GI_Forum 2018, Issue 1
Page: 47 - 64
Full Paper
Corresponding Author:
bernd.resch@sbg.ac.at
DOI: 10.1553/giscience2018_01_s47

Public Perception of Climate Change in Alaska: A Case Study of Opinion-Mining using Twitter

Helena Bergstedt[1], Alina Ristea[1], Bernd Resch[1] and Annett Bartsch[2,3]

[1]University of Salzburg, Austria
[2]b.geos, Korneuburg, Austria
[3]Austrian Polar Research Institute, Vienna, Austria

Abstract

The Arctic, and with it the State of Alaska, USA, is an area highly impacted by climate change. Changing environmental conditions have started to impact local communities, causing a need for changes ranging from new infrastructure to the relocation of entire towns. These changes connected to rising temperatures have been shown to affect people's overall health, and their mental health in particular. Previous studies using opinion-mining and Twitter data have focused on large areas, not distinguishing between regions within countries. In the course of the research presented in this paper, we analysed Twitter data for the period 2013–2017, from which we extracted opinions concerning climate change topics by applying sentiment analysis (polarity and feelings) and climate change dictionaries, on a 10 x 10 km grid for the State of Alaska, USA. The number of climate change-relevant tweets was found to be much lower than reported in previous studies, where the USA was only considered in its entirety. After applying a topic-modelling approach, we found little difference between the spatial distributions of hotspots for the different climate change topics. A comparison with population data showed considerable biases towards English-speaking communities, tweets in indigenous languages being excluded when pre-defined dictionaries in English were used.

Keywords:

Social Media, Climate Change, Alaska, opinion mining

1 Introduction

Climate change has been shown to affect the Arctic disproportionately highly compared to lower latitudes (Intergovernmental Panel on Climate Change, 2007), a phenomenon known as Arctic amplification (Serreze & Barry, 2011). This is visible through changes in a multitude of phenomena (Callaghan et al., 2010). Studies have found changes in precipitation and local surface evaporation (Bintanja & Selten, 2014), and yearly snow cover (Derksen & Brown, 2012). Permafrost covers large parts of the Arctic and is known to thaw under the influence of rising temperatures (Osterkamp & Romanovsky, 1999). This causes changes in ground stability (Rowland et al., 2010), hydrological conditions (Lawrence, Slater, & Swenson, 2012;

Woo, 1986), ecology (Jorgenson, Racine, Walters, & Osterkamp, 2001), and coastal erosion (Lantuit & Pollard, 2008), all of which influence the daily life of local communities (Bartsch & Meyer, 2016). This has led to new transportation and infrastructure needs caused by changing ground and ice conditions (Larsen et al., 2008; Liljedahl et al., 2016), altering the way people are able to traverse the landscape (Doré, Niu, & Brooks, 2016). Natural hazards are generally thought to become more frequent (Nelson, Anisimov, & Shiklomanov, 2002). In Alaska, the impact of climate change is already highly visible (Jorgenson, Shur, & Pullman, 2006; Raynolds et al., 2014) and has begun to impact public and private infrastructure and the daily lives of citizens (Hovelsrud, Poppel, van Oort, & Reist, 2011; Melvin et al., 2017). Recently, the impact of climate change on the health, and specifically mental health, of local communities has come under discussion (Cunsolo Willox et al., 2013; Cunsolo Willox et al., 2014). This is particularly important for people living in regions highly susceptible to climate change. Cunsolo Willox et al. (2014) argue that people living in the Circumpolar North are especially vulnerable to such developments as they rely on the environment for their livelihoods. In these regions, climate change has implications for resource availability and food security, acute or extreme weather events, landscape degradation, impacts on infrastructure, as well as cultural practices, especially those of indigenous communities (Cunsolo Willox et al., 2014). Burkett, Verchick and Flores (2017) have found that communities of Alaskan natives and Native Americans in Alaska are already experiencing climate change-induced relocation. In many cases, the need for relocation is caused by rising sea levels or coastal erosion threatening communities and infrastructure (Burkett, Verchick & Flores, 2017). All these factors are strongly connected to general health, and in particular to mental well-being (Friel & Ford, 2015; World Health Organization, 2011).

Analysis of Twitter and other social media data has been used to study a great variety of topics, such as public health and healthcare (Culotta, 2010), crime (Kounadi, Ristea, Leitner, & Langford, 2017; Lau, Xia, & Ye, 2014), politics (Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012), and crisis management (Cameron, Power, Robinson, & Yin, 2012). Algorithms using opinion-mining and topic-modelling have frequently been applied to social media data and have been further developed in recent years. Sentiment analysis has been applied at many levels of spatial and temporal granularity (e.g. hourly or daily), mainly by using Twitter data for understanding users' opinions. Various lexicons and dictionaries are used in the analysis of polarity (positive, negative and neutral) and feelings (such as sadness, joy), and some researchers are working on the differences between subjective and objective opinions (Pak & Paroubek, 2010), and between using unigrams and bigrams (Go, Bhayani, & Huang, 2009) – i.e., in text-mining and natural language processing, single words and 2-word sequences. From a spatial point of view, it is important to find clusters of different polarities on smaller scales in order to understand local rather than general behaviour. Several studies using data obtained from Twitter have focused on public opinions concerning climate change (An et al., 2014; Cody, Reagan, Mitchell, Dodds, & Danforth, 2015; Jang & Hart, 2015). However, these studies have generally included data for larger regions, several countries, or in many cases the whole world. For example, Pathak, Henry, and Volkova (2017) consider the whole world; nor do they distinguish between regions within countries. In their study, Alaska is considered as part of the United States as a whole; differences arising from vastly different populations and environmental conditions are not taken into account. To the best of our knowledge, no studies focusing on the analysis of social media in relation to the public perception of climate change and its impacts on a local or regional scale exist for Arctic regions, including Alaska.

Taking into consideration Arctic amplification and the direct implications of climate change for both Alaskan landscapes as well as local communities (Hovelsrud et al., 2011), our study focuses on the Alaskan mainland as a study site. We investigate the usability of Twitter data to study public perceptions of climate change and its implications on daily life in vulnerable communities through addressing the following research questions:

RQ 1: Are there differences in the temporal and spatial distributions between positive and negative tweets in Alaskan territory?
RQ 2: How are climate change-relevant Twitter messages distributed at a fine spatial scale in a sparsely populated area?
RQ 3: Which sentiments (polarity and emotions) are most predominant when talking about climate in Alaska?

The study is designed to explore a well-known and widely discussed topic (climate change) using an established method (sentiment analysis of tweets) on a regional level. Our analysis aims to show the limitations of these methods when applied to regional analysis. Our objective is to explore spatial patterns in the results, but also the advantages of using Twitter data to analyse public perception of climate change among people who are affected on a daily basis by its impacts on their environment as well as by related policy decisions.

