Global Environmental Change xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Global Environmental Change

journal homepage: www.elsevier.com/locate/gloenvcha

Public microblogging on climate change: One year of Twitter worldwide

Andrei P. Kirilenko [a,*], Svetlana O. Stepchenkova [b]

[a] Department of Earth Systems Science and Policy, University of North Dakota, Stop 9011, Grand Forks, ND 58202-9011, United States
[b] The Department of Tourism, Recreation and Sport Management, University of Florida, P.O. Box 118208, Gainesville, FL 32611-8208, United States

ARTICLE INFO

Article history:
Received 23 November 2013
Received in revised form 14 February 2014
Accepted 21 February 2014

Keywords:
Climate change
Communication
Blogging
Twitter
Content analysis

ABSTRACT

Public perceptions of climate change are traditionally measured through surveys. The exploding popularity of social networks, however, presents a new opportunity to research the spatiotemporal pattern of public discourse in relation to natural and/or socio-economic events. Among the social networks, Twitter is one of the largest microblogging services. The architecture of Twitter makes the question "what's happening?" the cornerstone of information exchange. This inspired the notion of using Twitter users as distributed sensors, which has been successfully employed in both the natural and social sciences. In 2012 and 2013, we collected 1.8 million tweets on "climate change" and "global warming" in five major languages (English, German, Russian, Portuguese, and Spanish). We discuss the geography of tweeting, weekly and daily patterns, major news events that affected tweeting on climate change, changes in the central topics of discussion over time, the most authoritative traditional media, blogging, and the most authoritative organizational sources of information on climate change referenced by Twitter users in different countries. We anticipate that social network mining will become a major source of data in the public discourse on climate change.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The term microblogging refers to Internet services that allow the users (bloggers) to post brief online messages ("status updates") that are visible to their social network. "The simplicity of publishing such short updates in various situations...makes microblogging an innovative communication method that can be seen as a hybrid of blogging, instant messaging, social networking and status notifications" (Ross et al., 2011). Currently, microblogging enjoys almost unprecedented popularity; by July 2012, over 0.5 billion accounts were registered with microblogging service Twitter alone (Semiocast, 2012). Discussions on microblogging platforms provide vast amount of data on various topics of social importance, which has helped to establish microblogging as a rich resource for academic research. A recent meta-analysis of 575 peer-reviewed publications that used Twitter microblogging data identified the wide range of science domains to which these studies belong: geography, marketing, natural disaster management, linguistics, politics, and many others (Williams et al., 2013). Notably, research on public discussions of climate change is entirely absent from this list.

Public perceptions of climate change are typically measured using surveys, such as the Climate Change in the American Mind project (Leiserowitz, 2012), but the exploding popularity of social networks presents a new opportunity to study the public discourse on climate change and its spatiotemporal patterns in relation to natural variability in weather patterns or social and economic events. Using the Academic Search Primer and Web of Science, we, however, found few studies that used microblogging as a data source to investigate phenomena only tangentially related to the climate change discourse. For example, Segerberg and Bennett (2011) examined the internal organization of climate change protests related to the 2009 15th United Nations Conference of the Parties on Climate Change (COP-15) in Copenhagen by analyzing the changes in origin and connectivity between Twitter messages related to the event. Satchwell (2012) researched children's understanding of climate change using a variety of methods, including observations of Twitter conversations. Bruns and Burgess (2011) investigated the public discourse during the 2010 Australian elections, including climate change as one of the many topics. Bosch (2012) investigated online coverage of climate change in South Africa by comparing the blogs of the Mail & Guardian newspaper's journalists with Twitter messages.

* Corresponding author. Tel.: +1 701 7776761; fax: +1 701 7772940.
E-mail address: andrei.kirilenko@und.edu (A.P. Kirilenko).

http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008
0959-3780/© 2014 Elsevier Ltd. All rights reserved.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

However, to date, no overarching studies of microblogging on Twitter about climate change have addressed the communication process of "who says what through which channel to whom with what effect" (Lasswell, 1948).

Among the social networks, Twitter is arguably the largest microblogging service: according to Twitter CEO Dick Costolo, Twitter users generate 500 million unique messages (tweets) daily (Costolo, 2012). Originally, Twitter was designed as a platform to help people stay in touch by publishing brief updates on their activities. However, over time, this concept has shifted dramatically. Twitter co-founder Biz Stone (2009) writes: "*The fundamentally open model of Twitter created a new kind of information network and it has long outgrown the concept of personal status updates. Twitter helps you share and discover what's happening now among all the things, people, and events you care about. 'What are you doing?' isn't the right question anymore—starting today, we've shortened it by two characters. Twitter now asks, 'What's happening?'*"

This change in Twitter's model led scholars to regard the interactions among Twitter users as a large, distributed network of sensors that react to external events by sending tweets, making the data especially conducive to studying natural phenomena, such as weather and climate. Studies of this type involve the data mining technologies of Big Data to collect large samples or the entire population of tweets on a topic of interest and analyze millions of Twitter messages (Williams et al., 2013). Fundamentally, however, the value of "people as sensors" information is not so much a product of its size, but rather its networked, interconnected structure (Boyd and Crawford, 2012). Adopting the "distributed network of sensors" concept, a number of scholars used Twitter as a real-time tool for monitoring natural and social phenomena. For example, Sakaki et al. (2013) reported that by monitoring tweets originating in Japan, they were able to detect 93% of earthquakes of seismic intensity 3 or higher and send warnings to subscribers much more rapidly than the Japan Meteorological Agency. In behavioral economics, Bollen et al. (2011) demonstrated that monitoring Twitter activity could predict changes in the Dow Jones Industrial Average. In epidemiology, Achrekar et al. (2011) found that their Twitter-based predictions of the number of flu cases were highly correlated with seasonal influenza-like illness statistics from the US Center for Disease Control. In political science, Wang et al. (2012) conducted an analysis of events leading to changes in the public's political sentiments by classifying Twitter messages related to the 2012 US presidential election. In studying perceptions of climate change, we (Molodtsova et al., 2013) were able to show that the weekly number of tweets on climate change correlates with local weather and extreme events, in line with recent results obtained through traditional surveys (Howe et al., 2012).

The high utility of Twitter messages for research on climate change and climate change communications is connected to the internal architecture of the Twitter message and the fact that they are augmented by ample auxiliary information. Two primary domains in Twitter data are (1) the user domain, which contains various statistics on the registered user, such as geographical location, time zone, account age, "social status" (the number of followers) and (2) the message domain, which encompasses the content and properties of individual messages, including the referrals from and references to messages by other Twitter users, hash tags, and URL references to Internet resources (Cheong and Lee, 2010). This combination of data on the message itself, its author, and its connectivity network makes Twitter data advantageous to a research of climate change discourse, while the locational information that communicate the geographical coordinates of a tweeting person makes Twitter data especially compatible with studying the variations of spatial patterns of climate change messaging e.g. in relation to local weather.

