ as climate protest, or $Y = 0$ as non-climate protest by manually checking its description meaning. In this way, each GSR record is converted to a corresponding vector based on the bag of words. We estimate k_i and b based on maximum likelihood criterion. This process is usually converted to convex optimization problem with efficient solutions [20]. Once the coefficients of k_i and b are estimated, the probability for each class is calculated by:

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

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

If the probability of $P(Y = 1|X = x)$ is larger than 0.5, then the protest event will be classified as climate protest, otherwise, non-climate protest.

b. Evaluation. We manually labeled 1700 GSR protest records as climate or non-climate protests. We used 70% of the dataset as training, and the rest as test data. To ensure that we have trustworthy classification results, we evaluate the performance carefully by cross validation. 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). The climate events only 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). The experiment results are shown in Table 1, where we demonstrate logistic regression method outperforms other methods uniformly.

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.

- 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.
- Akil L, Ahmad HA, Reddy RS (2014) Effects of climate change on salmonella infections. *Food-borne pathogens and disease* 11(12):974–980.
- Mitigation CC (2011) Ippc special report on renewable energy sources and climate change mitigation.
- Warner K et al. (2009) In search of shelter: Mapping the effects of climate change on human migration and displacement. *In search of shelter: mapping the effects of climate change on human migration and displacement*.
- Le Billon P (2001) The political ecology of war: natural resources and armed conflicts. *Political geography* 20(5):561–584.
- Gleick PH (2014) Water, drought, climate change, and conflict in syria. *Weather, Climate, and Society* 6(3):331–340.
- 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.
- Hsiang SM, Meng KC, Cane MA (2011) Civil conflicts are associated with the global climate. *Nature* 476(7361):438–441.
- Scheffran J, Brzoska M, Kominek J, Link PM, Schilling J (2012) Climate change and violent conflict. *Science* 336(6083):869–871.
- Burke M, Hsiang SM, Miguel E (2014) Climate and conflict, (National Bureau of Economic Research), Technical report.
- IPCC (2007) *Climate Change 2007: Climate Change Impacts, Adaptation and Vulnerability*. (Cambridge University Press).
- Johnstone S, Mazo J (2011) Global warming and the arab spring. *Survival* 53(2):11–17.
- Sternberg T (2012) Chinese drought, bread and the arab spring. *Applied Geography* 34:519–524.
- BUSH R (2010) Food riots: Poverty, power and protest1. *Journal of Agrarian Change* 10(1):119–129.
- Wischnath G, Buhaug H (2014) On climate variability and civil war in asia. *Climatic Change* 122:709–721.
- Hendrix C, Haggard S, Magaloni B (2009) 1 grievance and opportunity: Food prices, political regime, and protest.
- Sarsons H (2011) Rainfall and conflict.
- Ramakrishnan N et al. (2014) 'beating the news' with embers: Forecasting civil unrest using open source indicators in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. (ACM), pp. 1799–1808.
- 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.
- James G, Witten D, Hastie T, Tibshirani R (2013) *An introduction to statistical learning*. (Springer) Vol. 112.

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
Logisitic Regression	0.7513	0.5102	0.6018	0.9322