2 Data and Methods

Study area

The study area is defined by the Alaska, USA administrative border, with a southern limit of 60°N and a western limit at 170°W (Figure 1). In the following analysis, we adopt a 10 x 10 km grid) according to a method proposed by Eck, Chainey, Cameron, and Wilson (2005). Twitter data represented by points of the geotagged locations and population data were aggregated to the grid, which covers the entire study area, resulting in 13,425 grid cells. To identify clusters and outliers for the spatial distribution of tweets and population, we use this grid projected in the NAD83/Alaska Albers coordinate system.

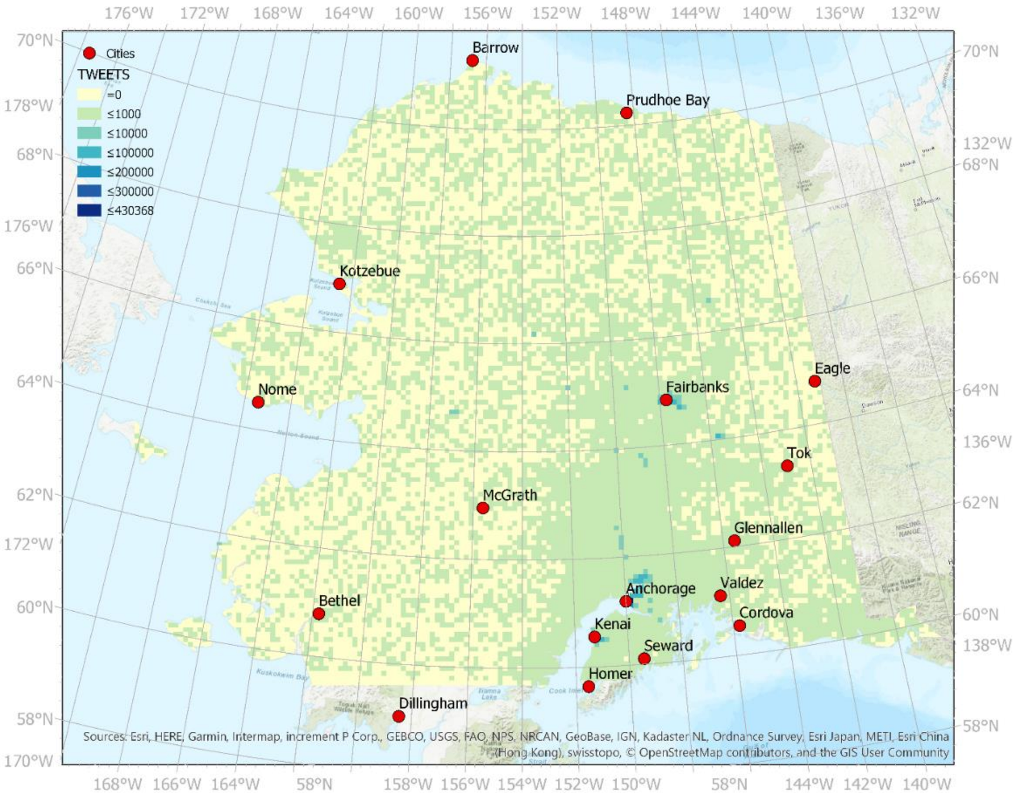

Figure 1: Study Area as defined by the Alaskan state border and additional limits of 60°N and 170°W (WGS1984). Colours represent the number of tweets used in the 10 x 10 km grid. Locations of towns and cities as provided by Alaska State Department of Natural Resources (http://www.asgdc.state.ak.us/)

Data

Twitter data were obtained through the Twitter Streaming Application Programming Interface for the period 2013–2017 (with gaps in August–September 2013, November 2013–March 2014, and the first ten days of July 2014). Only geotagged tweets were considered, in order to allow for geospatial analysis. Twitter data are frequently used in research. However, practical concerns may arise because geotagged tweets represent approximately just 1% (Li, Goodchild, & Xu, 2013; Morstatter, Pfeffer, Liu, & Carley, 2013) to 10% of all posted tweets (Anselin & Williams, 2015; Zhang, Ni, He, & Gao, 2016). These limitations are discussed in a growing body of literature (Steiger, Westerholt, Resch, & Zipf, 2015; Sui & Goodchild, 2011).

Population data included in the study was acquired from the online database of the U.S. Geological survey (U.S. Geological Survey, 2014), which includes the 2010 Census population information for each town/city. The data from each town/city centroid in Alaska were joined in the spatial unit of analysis (10 x 10 km grid cell).

Methodology

In the first step, the tweets' text was cleaned by pre-processing operations. Among other things, this pre-processing removed stop words (i.e non-discriminative words that can reduce the accuracy of the classifier, such as 'that', 'other'), URLs, prepositions, punctuation symbols and numbers; tokenization (the process of breaking a text up into words and other meaningful elements), and all letters were converted to lowercase, in accordance with Resch et al. (2017). Secondly, after checking the tweeting intensity per username, the most active users were excluded. While it is not the purpose of this study to identify personal and non-personal users, the non-personal accounts with high volumes of tweets would introduce bias into the analysis. The eight most active users were excluded from the analysis, including a recruitment company (AlaskaTourjobs) and users tweeting about earthquake occurrences reported by the USGS (everyEarthquake, QuakesToday).
After the pre-processing steps, we followed the workflow presented in Figure 2. The geotagged tweets included approximately 1.9 million messages. These tweets were divided into nine categories, following the approach of Pathak et al. (2017). In order to analyse the topic of climate change, Pathak et al. categorized the tweets into one dataset about 'climate change' and nine subsets (weather, energy, air issues, security, economy, food, animals, water and climate denial), of which five are included in Global Pulse data (United Nations Global Pulse, 2016). For our case study, we used their main 'Climate Change' category and all of the subsets apart from the last, as our tweet dataset did not contain significant climate denial statements.