In addition, the condensed structure of Twitter messages makes a perfect fit for communication research following the framework of content analysis, a methodology that "aims at describing, with optimum objectivity, precision, and generality, what is said on a given subject in a given place at a given time" (Lasswell et al., 1952). Content analysis involves the systematic reduction of the content flow, whether textual or otherwise symbolic, and can be subjected to quantitative data analysis techniques. One of the primary components of content analysis is dividing the content into units that are "as large as is meaningful...and as small as is feasible..." (Krippendorff, 2004, p. 102). The meaning of a text, however, may be changed in this process, as the size of context units affects the results of their analysis (Geller et al., 1942; Krippendorff, 2004). Complete, yet brief, Twitter messages have already undergone the process of data reduction, i.e., standardizing the size of the messages, "shedding off" irrelevant content, and incorporating conventions and shortcuts the "meaning" of which is shared by all Twitter users. Further, Twitter users frequently highlight the key words in their messages through the hash tag mechanism, accentuating the primary message of their tweet. All of these features make Twitter messages superior to comparable data from other sources for studying traditional and user-generated media communications about climate change (Bastos and Zago, 2013).

Following Lasswell (1948), this study analyzes and comprehensively describes the current state of the global public discourse on climate change, as reflected in microblogging on Twitter. In 2012 and 2013, we collected 1.8 million tweets that included the words "climate change" and/or "global warming" in five major languages, namely, English, German, Russian, Spanish, and Portuguese. This vast amount of Twitter data was summarized along the following dimensions (1) geography of tweeting; (2) changes in the primary topics of discussion over time; (3) the most authoritative Internet sources of information; (4) the most authoritative traditional media, bloggers, organizations and celebrities tweeting on climate change in different countries; and (5) major news events that affected tweeting on climate change worldwide. Prior to the analysis of the climate change discourse, several methodological issues were resolved, including assigning geographical locations to tweets and correcting for natural temporal cycles in Twitter activity.

2. Methods

The overall methodology used to collect and process the data, described in this section, is depicted in Fig. 1.

2.1. Data collection

Throughout 2012, we systematically searched Twitter for messages related to climate change in four languages: English, German, Russian, and Spanish (henceforth – EGRS), with search frequency corresponding to tweeting intensity. The language selection was made based on population and the number of countries speaking the language, and on the number of tweets in that language. For that purpose, we developed Python code containing pre-set search phrases using Twitter REST API v1.0. In total, 1,853,392 tweets were collected. After the initial data analysis revealed an active and distinct discussion on climate change in Latin America, we collected an additional set of 18,871 messages in Portuguese (henceforth – P) to analyze the discussion of climate change in Brazil; due to differences in data collection periods, we did not use these tweets in the analysis of changes in climate change discussion over time. Table 1 provides details on search words, sample periods, periodicity, and the number of days on which tweets in a given language were collected. Collected

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

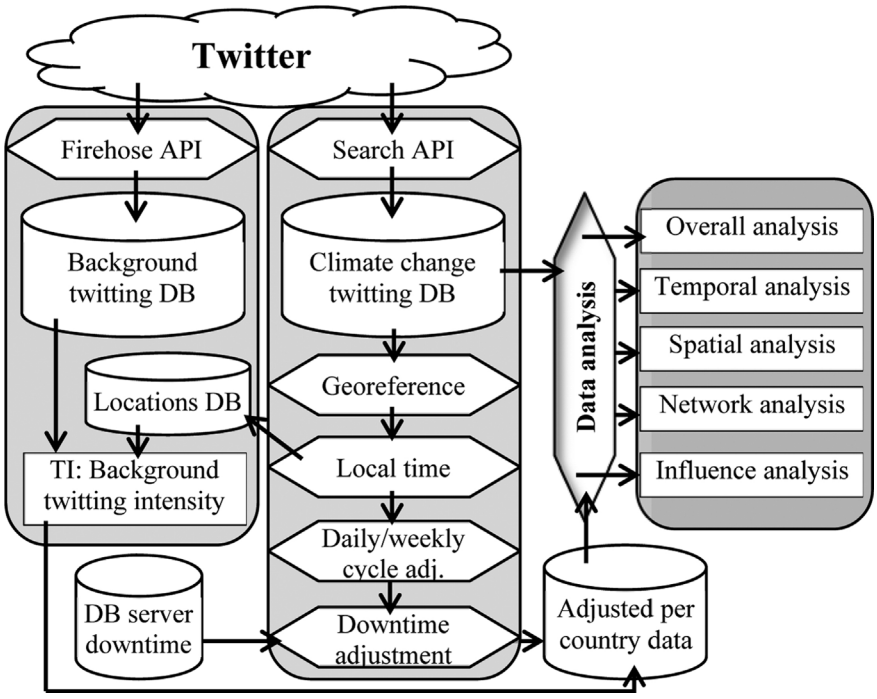

Fig. 1. Data collection and processing.

tweets were filtered for presence of search words, eliminating erroneous tweets, e.g., those selected solely due to the presence of the search words in a URL link.

2.2. Geolocation

Knowing the geographical location of a tweet, e.g., the country of origin, is crucial for studying the geographical pattern of microblogging on a topic of interest. Twitter provides its users with a direct method to explicitly define their location by providing the latitude and longitude of their GPS-enabled device (e.g., mobile phone; henceforth – device coordinates) and/or the neighborhood in which this location is located (henceforth – status coordinates). However, both of these options are rarely employed: estimates of the presence of structural locational information in tweets by various authors tend to represent less than 1% of their tweet databases. For example, Graham et al. (2013) found that of the 19.6 billion tweets they collected, only 0.7% contained geolocational information. In our sample, we found that the shares of tweets with geographical location data in English and German were close to that observed for the entire population of tweets (0.82% and 0.61%, respectively); however, the percentages of such tweets in Spanish and Russian were significantly higher (1.46% and 1.47%). It is important to note that tweets with geographical location data are unlikely to be representative of the entire population, as it is assumed that they reflect the views of the most technologically advanced group (Graham et al., 2013).

As opposed to stating geographical location explicitly, a considerable share of Twitter users provides their location in textual form. We developed Python code, which utilized the online geographical database GeoNames (www.geonames.org), to resolve 105,378 unique location names provided by Twitter users and used tweets with GPS coordinates to validate the algorithm. For methodological details, validation results and a review of geolocation methods, see (Cheng et al., 2010; Eisenstein et al., 2010; Gao and Barbier, 2012; Appendix A.

2.3. Data adjustment

Over the course of 2012, we experienced a few Internet service outages. Despite significant redundancy built into the data collection algorithm to guard against losing data, these outages occasionally led to reduced data volume. The percentage of lost tweets depended on the number of tweets per hour, which in turn strongly depended on the time of the day and the language of the tweets. To correct for the error in tweet volume, we performed the following procedures:

1 The original tweets are GMT time-stamped. We computed local date/day of the week/time d_{loc}, w_{loc}, t_{loc} using the time zones of the estimated geographical location;
2 The days with 24 h of stored data were used to compute relative hourly tweeting intensity $f_h(t_{loc})$ (Fig. 2A) and weekly tweeting intensity $f_w(w_{loc})$ (Fig. 2B) – note the difference with background Twitter activity (Fig. 2C and D, respectively).

Table 1
Twitter search details: search words (without inflected word forms), sample start date, periodicity, and the total number of days with more than one tweet.