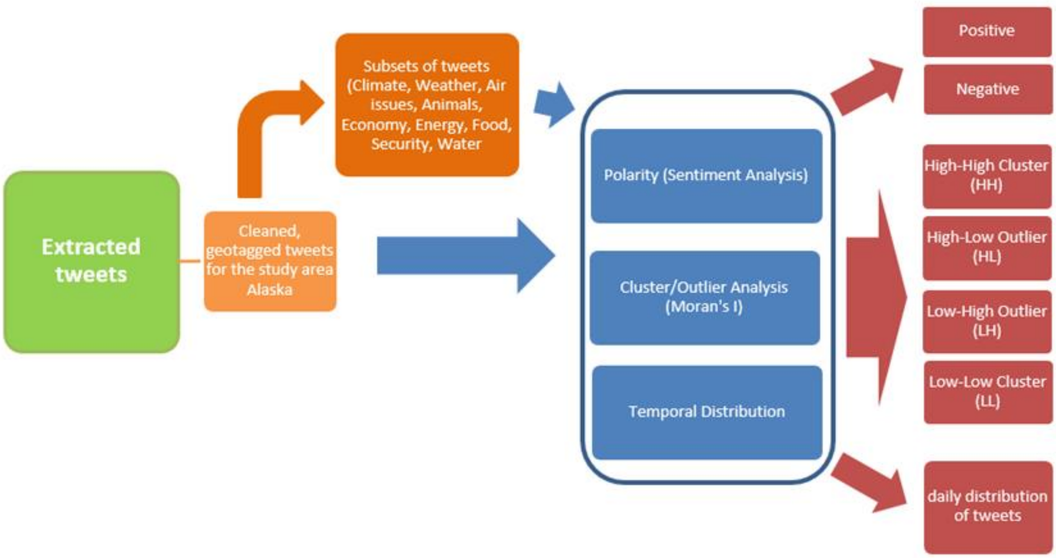

Figure 2: Workflow

Sentiment analysis

For the sentiment analysis of the geotagged tweets, we selected the lexicon created by Hu and Liu (2004), which includes around 6,800 words. This approach uses negative and positive dictionaries in order to calculate the difference between positive and negative polarity for each tweet (Hu & Liu, 2004):

$$score = Number\ of\ positive\ words - Number\ of\ negative\ words$$

A score above zero means that the sentence has an overall 'positive sentiment', a score lower than zero that the sentence has an overall 'negative sentiment'. Tweets with a score of zero are considered 'neutral'. However, misclassifications arise when scores are close to zero due to the low polarity. In the following analysis we define 'positive opinion' tweets as ones with a score higher than or equal to 2, and 'negative opinion' tweets as ones with a score lower than or equal to -2. We also introduce the possibly biased values ($\leq 1\ and\ \geq -1$) as 'low positive' and 'low negative' polarity. The sentiment polarity was analysed for all nine subsets.

In addition, we followed a second approach for sentiment analysis. The National Research Council Canada (NRC) Emotion Lexicon (Emolex) (Mohammad & Turney, 2010, 2013) includes 14,182 unigrams (words) and ~25,000 senses (meanings of the words, e.g. different senses of the word 'play', based on the context in which it is used). This method identifies emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, trust) and polarity (negative and positive) for each tweet, and defines a scale of association between emotions-sentiments and the tweet text (not associated, weakly, moderately or strongly associated). In order to compare the Hu and Liu lexicon with the NRC for polarity variables, we created two categories: one including 'positive' and 'negative' opinions as the ones moderately and strongly associated, the other including 'low positive' and 'low negative' opinions. The NRC algorithm was implemented using the Syuzhet package in R (Jockers, 2015).

Spatial analysis

In order to spatially analyse the Twitter data, a cluster/outlier analysis was performed using the Anselin Local Moran's I statistic (Anselin, 1995; Mitchel et al., 2005). This test identifies concentrations of high and low values and spatial outliers, or clusters of features with values similar in magnitude. The analysis was done using the 'contiguity edges corners' scheme for conceptualizing the spatial relationships, involving 499 permutations, in ArcMap. The model was run for all geolocated tweets (including all subsets) aggregated to the 10 x 10 km grid cells projected in the NAD83 / Alaska Albers coordinate system. The Cluster/Outlier Analysis identified four elements with a 95% significance. Two statistically significant cluster types (of high (HH) and low (LL) values) were identified, as well as two outlier types (one in which a high value is surrounded primarily by low values (HL), and the other in which a low value is surrounded primarily by high values (LH)). The rest of the cells were classified as not significant.

Temporal analysis

For the temporal analysis, the geotagged tweets and the nine climate categories were aggregated into daily bins. The first category included a temporal distribution using daily count values; for the subcategories, a percentage of the total number of geotagged tweets per day was calculated (for example, climate tweets represented 1% of 363 geotagged tweets recorded on 31 August 2015). Additionally, for the geotagged tweets, the positive and negative tweets were analysed considering their time stamp, calculating their percentage of the daily total.

3 Results and Discussion

In this section, we present the following categories of results: opinion-mining through sentiment analysis, significant distribution of tweets, positive/negative tweets, climate tweet clusters and outliers, and temporal analysis of tweet density for all geotagged messages and for the climate subsets. Table 1 shows a summary of statistics for the datasets used. The number of tweets varies between the different subsets; the topics 'Economy' and 'Food' are the most frequent ones, while the topic 'Air Issues' has the lowest number of tweets. Using the lexicon of Hu and Liu (2004), 4.63% of geotagged tweets (~1.9 million messages) were identified as expressing a positive opinion, and 3.89% a negative opinion; 18.87% are characterized by a low positive and 14.57% a low negative polarity.

Table 1: Descriptive statistics for all 13,425 grid cells. Minimum and maximum numbers of tweets per grid cell, the sum of all tweets belonging to one category over all grid cells, mean number of tweets per grid cell, and standard deviation

Variable	Min.	Max.	Sum	Mean	Std. Dev.	Unique users
Tweets	0	430,368	1,893,142	141.016	4835.113	32,242
Positive	0	21,245	87,632	6.527	233.1528	9,927
Negative	0	18,073	73,572	5.480	198.8705	5,893
Low positive	0	86,471	357,189	26.606	972.656	16,987
Low negative	0	65,290	275,833	20,546	731,4486	11,351
Population	0	291,826	461,202	34.354	2541.726	-
Subcategories						
Climate	0	43	148	0.011	0.455	100
Weather	0	22	129	0.009	0.312	101
Air issues	0	7	33	0.002	0.088	31
Animals	0	82	364	0.027	0.931	294
Economy	0	1476	5,805	0.432	16.299	1,784
Energy	0	187	780	0.058	2.064	521
Food	0	1440	5,795	0.432	15.797	2,092
Security	0	266	556	0.041	2.524	246
Water	0	781	4,030	0.300	8.953	1,925

Figure 3 shows the daily number of geotagged tweets for two time periods: 1 January 2013 – 31 April 2015 and 1 May 2015 – 31 December 2017 (including the gaps explained in subsection "Data"). The y-axis shows differing values between the two periods, the number of tweets in the second one being noticeably lower. In April 2015, Twitter changed its policy and the way of geotagging, resulting in a lower proportion of tweets containing coordinates after that date. After April 2015, the positive tweets represent a clearly higher percentage of the total geotagged tweets than the negative ones, while before that date the two polarities were only slightly different (see Figure 4). This suggests that the change in Twitter's policy caused users to geotag disproportionally fewer negative tweets. Figure 5 shows the daily temporal distribution of tweet counts regarding the nine climate subsets. After 2016, tweets about water and food represent an increased percentage of the geotagged tweets for those specific days (e.g. 2.5% of the tweets posted on 17 October 2017 are from the food category). Before the Twitter policy change in 2015, the categories had similar daily distributions (see Figure 5).

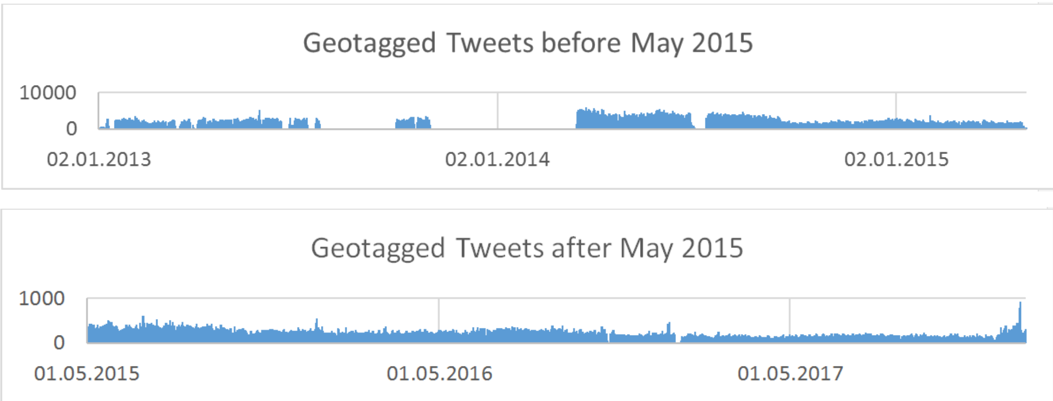

Figure 3: Geotagged tweets before 1 May 2015 (after Twitter policy changed) and after 1 May 2015. The y-axes have different ranges to account for the difference in number of tweets for the two time periods

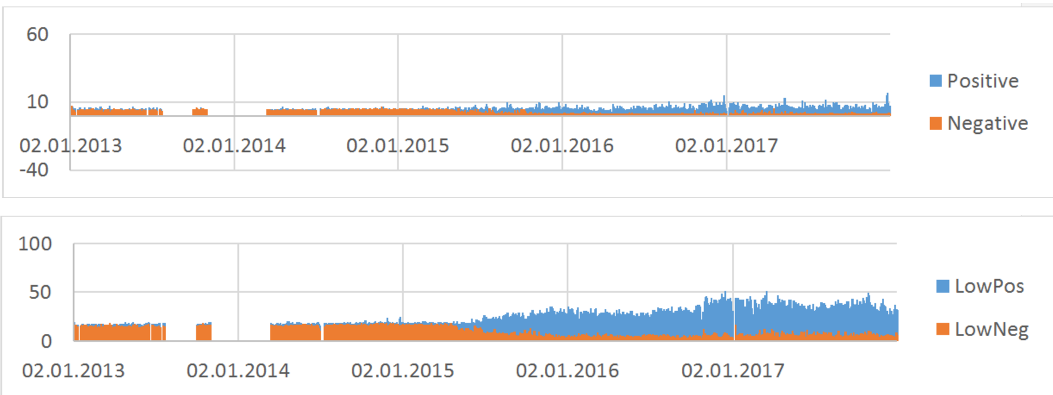

Figure 4: Percentage of Positive/Low Positive (blue) and Negative/Low Negative (orange) tweets per day, as obtained using Hu and Liu's (2004) lexicon

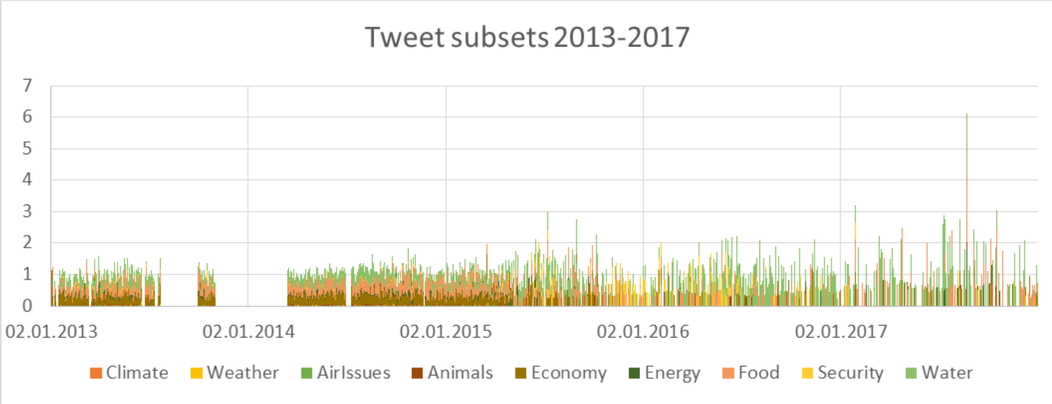

Figure 5: Percentage of tweet subsets per day obtained using NRC Lexicon for emotions (Mohammad & Turney, 2010, 2013)

Figure 6 shows density and cluster/outlier maps for all geotagged tweets on the 10 x 10 km grid and population data from the Census (U.S. Geological Survey, 2014). As expected, areas around the main towns/cities have a higher tweet density. However, in the western part, where LL clusters and LH outliers for the population are located, the Twitter data density is low and not representative of the patterns observed in the population data. According to the United States Census Bureau, these areas (especially on the coast) have a large indigenous population, possibly speaking languages other than English (Norris, Vines, & Hoeffel, 2012). As our analysis considers only English-language tweets, we introduce bias, possibly excluding native communities.