Language	Search phrase	Sample period	Sample frequency (min)	Days with >1 tweet
English	climate change, global warming	1/1 to 31/12	10	364
German	klimawandel, globale erwärmung	6/2 to 31/12	30	315
Russian	изменение климата, глобальное потепление	1/1 to 31/12	120	366
Spanish	cambio climático, calentamiento global	10/10 to 31/12	30	81
Portuguese	mudança clima, alteração climática, aquecimento climática	5/6 to 7/11 (2013)	20	66

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

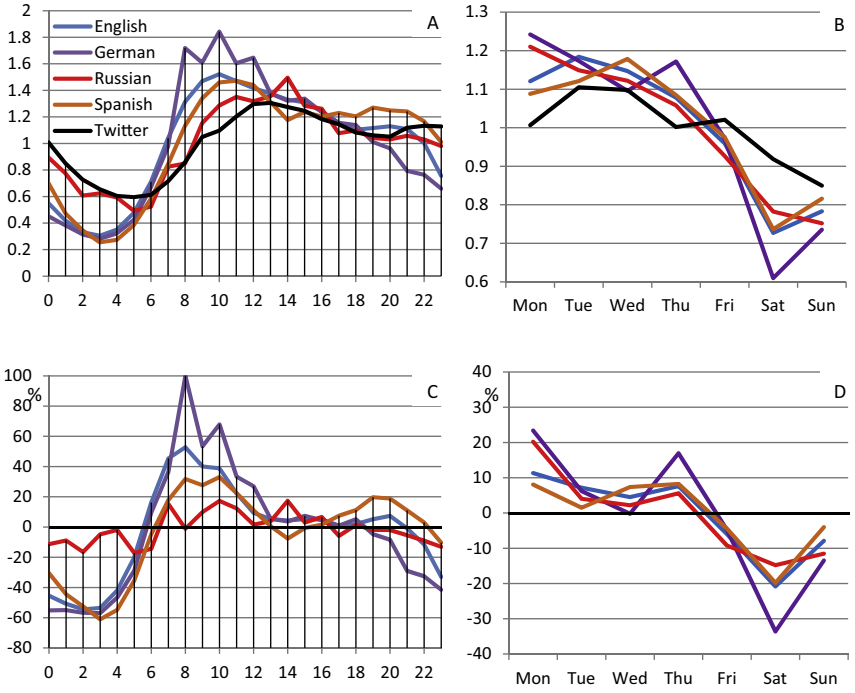

Fig. 2. Relative number of tweets on climate change R_h in four languages over the hours of the day h (A) and over the days of the week (B). Note the shifted and significantly more prominent daily and weekly cycles in climate change discussions relative to background Twitter activity TI_h (shown in legend as "Twitter", http://www.sysomos.com/insidetwitter). Plots C and D present the percentage deviation P_h from background Twitter activity, computed as $P_h = (R_h - TI_h)/TI_h \times 100\%$.

3 The days with less than 12 h of stored data were discarded; for days with $n > 12$ h $H = \{h_1, \ldots, h_n\}$ of tweets missing, the daily number of tweets T_{24} was extrapolated from the number of collected tweets T_n as follows:

$$T_{24} = T_n f_w(w_{loc}) \frac{n}{24} \sum_{h_i \in H} f_{h_i}(t_{loc})$$

2.4. Comparison of Twitter activity across countries

The number of collected tweets on climate change is highly heterogeneous (Section 3.1) across the 113 countries in the EGRS+P database, with the differences influenced not only by the prominence of the topic of climate change in the public discourse in a particular country but also by such factors as languages spoken, population size, rates of Internet penetration, and overall Twitter popularity. Several procedures were conducted to isolate the effect of the latter factors on the number of tweets about climate change, i.e., to obtain a measure of "true" popularity of the climate change discourse in each country. First, we estimated the "background Twitter activity" for each country in the database, which was defined as the number of tweets from a particular country across all languages and topics, including but not limited to, the topic of climate change. For that purpose, a sample of 1000 tweets was extracted every 5 min over a three-week period from June 13 to July 3, 2013. The data were collected using Twitter Firehose, a streaming API that returns a real-time, random sample of tweets (~1% of all tweets). The total number of tweets collected was 3,274,264. Then, for each country in the EGRS+P database, we estimated the EGRS+P tweeting intensity (TI), i.e., the number of tweets in the EGRS+P languages divided by its background twitter activity. Of the countries in the database, 78 had TI values above 0.5, while nine countries had TI values of less than 0.2. The latter countries were excluded on the grounds that over 80% of the tweets from these countries were in languages other than EGRS+P. Using these procedures, we removed the effects of languages other than EGRS+P from the data.

Finally, we removed the effects of differences in the data collection periods for various countries, population size, and overall Twitter popularity by computing the normalized tweeting intensity on the topic of climate change, which we consider the "true" measure of the climate change discourse in a particular country. The following procedures were performed to that effect:

1 The number of collected tweets on climate change for a country c from the EGRS+P database was divided by the number of days over which the tweets were collected to determine the mean daily tweeting intensity I_c for a country c;
2 Similarly, the mean daily background tweeting intensity I'_c was computed for a country c;
3 The relative climate change tweeting intensity for a country c was computed as $R_c = I_c/I'_c$, and
4 It was normalized as: $N_c = R_c/mean(R_i)$.

3. Results

3.1. Spatial distribution of tweeting on climate change

We evaluated the importance of the topic of climate change in Twitter discussions in different countries by comparing of the daily number of tweets originating from these countries, as determined by geolocation algorithm. We excluded countries with limited Twitter penetration (background Twitter activity below 10 tweets per day) and few tweets on climate change (fewer than one climate-related tweet per day). The direct comparison of the daily number of tweets (Appendix Figure B1) is skewed toward the countries where EGRS+P are the primary languages. We corrected for this bias by dividing the daily number of tweets on climate by the background tweeting intensity TI in the EGRS+P languages.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

Table 2
The countries contributing over 1% of the total number of tweets on climate change globally. Columns: average number of tweets per day (TPD), relative EGRS+P tweeting intensity (TI, i.e., the number of tweets in the EGRS+P languages relative to the total number of tweets originated in the country), and the percentage of the total number of daily tweets (PT), adjusted for TI (PT adj.).

Country	TPD	TI	PT	PT adj.
USA	1837	0.88	40.8	38.7
UK	561	0.87	12.5	12.0
Canada	330	0.86	7.3	7.1
Australia	233	0.87	5.2	5.0
Brazil	174	0.90	3.9	3.6
Indonesia	125	0.12	2.8	–
Spain	113	0.90	2.5	2.3
Mexico	90	0.92	2.0	1.8
Venezuela	81	0.93	1.8	1.6
Colombia	76	0.93	1.7	1.5
Philippines	61	0.48	1.4	2.3
India	61	0.78	1.4	1.4
Germany	58	0.79	1.3	1.4
S. Africa	44	0.88	1.0	0.9
Netherlands	32	0.32	0.7	1.9
Italy	27	0.33	0.6	1.5
France	26	0.21	0.6	2.3

Two countries, the USA and the UK, dominate the climate change discussion on Twitter: on average, 41% of the daily discussion of climate change on Twitter originates from the USA and 13% comes from the UK (Table 2). Other countries contributing over 5% to the climate change discussion include Canada and Australia; together, these four countries are responsible for 66% of the daily climate change discussion in Twitter. Two Portuguese-speaking countries, Brazil and Portugal, contribute 3.9% and 0.3%, respectively, to daily tweeting on climate. The abovementioned bias correction made few changes to the list of the countries responsible for the majority of the climate change discussion (Table 2 and Fig. 3).

A similar distribution of tweets among countries was observed when we selected only those tweets published in one of a country's major languages. This shortened the list of countries for which at least 100 tweets were collected from 113 to 40 countries where English, German, Russian, Spanish, or Portuguese are the major languages. The general pattern of tweeting holds: 46% of the daily discussion of climate change on Twitter originates from the USA, 15% comes from the UK, followed by Canada and Australia; these countries are together responsible for 76% of the climate change discussion.

This daily tweeting intensity does not characterize the relative prominence of the topic of climate change in Twitter discussions, as (1) the countries differ in population, Internet penetration, and Twitter popularity and (2) the discussion can be conducted in languages other than the five selected. To account for these differences, we computed the normalized climate change tweeting intensity N_c. The normalized climate change tweeting intensity represents the relative abundance of the topic of climate change in microblogging, where the countries with the average topic abundance have the intensity $N_c = 1$, and the countries with over twice the average rate of climate change discussions with respect to their Twitter penetration have $N_c > 2$ (Fig. 3). The most-tweeting countries, the USA and UK, both have the above-average N_c (1.31 and 1.23, respectively). Discussions of climate change are also more prominent in the OECD countries (average $N_c = 1.42$) and Kyoto Protocol Annex 1 countries (average $N_c = 1.43$). Among the Annex 1 countries, Russia and Belarus have the smallest N_c (0.32 and 0.30, respectively) and are among the ten countries least-interested in the climate change discussion. The next smallest N_c between the Annex 1 countries is nearly twice as large (0.56) and belongs to Ireland.

The set of countries discussing climate change the most includes Qatar, Switzerland, Norway, Australia, Belgium, Canada, and the Netherlands (in decreasing order of N_c). We hypothesize that the abnormal number of climate change-related tweets in Qatar ($N_c = 4.1$; compare to $N_c = 0.27$ in Saudi Arabia) is an artifact created by microblog discussions of the visitors to the UN Climate Change Conference held in Doha, Qatar in November 2012, on the one hand, and low Twitter penetration in that country (0.2% of background tweeting intensity), on the other.

The distribution of discussions on climate change between the countries is highly uneven. Fig. 4 depicts the locations of urban areas (ESRI World Urban Areas) coincidental with climate tweeting intensity of at least one tweet per day. The majority of these places are located in Western Europe and the coastal areas of the USA. Of the 476 urban areas present in the database, 22 areas are responsible for over half of the total global discussion on climate

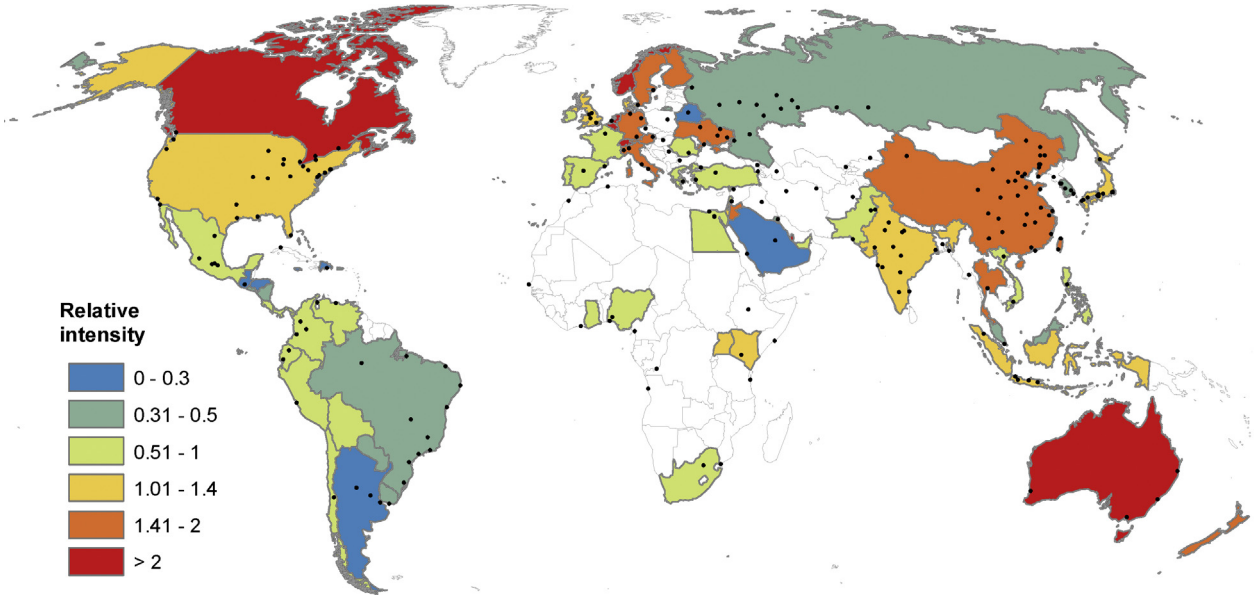

Fig. 3. Relative number of climate change tweets originating from different countries, adjusted for Internet and Twitter penetration and language proliferation. The dots indicate cities with populations over 1 million.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

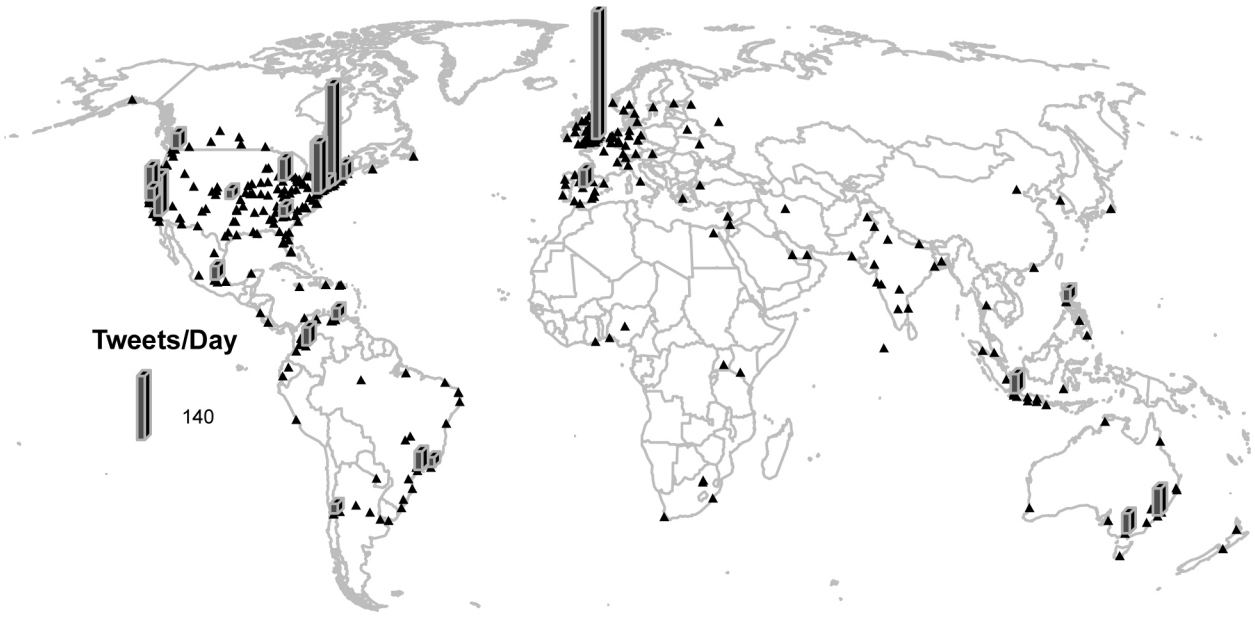

Fig. 4. Main urban areas in terms of tweeting on climate change. The triangles represent individual locations in the database, merged into urban areas, for which at least 1 tweet per day was collected on average. The bars indicate the most active individual locations with daily numbers of tweets over 20; the height of the bars represents the number of tweets per day with a maximum of 277 tweets/day in the London metropolitan area.

change (Table 3). When the tweeting intensity language bias (TI) is removed as described above, this has little effect on the list of the urban areas primarily responsible for tweeting on climate change (Table 3).