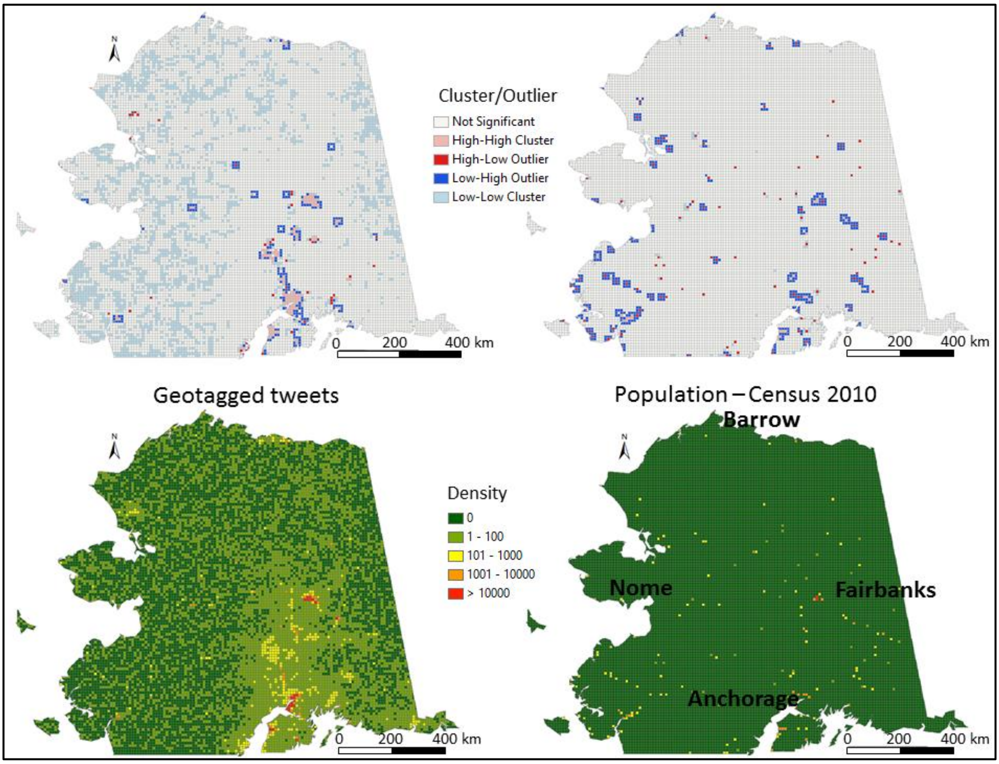

Figure 6: Density (bottom) and Cluster/Outlier Mapping (top) for all geotagged tweets (left) and population (right)

Figure 7 depicts the positivism and negativism from the geotagged tweets. From a general visualization, the density and spatial clustering are similar. In the southeast area (around Glenallen and McCarthy – red circle in Figure 7), high-low outliers for the positive tweets could be found, while no outliers concerning negative tweets were identified in this area. Moreover, more negative tweets were found in the northern area, around Wainwright and Atqasuk (blue circle in Figure 7).

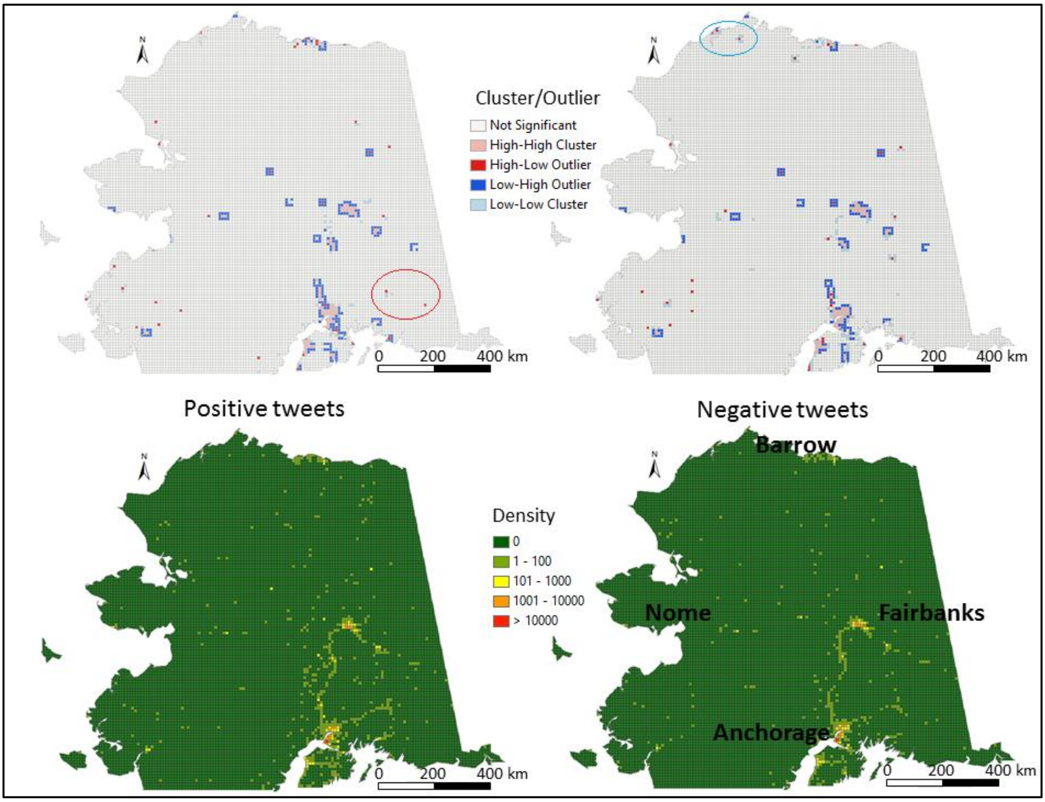

Figure 7: Density (bottom) and Cluster/Outlier mapping (top) for all geotagged tweets, positive and negative. Red and blue circles highlight areas differing between positive and negative tweet clusters/outliers (Red: cities of Glenallen and McCarthy; blue: cities of Wainwright and Atqasuk)

The nine climate categories were spatially analysed for clusters/outliers (Figure 8). LL clusters are more common when the dataset is fairly small (for the sizes of the different groups, see Table 1). Figure 8 displays the areas where people are more concerned about certain topics (or, rather, where they are using words from the above-mentioned dictionaries). It shows many similarities between the nine categories. For the topics 'Economy', 'Food' and 'Water', where the data volume is larger, we notice many non-significant areas. The city of Galena in the west is defined as an HH cluster, except for 'Air Issues' and 'Security'. Nome, located on the west coast, is an HH cluster except for 'Animals' and 'Weather' (see Figure 8). The cities with more inhabitants, such as Anchorage, Palmer and Kenay (all in the south) and Fairbanks (in the centre), show tweets from all subsets. This was expected considering these areas have larger populations.