3.2. Temporal distribution and the most popular topics

The discussion of climate change on Twitter exhibits high temporal variability, which is partially explained by major news events (Section 3.3) and local weather (Names withheld, 2012). Correcting for data collection outages and weekly cycles (Section 2.3) resulted in a significant reduction in data variability, with the standard deviation declining from 1556 for the raw data to 1378 for the outage-adjusted data to 1241 for the weekly cycle-adjusted data (Fig. 5). The largest spikes in tweeting activity were linked to news regarding the Intergovernmental Panel on Climate Change (IPCC) report "Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation" (March 27th, event 2), the reversal of Richard A. Miller's climate change skepticism (July 28th, event 3), the op-ed in the Washington Post by James Hansen "Climate change is here" (August 4th, event 4), and regarding the landfall of Hurricane Sandy in New Jersey (October 30th, event 5). Other spikes in public interest in climate change were related to various news, weather events, and their combinations. For example, the spike in early February 2012 (event 1) was linked to several factors: the launch of a new UN climate change adaptation database, news regarding the Climate Reality project on Antarctic ice melt, the Daily Mail publication on a new ice age, and other events.

Table 3
Metropolitan areas producing half of the tweets on climate change daily.

Urban Area	Country	Tweets/day	Percentage of all tweets	Percentage adj. for TI
London	UK	277.3	10.2	9.8
New York	USA	218.4	8.0	7.6
Washington	USA	109.6	4.0	3.8
Los Angeles	USA	87.4	3.2	3.1
San Francisco	USA	69.9	2.6	2.4
Toronto	Canada	66.2	2.4	2.4
Sydney	Australia	58.7	2.2	2.1
Paris	France	37.4		2.0
Chicago	USA	47.9	1.8	1.7
Melbourne	Australia	43.2	1.6	1.5
Jakarta	Indonesia	41.6	1.5	1.5[a]
San Paulo[b]	Brazil	40.6	1.5	1.4
Bogota	Columbia	38.8	1.4	1.3
Madrid	Spain	37.1	1.4	1.3
Vancouver	Canada	33.2	1.2	1.2
Boston	USA	31.7	1.2	1.1
Mexico City	Mexico	29.1	1.1	1.0
San Francisco	USA	26.7	1.0	0.9
Manila	Philippines	26.3	1.0	1.7
Rome	Italy	40.7		1.6
Caracas	Venezuela	23.6	0.9	
Atlanta	USA	22.6	0.8	0.8
Santiago	Chile	21.2	0.8	
Richmond	USA	21.0	0.8	

[a] Not adjusted for TI.
[b] Values are based on the 2013 dataset.

The topics of discussion are indexed by Twitter users through hash tags (keywords) embedded in tweets. In total, 58,746 hash tags were used in tweets. The top hash tags were primarily relevant to climate change in all languages, with exception of Russian, in which all of the most frequent hash tags were generic (e.g., "News", "FOLLOWBACK"). Aside from general climate-related keywords (e.g., "climate" – see Table 4), among the most frequent hash tags in all languages were COP-18 meeting in Doha and Hurricane Sandy. Among the country- and language-specific hash tags are those related to local politics and activism, for example, the Wrinkled Shirt Day, an Argentinean NGO initiative to curb energy use by not ironing clothes (hash tag DiaDeLaCamisaArrugada).

The temporal distribution of hash tags reveals changes in interest regarding specific topics of discussion over time (Fig. 6). Several events produced short-term, high-intensity spikes of interest. Among those events are two high-profile international meetings, the United Nations Conference on Sustainable Development in June and the UN Climate Change Conference in November. Several election campaigns in which the candidates debated climate change are also present in the data. Those events include the Alberta provincial elections, US elections, and discussions on the Keystone pipeline. Among the political forums, the Top

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

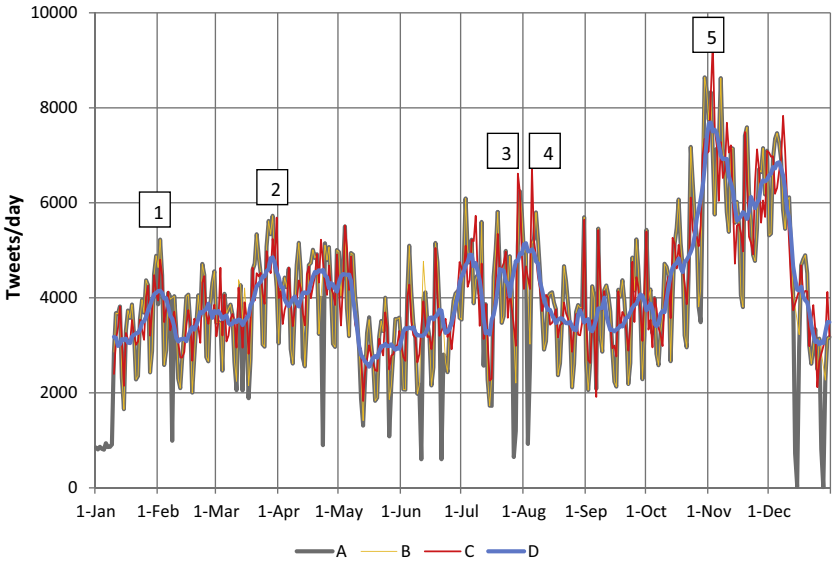

Fig. 5. Intensity of tweeting on climate change (tweets/day): raw data (A); adjusted for server downtime (B) and for the weekly cycle (C); weekly running mean (D). The numbers identify major spikes in climate discussion activity: (1) Warm winter spell in the U.S. and other events, e.g. the launch of the UN climate change adaptation database, Climate Reality project on Antarctic ice melt, and Daily Mail publication on a new ice age; (2) IPCC report "Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation"; (3) Reversal of Richard A. Miller's climate change skepticism; (4) Op-ed in the Washington Post by James Hansen "Climate change is here"; (5) Hurricane Sandy landfall.

Conservatives on Twitter hash tag, tcot, which has over 1 million followers, was the most frequent. The most frequently used in terms of the number of days and subject of tweets was, however, Hurricane Sandy, which made it the most discussed event related to global climate change.

3.3. The most authoritative Internet sources of information

The tweets frequently refer to external resources. Of the entire database, 34% of the tweets contained references to at least one resource in the Internet. In total, these URLs pointed to over 330,000 unique pages at 46,863 different domains; however, 50% of the references point to just 129 (0.28%) of the most popular domains. The majority of the most popular resources included traditional news sources, news aggregators, popular science journals, personal blogs, and political science sites. To locate original content providers, we excluded generic domains such as social networks other than Twitter (e.g., Facebook), picture and video sharing sites (e.g., YouTube), web analysts (e.g., SNSanalytics), and advertisement sites (e.g., credit-help24.org).