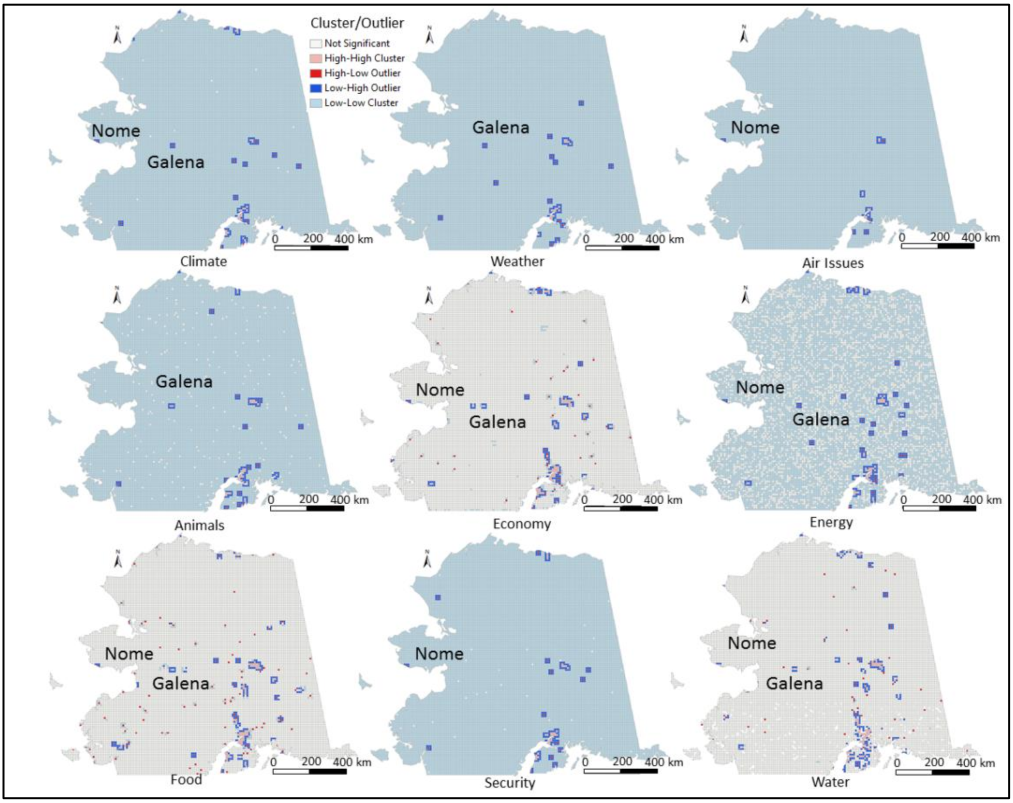

Figure 8: Cluster/Outlier mapping for the nine climate categories 'Climate', 'Weather', 'Air Issues', 'Animals', 'Economy', 'Energy', 'Food', 'Security' and 'Water'. Nome and Galena show notable differences between the results for the subsetted categories

For the nine climate categories, we adopted another sentiment analysis algorithm, from the NRC lexicon (see section "Temporal analysis" for further description). Figure 9 compares the two methods for defining positive and negative tweets, as well as ~~for~~ low positive and low negative tweets. The NRC method shows higher percentages in all cases compared with the first method. 'Security', 'food' and 'economy' include a high percentage of positive tweets. For 'Air Issues', 'Economy' and 'Energy', the first lexicon identifies more negative tweets than positive ones.

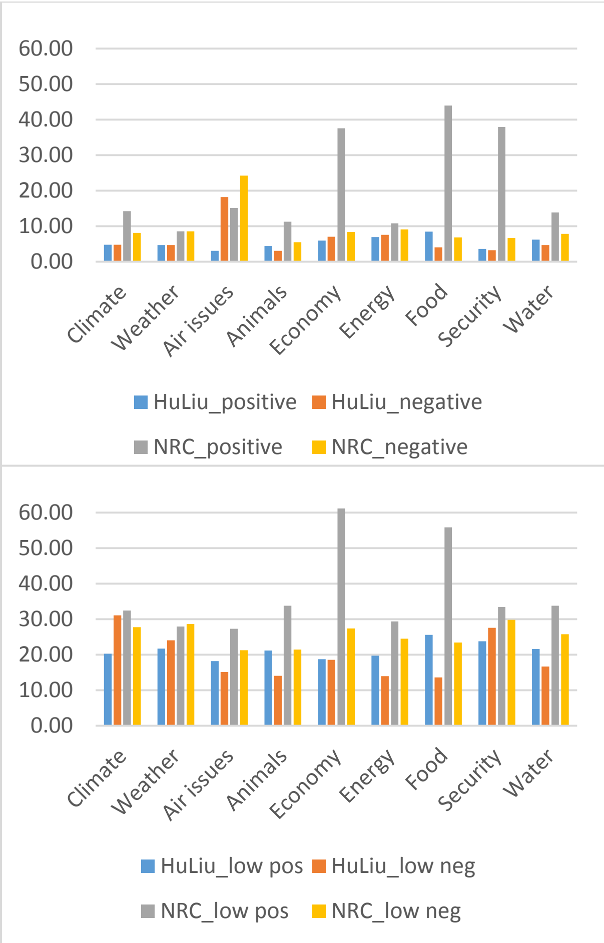

Figure 9: Percentage of Positive and Negative (left) and low positive and low negative (right) tweets per category using HuLiu (Hu & Liu, 2004) and NRC methods (Mohammad & Turney, 2010, 2013)

The NRC method also defines emotions for each tweet (Figure 10). Climate and weather topics elicit more fear, while food has a higher percentage of joy and a low percentage of fear. The link between fear and the topics of climate and weather agrees with the overall descriptions of the effect climate change can have on people's daily lives (Cunsolo Willox et al., 2013). Figure 10 also shows the most frequently used words in each subset. While for some categories the connection to climate change is clear (words like 'climate', 'change', 'global warming' and 'heatwave' being used frequently), other categories, like 'Food', are more difficult to attribute to the topic of climate change. The subset 'Weather' shows frequent use of words such as 'flood' and 'heatwave', both of which are phenomena commonly attributed to a changing climate in the Arctic (Frich et al., 2002; Radosavljevic et al., 2016). These results show that it is difficult to extract text referring to a specific subject just by using a defined dictionary; for example, people who talk about food in social media show a high probability of not discussing food in relation to global warming or climate, which is why we notice the emotion of joy being highlighted here.

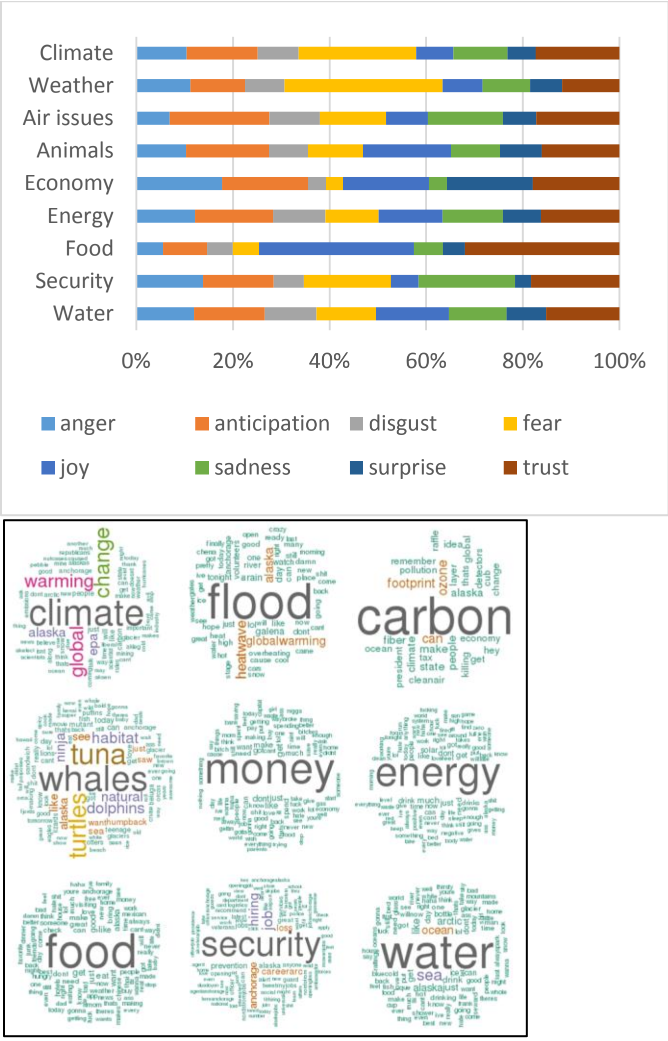

Figure 10: Sentiment distribution using NRC method over the nine climate categories (left); word-cloud visualizations of the most common words per category (right)

4 Conclusion

In this study, we presented the use of tweets to identify opinions about climate change on a 10 x 10 km grid. In contrast to other studies in this field, which commonly cover larger areas, we focused on just one region which is highly impacted by climate change, and performed our analysis on a fine spatial scale.