Table 4
The most frequently used English-language hash tags, defined as tags used at least 50 times over at least 4 days in 2012. Tags with identical meanings (e.g., Climate, climate change, global warming) are aggregated. Keyword explanations were obtained from http://tagdef.com.

Hash tag	*N*	Explanation
Climate	150,970	Tweets concerning climate change
tcot	10,205	Top conservatives on Twitter is a coalition of conservatives on the Internet
auspol	8800	Australian politics
Sandy	8697	Hurricane Sandy, which struck New York in October 2012
COP18	8488	UN Climate Change Conference of the Parties
p2	6717	Progressives using social media
cdnpoli	4467	Generic Canadian political issues
qanda	1906	Tweets relating to the Australian ABC1 panel discussion TV program Q & A
RioPlus20	1548	Rio+20 United Nations Conference on Sustainable Development
abvote	1154	Alberta 2012 provincial election
Keystone	745	Keystone pipeline
Current 2012	690	US presidential elections
Debates	598	US presidential elections – debates

Overall, the most authoritative (i.e., the most frequently referenced) source of information on climate change in Twitter discussions was *The Guardian* (5% of all climate change references) followed by news aggregator *Huffington Post* and political blog *ThinkProgress*. Other popular traditional news sources on climate change included *The Washington Post*, *Forbes*, and the *Daily Mail*. This list does not include *The New York Times*, as the relevant content was primarily published in its blogs *Green* and *Dot Earth*. The most authoritative climate change skepticism web sites included *Watts Up With That?* and *Climate Depot*. Finally, among the popular science and science sources, the leading news outlets were *Scientific American*, *Science Daily*, and *National Geographic* (Table 5). Table 6 lists the three most frequently referenced news resources in the five languages considered.

With respect to the most frequently referenced web pages, many belonged to traditional media, such as *The Guardian*; however, political activism pages were also common (Table 7). We observed that some blogs are closing the popularity gap with traditional media as the top source of information on climate change; nevertheless, they are an order of magnitude less referenced. The most popular blogs, such as the *New York Times* blogs *Green* and *Dot Earth*, however, are operated by professional journalists, and therefore are not entirely independent, filling the niche between the traditional news sources and the new, grassroots web-based journalism. Publications in the top standalone blogs, *Watts Up With That?* and *Climate Depot*, both devoted to climate change skepticism, were referenced 20 times less frequently than the top-cited traditional newspaper, *The Guardian*. Overall, the weekly number of tweets on climate change is strongly correlated with the number of climate change news articles published in 80 major world newspapers[1] (Pearson's $r = 0.46^{**}$,

[1] 2012 weekly numbers of newspaper publications were obtained using LexisNexis Academic searches for "climate change" or "global worming" using the databases Major US Newspapers and Major World Newspapers.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

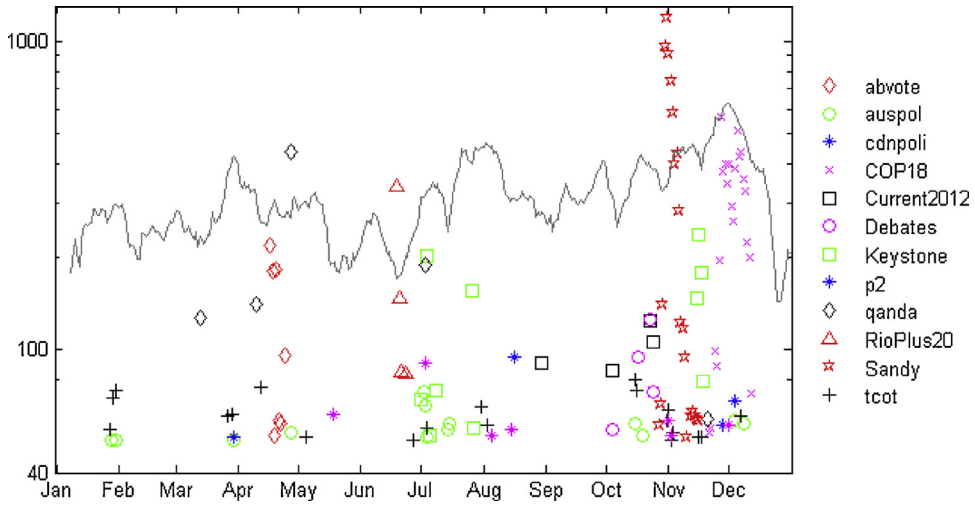

Fig. 6. Distribution of the most frequent hash tags over time. For legend explanations, see Table 4. The line represents the weekly running mean of the cumulative use of the tags Climate, Climate Change and Global Warming.

Table 5
The most frequently referenced domains in tweets on climate change (percentage of the total number of tweets with references).

Traditional news	%	News aggregators	%	Science	%	Blogs	%
guardian.co.uk	5.0	huffingtonpost.com	3.5	scientificamerican.com	0.7	green.blogs.nytimes.com	0.5
washingtonpost.com	1.2	thinkprogress.org	2.2	sciencedaily.com	0.7	wattsupwiththat.com	0.2
forbes.com	0.8	grist.org	1.6	nationalgeographic.com	0.5	climatedepot.com	0.2
dailymail.co.uk	0.8	treehugger.com	1.3	thecarbonfootprint.info	0.5	blogs.discovermagazine.com	0.1
bbc.co.uk	0.7	motherjones.com	1.2	phys.org	0.4	skepticalscience.com	0.1
rollingstone.com	0.6	scoop.it	0.6	newscientist.com	0.3	**Political action**	**%**
smh.com.au	0.6	mediamatters.org	0.3	rtcc.org	0.3	greenpeace.org	1.1
telegraph.co.uk	0.6			climatecentral.org	0.3	care2.com	0.4
latimes.com	0.6			livescience.com	0.3	credoaction.com	0.3
independent.co.uk	0.5			nature.com	0.2	good.is	0.2

Table 6
The most frequently referenced resources in tweets on climate change, in different languages.

English	German	Russian	Spanish	Portuguese[a]
Guardian	Spiegel	news.yandex.ru	diarioEcologia.com	uol.com.br
Huffington Post	Zeit	vk.com	efeverde.com	envolverde.com.br
ThinkProgress	Sueddeutsche	meteopathy.ru	actualidad.rt.com	abril.com.br

[a]Based on 2013 sample.

Table 7
Ten most discussed 2012 news events on Twitter.

Cir. date	Resource	Name	Content
29/1	Daily Mail	Forget global warming – it's Cycle 25 we need to worry about (and if NASA scientists are right the Thames will be freezing over again)	From 1997, there was no discernible rise in aggregate global temperatures; mini-ice age may be coming
17/3	Forbes	Scientists Call For Stronger Global Governance To Address Climate Change	The paper in Science signed by 32 scientists
3/4	credoaction.com	Tell the EPA: We want stronger rules to fight climate change	Online petition to the US EPA
18/6	endfossilfuelsubsidies.org	Twitterstorm	Petition to the Rio Earth Summit to stop fossil fuel subsidies
20/6	Greenpeace	Petition "Global Sanctuary in the Arctic"	A call to sign the petition
28/6	ThinkProgress	Bombshell: Koch-Funded Study Finds 'Global Warming Is Real', 'On The High End' And 'Essentially All' Due To Carbon Pollution	The BEST publication and BEST director Richard Muller's conversion from climate change denial
19/7	Rolling Stone	Global Warming's Terrifying New Math	US warmest spring on record
3/8	Washington Post	Climate change is here—and worse than we thought	Op-ed by Director of NASA GISS J. Hansen
13/10	Daily Mail	Global Warming Stopped 16 Years Ago, Reveals Met Office Report Quietly Released... And Here Is The Chart To Prove It	From 1997, there was no discernible rise in aggregate global temperatures
1/11	Businessweek	It's Global Warming, Stupid	Hurricane Sandy aftermath

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

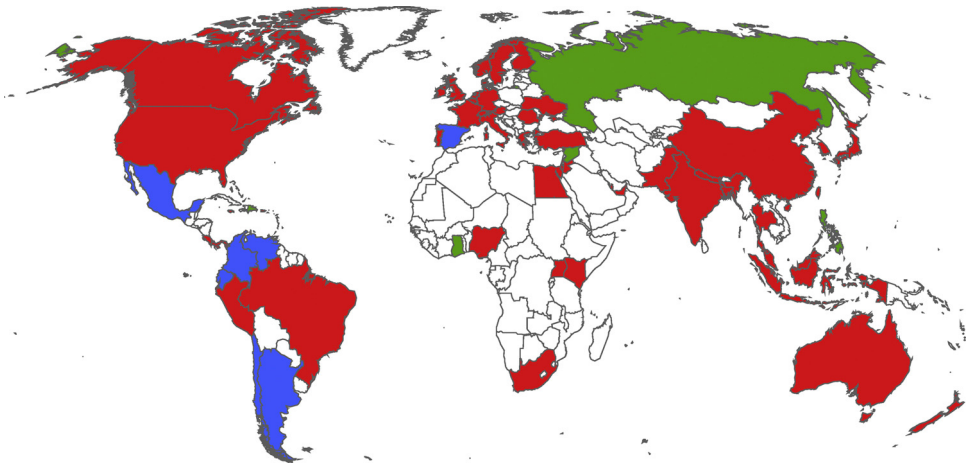

Fig. 7. Clustering of the countries with respect to the most frequently cited web sources.

Spearman's $\rho = 0.45^{**}$). When a one-week lag is applied to newspaper publications, the correlation increases ($r = 0.62^{**}$, $\rho = 0.47^{**}$). The correlation is even higher for major US newspapers ($r = 0.64^{**}$, $\rho = 0.58^{**}$ without lag; $r = 0.76^{**}$, $\rho = 0.51^{**}$ with lag), which can be explained by the US dominance in the climate change discussion on Twitter.

Among the most frequently referenced sources of information on climate,[2] some of the web resources are either frequently cited in almost all (*Huffington Post* and *The Guardian*) or in the majority of countries (*BBC*, *Grist*, *ThinkProgress*, *The Washington Post*, *TreeHugger*, *Scientific American*, *Forbes*, and *Science Daily*). However, other frequently cited web resources differ across countries. We performed hierarchical cluster analysis to aggregate countries into a few homogenous groups based on their patterns of referencing web resources. Citation frequencies for each country are represented by a sparse frequency matrix (sparsity 0.11); thus, we selected between distance measures that exclude joint absences from consideration; otherwise, the absence of the same web resource in two countries would artificially inflate the similarity of those countries. We rescaled the frequency matrix into a 0/1 presence/absence matrix and applied the centroid and average (within-group) linkage clustering methods using the Jaccard and Dice coefficients as distance measures with similar outcomes. Fig. 7 illustrates a robust, 3-cluster solution that used the centroid method and Dice distance measure. Generally, the countries are clustered into (1) a group primarily comprising English-speaking and developed countries, (2) Spanish-language countries, and (3) outliers that primarily use their own specific information sources (e.g., Russia). The increase in the number of groups divides the first group into smaller groups of countries with primary languages other than English.

3.4. The most authoritative Twitter bloggers

Twitter users can link their messages to existing tweets or other users through the mechanisms of a retweet (retransmitting the message), a reply (replying to a message), or a mention (mentioning a Twitter user, including a retweet), thereby creating conversations. Similar to external references, among the 0.5 million retweets in the database, the majority links to a limited number of Twitter users, with 50% of all retweets linking to just 0.37% of all users. Three Twitter users, *Mother Jones* magazine, *Climate Reality*, *a* climate advocacy organization, and *Grist*, an online news aggregator, are mentioned in at least 1% of all retweets. The Appendix Table C1 presents the 36 top Twitter users, which are mentioned in at least 0.25% of all retweets. As a group, they are mentioned in 20% of all retweets. While traditional news organizations are by far the most common external Internet references, internally, Twitter users' messages refer to the Twitter accounts of traditional news outlets, popular science magazines, advocacy organizations, bloggers, and celebrities with equal frequency. The order in this list, however, is entirely determined by the dominance of the English language on Twitter.

The list of the most authoritative Twitter bloggers exhibits much higher country-to-country variability, relative to the most authoritative external Internet sources. In German-speaking countries, in addition to English resources, the *Klimawandel-CO2* (Dritter_Planet) climate change science popularization web site is mentioned the most. In Spanish, the most authoritative resource is *Diario Ecologia* (Diario_Ecologia). Only one Twitter account, Al Gore's *Climate Reality* project, is frequently mentioned in at least half of the countries present in the database, which makes it the single most important internationally Twitter resource on climate change. Two of the other most authoritative Twitter bloggers, *Mother Jones* magazine and *Grist*, a news aggregator, are almost exclusively mentioned in English-speaking countries. Overall, large international climate advocacy organizations (e.g., Greenpeace) and celebrities (e.g., Ian Somerhalder) prevail among the most authoritative Twitter users blogging on climate change.

4. Discussion

This study analyzed Twitter data on climate change using the geographical distribution of tweets, their weekly and daily patterns, major news events that affected tweeting on climate change, frequently referenced web resources, and the most authoritative Twitter users. Geographically, the discussion of climate change is concentrated in the industrially developed countries of Europe, Australia, North America, India, and the Philippines, with the highest percentage of tweets coming from the London and New York urban areas. While the daily tweets from the US and UK account for a striking 41% and 13% of all tweets, respectively, the relative weights of these countries in climate change discussions are not the largest. The countries of Latin America participate in the discussion using their own distinct set of authoritative Spanish-language web resources, with majority of

[2] We arbitrary defined the most frequently cited resources as those referenced in at least 0.5% (but no less than 500) of messages, containing references. Additionally, we discarded the references to sharing and social networks (e.g., twitter.com, youtube.com), advertising and web analytics (credit-help24.org), and redirections that failed to resolve (tinyurl.com).

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

tweets originating from the largest cities in these countries. Among the five languages, the number of tweets originating from countries with significant Russian-speaking populations was the smallest. Tweets' content and the most authoritative sources for Russia are also dissimilar to other language clusters, primarily concentrating on jokes or expressing climate change skepticism and irrelevance. Overall, a significant percentage of tweets referenced climate skepticism sources, mainly Andrew Watts' blog *Watts Up With That?* and the web portals *Climate Depot* and *Skeptical Science*. Similarly, many tweets mentioned celebrities and organizations known for their climate change skepticism. This is especially relevant for tweets originating in the USA.