This case study shows the usability of Twitter data in sparsely populated areas for studying people's general opinions and emotions about climate topics. Spatial analysis was used to determine areas of positive and negative tendencies in messaging, and others where climate-related information is discussed. Areas showing clusters and outliers differed only slightly between the subcategories studied.

Compared with the total number of geotagged tweets, results show a moderate volume of tweets that used words included in the climate change dictionary. This regional study adds to other global studies, where Alaska is commonly included in the United States boundaries. Despite the plausible results, our study shows several limitations. First, Twitter data is restricted to a low volume of geotagged tweets, which might cause biases in opinion-mining, including social and demographic biases. In addition, this analysis focused on messages written in English, missing out important groups such as indigenous people speaking other languages. This bias needs to be addressed when applying Twitter-based methods to these areas and calls for the creation and use of multi-language dictionaries, including indigenous languages. The changes in Twitter policy from April 2015 may also have played a role in restricting the amount of information with spatial attributes that was collected.

References

An, X., Ganguly, A. R., Fang, Y., Scyphers, S. B., Hunter, A. M., & Dy, J. G. (2014). Tracking climate change opinions from twitter data. In *Workshop on Data Science for Social Good.*

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115.

Anselin, L., & Williams, S. (2015). Digital neighborhoods. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *9*(4), 305–328. https://doi.org/10.1080/17549175.2015.1080752

Bartsch, A. & Mayer, A~~lexandra~~. (2016). Klimawandel in der Arktis - Perspektiven aus den Natur- und Sozialwissenschaften. 166-182.

Bintanja, R., & Selten, F. M. (2014). Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat. *Nature*, *509*(7501), 479–482. https://doi.org/10.1038/nature13259

Burkett, M., Verchick, R.R.M., Flores, D. (2017). Reaching Higher Ground: Avenues to Secure and Manage New Land for Communities Displaced by Climate Change. Center for Progressive Reform (May 2017); Loyola University New Orleans College of Law Research Paper No. 2017-07.

Callaghan, T. V., Bergholm, F., Christensen, T. R., Jonasson, C., Kokfelt, U., & Johansson, M. (2010). A new climate era in the sub-Arctic: Accelerating climate changes and multiple impacts. *Geophysical Research Letters*, *37*(14), n/a-n/a. https://doi.org/10.1029/2009GL042064

Cameron, M. A., Power, R., Robinson, B., & Yin, J. (2012). Emergency Situation Awareness from Twitter for Crisis Management. In : *WWW '12 Companion, Proceedings of the 21st International Conference on World Wide Web* (pp. 695–698). New York, NY, USA: ACM. https://doi.org/10.1145/2187980.2188183

Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2015). Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll. *PloS one*, *10*(8), e0136092. https://doi.org/10.1371/journal.pone.0136092

Culotta, A. (2010). Towards detecting influenza epidemics by analyzing Twitter messages. In P. Melville, J. Leskovec, & F. Provost (Eds.), *Towards detecting influenza epidemics by analyzing Twitter messages* (pp. 115–122). New York, New York, USA: ACM Press. https://doi.org/10.1145/1964858.1964874

Cunsolo Willox, A., Harper, S. L., Ford, J. D., Edge, V. L., Landman, K., Houle, K.,. . . Wolfrey, C. (2013). Climate change and mental health: An exploratory case study from Rigolet, Nunatsiavut, Canada. *Climatic Change*, *121*(2), 255–270. https://doi.org/10.1007/s10584-013-0875-4

Cunsolo Willox, A., Stephenson, E., Allen, J., Bourque, F., Drossos, A., Elgarøy, S.,. . . Wexler, L. (2014). Examining relationships between climate change and mental health in the Circumpolar North. *Regional Environmental Change*, *15*(1), 169–182. https://doi.org/10.1007/s10113-014-0630-z

Derksen, C., & Brown, R. (2012). Spring snow cover extent reductions in the 2008-2012 period exceeding climate model projections. *Geophysical Research Letters*, *39*(19), n/a-n/a. https://doi.org/10.1029/2012GL053387

Doré, G., Niu, F., & Brooks, H. (2016). Adaptation Methods for Transportation Infrastructure Built on Degrading Permafrost. *Permafrost and Periglacial Processes*, *27*(4), 352–364. https://doi.org/10.1002/ppp.1919

Eck, J., Chainey, S., Cameron, J., & Wilson, R. (2005). Mapping crime: Understanding hotspots.

Frich, P., Alexander, L. V., Della-Marta, P. M., Gleason, B., Haylock, M., Tank, A. K.G., & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate research*, *19*(3), 193–212.

Friel, S., & Ford, L. (2015). Systems, food security and human health. *Food Security*, *7*(2), 437–451. https://doi.org/10.1007/s12571-015-0433-1

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, *1*(12).

Hovelsrud, G. K., Poppel, B., van Oort, B., & Reist, J. D. (2011). Arctic Societies, Cultures, and Peoples in a Changing Cryosphere. *Ambio*, *40*(S1), 100–110. https://doi.org/10.1007/s13280-011-0219-4

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168–177).

Intergovernmental Panel on Climate Change. (2007). Climate change 2007: The physical science basis. *Agenda*, *6*(07), 333.

Jang, S. M., & Hart, P. S. (2015). Polarized frames on ‘climate change’ and ‘global warming’ across countries and states: Evidence from Twitter big data. *Global Environmental Change*, *32*, 11–17. https://doi.org/10.1016/j.gloenvcha.2015.02.010

Jockers, M. L. (2015). *Syuzhet: Extract Sentiment and Plot Arcs from Text*. Retrieved from https://github.com/mjockers/syuzhet

Jorgenson, M. T., Racine, C. H., Walters, J. C., & Osterkamp, T. E. (2001). Permafrost Degradation and Ecological Changes Associated with a WarmingClimate in Central Alaska. *Climatic Change*, *48*(4), 551–579. https://doi.org/10.1023/A:1005667424292

Jorgenson, M. T., Shur, Y. L., & Pullman, E. R. (2006). Abrupt increase in permafrost degradation in Arctic Alaska. *Geophysical Research Letters*, *33*(2), L02503. https://doi.org/10.1029/2005GL024960

Kounadi, O., Ristea, A., Leitner, M., & Langford, C. (2017). Population at risk: Using areal interpolation and Twitter messages to create population models for burglaries and robberies. *Cartography and Geographic Information Science*, *55*(1), 1–16. https://doi.org/10.1080/15230406.2017.1304243

Lantuit, H., & Pollard, W. H. (2008). Fifty years of coastal erosion and retrogressive thaw slump activity on Herschel Island, southern Beaufort Sea, Yukon Territory, Canada. *Geomorphology*, *95*(1-2), 84–102. https://doi.org/10.1016/j.geomorph.2006.07.040

Larsen, P., Goldsmith, S., Smith, O., Wilson, M., Strzepek, K., Chinowsky, P., & Saylor, B. (2008). Estimating future costs for Alaska public infrastructure at risk from climate change. *Global Environmental Change*, *18*(3), 442–457. https://doi.org/10.1016/j.gloenvcha.2008.03.005

Lau, R. Y.K., Xia, Y., & Ye, Y. (2014). A Probabilistic Generative Model for Mining Cybercriminal Networks from Online Social Media. *IEEE Computational Intelligence Magazine*, *9*(1), 31–43. https://doi.org/10.1109/MCI.2013.2291689

Lawrence, D. M., Slater, A. G., & Swenson, S. C. (2012). Simulation of present-day and future permafrost and seasonally frozen ground conditions in CCSM4. *Journal of Climate*, *25*(7), 2207–2225. https://doi.org/10.1175/JCLI-D-11-00334.1

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, *40*(2), 61–77. https://doi.org/10.1080/15230406.2013.777139

Liljedahl, A. K., Boike, J., Daanen, R. P., Fedorov, A. N., Frost, G. V., Grosse, G.,. . . Zona, D. (2016). Pan-Arctic ice-wedge degradation in warming permafrost and its influence on tundra hydrology. *Nature Geoscience*, *9*(4), 312–318. https://doi.org/10.1038/ngeo2674

Maxine Burkett, Robert R.M. Verchick, David Flores. (2017). *Reaching Higher Ground: Avenues to Secure and Manage New Land for Communities Displaced by Climate Change*. Retrieved from Center for progressive Reform website: http://progressivereform.org/articles/ReachingHigherGround_1703.pdf

Melvin, A. M., Larsen, P., Boehlert, B., Neumann, J. E., Chinowsky, P., Espinet, X.,. . . Marchenko, S. S. (2017). Climate change damages to Alaska public infrastructure and the economics of proactive adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(2), E122-E131. https://doi.org/10.1073/pnas.1611056113

Mitchel, A., & others. (2005). The ESRI Guide to GIS analysis, volume 2: Spartial measurements and statistics. *ESRI Guide to GIS analysis*.

Mohammad, S. M., & Turney, P. D. (2010). Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. In : *CAAGET '10, Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* (pp. 26–34). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1860631.1860635

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word‐emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. In *ICWSM*.

Nelson, F. E., Anisimov, O. A., & Shiklomanov, N. I. (2002). Climate Change and Hazard Zonation in the Circum-Arctic Permafrost Regions. *Natural Hazards*, *26*(3), 203–225. https://doi.org/10.1023/A:1015612918401

Norris, T., Vines, P. L., & Hoeffel, E. M. (2012). *The American Indian and Alaska Native Population: 2010*: US Department of Commerce, Economics and Statistics Administration, US Census Bureau Washington, DC.

Osterkamp, T. E., & Romanovsky, V. E. (1999). Evidence for warming and thawing of discontinuous permafrost in Alaska. *Permafrost and Periglacial Processes*, *10*(1), 17–37. https://doi.org/10.1002/(SICI)1099-1530(199901/03)10:1<17::AID-PPP303>3.0.CO;2-4

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10).

Pathak, N., Henry, M., & Volkova, S. (2017). Understanding Social Media ' s Take on Climate Change through Large-Scale Analysis of Targeted Opinions and Emotions. Retrieved from https://www.cs.jhu.edu/ svitlana/papers/PHV_17.pdf

Radosavljevic, B., Lantuit, H., Pollard, W., Overduin, P., Couture, N., Sachs, T.,. . . Fritz, M. (2016). Erosion and Flooding—Threats to Coastal Infrastructure in the Arctic: A Case Study from Herschel Island, Yukon Territory, Canada. *Estuaries and Coasts*, *39*(4), 900–915. https://doi.org/10.1007/s12237-015-0046-0

Raynolds, M. K., Walker, D. A., Ambrosius, K. J., Brown, J., Everett, K. R., Kanevskiy, M.,. . . Webber, P. J. (2014). Cumulative geoecological effects of 62 years of infrastructure and climate change in ice-rich permafrost landscapes, Prudhoe Bay Oilfield, Alaska. *Global Change Biology*, *20*(4), 1211–1224. https://doi.org/10.1111/gcb.12500

Resch, B., Usländer, F. and Havas, C. (2017) Combining Machine-learning Topic Models and Spatio-temporal Analysis of Social Media Data for Disaster Footprint and Damage Assessment. Cartography and Geographic Information Science (CaGIS), DOI: 10.1080/15230406.2017.1356242.

Rowland, J. C., Jones, C. E., Altmann, G., Bryan, R., Crosby, B. T., Hinzman, L. D., . . Geernaert, G. L. (2010). Arctic Landscapes in Transition: Responses to Thawing Permafrost. *Eos, Transactions American Geophysical Union*, *91*(26), 229–230. https://doi.org/10.1029/2010EO260001

Serreze, M. C., & Barry, R. G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, *77*(1-2), 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004

Steiger, E., Westerholt, R., Resch, B., & Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Computers, Environment and Urban Systems*, *54*, 255–265. https://doi.org/10.1016/j.compenvurbsys.2015.09.007

Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, *25*(11), 1737–1748. https://doi.org/10.1080/13658816.2011.604636

U.S. Geological Survey. (2014). USGS Small-scale Dataset - Global Map: Cities and Towns of the United States 201403 Shapefile.

United Nations Global Pulse. (2016). Global Pulse.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle. In : *ACL '12, Proceedings of the ACL 2012 System Demonstrations* (pp. 115–120). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2390470.2390490

Woo, M.-k. (1986). Permafrost hydrology in North America 1. *Atmosphere-Ocean*, *24*(3), 201–234. https://doi.org/10.1080/07055900.1986.9649248

World Health Organization (Ed.). (2011). *Impact of economic crises on mental health*. Geneva: World Health Organization.

Zhang, Z., Ni, M., He, Q., & Gao, J. (2016). *Mining Transportation Information from Social Media for Planned and Unplanned Events*. Retrieved from https://www.buffalo.edu/content/www/transinfo/Research/transportation-operations/social-media-mining-for-events/_jcr_content/par/download/file.res/MiningSocialMediaEvents_FinalReport.pdf