The climate change discussion on Twitter exhibits significant temporal variability, with strong daily and weekly cycles. The distribution of Twitter discussions over the time of the day and days of the week differ markedly from Twitter's baseline, with a substantially higher number of messages on climate change being published during the morning working hours (Fig. 2C) and during the first four days of the week (Fig. 2D). For example, in the English language, on average 40% more tweets on climate change are published between 7 am and 12 pm. We speculate that this difference suggests that a sizable amount of tweeting on climate change originates from a workplace, possibly as part of Twitter users' work duties. This interpretation seems entirely feasible, as a substantial volume of messages comes from the Twitter arm of long-established media, such as, for example, *The Guardian*. Interestingly, for the Russian language, tweeting on climate change exhibits little deviation from the baseline activity level, which is in line with the limited number of references to the traditional media in this country. Throughout the year, there were large changes in the intensity of tweeting on climate change, governed by major news events. By far, the most significant of these events was Hurricane Sandy, which almost doubled the daily number of messages. Other major events that ignited online discussions of climate change included election campaigns in the USA, Canada, and Australia and major international meetings such as the COP-18 conference in Dubai and the UN conference on sustainable development in Rio de Janeiro.

The frequency of the use of URL references to other domains and mentions can be utilized as an indicator of the most authoritative sources of information on climate change. Half of the messages with external references point to just 129 domains (0.28% of all domains). Yet, for the majority of these domains, their popularity in the climate change discussion is limited to relatively few countries. Only ten of these domains are present in the list of the most popular domains in at least of half of the countries included in this analysis. Of these ten domains, nine belong to the prestige press, top popular science magazines, and news aggregators, and only one (ThinkProgress.org) is a blog. These findings are in line with those reported by Brulle et al. (2012) in a study of 74 surveys of public concerns regarding climate change; they found that mass media coverage and elite cues were among the three most important factors in forming public opinion (the third factor was related to economic conditions). The retransmission of Twitter messages is likewise highly concentrated; with 50% of the retweets pointing to just 0.4% of Twitter users mentioned in climate change tweets. The top mentions are an almost identical list of online prestige press outlets and news aggregators, with addition of popular bloggers and celebrities. The evidence of "core" resources is in agreement with prior studies. Hindman (2008) noticed that the political discussions in the Internet tend to be highly concentrated, with the majority of hyperlinks pointing to very few newspapers, journalists and prominent bloggers, which makes is hard for minorities and non-mainstream messages to be heard. In the context of Twitter discussions of the news, Heim (2013) found that the elite journalists dominate the discourse. Wu et al. (2011) found that with respect to the overall activity of Twitter users across all topic areas 50% of references point to less than 0.05% of users' population. Hindman (2008) argued that Web infrastructure itself is conducive to the apparent bias toward the most popular Web resources, e.g. Google's PageRank search algorithm serves as a positive feedback, discriminating against all but the most visible content providers. His observation that "it may be easier to speak in cyberspace, but it remains difficult to be heard" is very pertinent to our findings. Overall, the online climate change discussion seems to be highly concentrated and mainly representing the views expressed by elite journalists, organizations, and celebrities, which might be influenced by the dominance of the political component in the discussion.

The public skepticism of general media, the fragmentation of news sources, and news customization created by advances in technology, such as the Internet, social networking sites, and RSS streaming capabilities, have led communication scholars to argue that we are witnessing the rise of more egalitarian, bottom-up news making characterized by minimal effects of media messages on audiences (Bennett and Iyengar, 2008). This view has precursors in the "two-step flow" model (Katz, 1957; Lazarsfeld et al., 1948), albeit at a new level, which posited that the flow of mass communication and, consequently, its effects on public opinion were not direct but mediated by *opinion leaders*, knowledgeable, media-savvy individuals who could influence the opinions of people in their circle of friends and relatives. Thus, the model emphasized the interpersonal character of information exchange and limited effects of mass media and their institutions. Other authors, however, noted that in reality the majority of alternative media discuss news reports initially published by the traditional set of "prestige press" outlets. Both tendencies are observed in this study. While Twitter's model is based on interpersonal communication, the majority of messages with URL references and mentions refer to a relatively small (fewer than 500) sources of information, divided among the "prestige press" (Stempel, 1961), news aggregators, celebrities, organizations, and a recently emerged group of individual newsmakers and commentators, e.g., popular bloggers (cf. Wu et al., 2011). Thus, on Twitter, opinion leaders remain important to the discourse on climate change.

The aim of this study was to extract the entire population of publicly available tweets on climate change over the reference periods. An alternative approach would have been to extract a sample of all messages using Twitter Firehose API. However, we believe that working with the entire population of messages, while challenging in terms of data management and processing, provides an important advantage over using a sample, especially for the analysis of less-represented geographical regions, population groups, or views. Unfortunately, the current functionality of the Twitter API is not advantageous to developing a search design that would solve the sampling problem. With the increased proliferation of GPS-enabled devices, more sophisticated message extraction methods will be possible; however, only approximately 1% of all tweets are currently georeferenced. While the majority of studies on Twitter geography only consider georeferenced tweets, excluding non-georeferenced messages inevitably skews the analysis of the climate change discussion toward the most technically savvy population. Thus the algorithms resolving tweets' geographical location from textual data seem highly important for future studies.

Data availability is a major consideration in research involving Twitter data. Currently, there are no free, publicly accessible archives of Twitter data, and Twitter's search capabilities are limited in terms of both the volume of returned data and time frame. Many researchers have developed various tools based on the Twitter application programming interface (API) (Russell, 2011) to aid data collection. The rapid evolution of the programming interface,

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

however, requires regular updates to data harvesting software. Even when a study attempts to collect all Twitter data, in reality only the publicly available component of Twitter (or any other social network) data are accessible for extraction. It is unknown if and how the pattern of private online discussions differs from the open ones. In addition, we found that some data from the open discussions are also not extracted. For example, while we attempted to collect all tweets meeting the search criteria, some tweets were lost. To estimate the volume of tweets missed by our search, we simultaneously collected data for one day from geographically distant locations, using different Internet providers. This resulted in an approximately 1% difference in collected data.

An important limitation of using Twitter for research on the climate change discussion is variation in Twitter penetration in different countries. Notably, the Chinese component of the dataset is extremely unreliable, as Twitter is officially banned in that country; therefore, the collected data reflect the opinions of the Chinese population accessing the Twitter through proxy servers or living outside the country. We attempted to mitigate this problem by (1) adjusting the number of collected tweets per country using Twitter penetration rates and (2) conducting a separate analysis of the 40 countries where EGRS+P are the main languages. Nevertheless, the collected data may be biased toward the English-speaking population.

Compared to traditional surveys, user-generated content, such as Twitter, has the definite advantage of being a non-intrusive research mode that lacks the complications associated with interactions with human subjects. Additional benefits are data availability (albeit not data completeness) and speed and the relative simplicity of the data collection process. Twitter data allow studying manifestations of public opinion as they naturally occur. Geolocation information makes Twitter data particularly suitable for comparative, cross-cultural research, the area in which the traditional survey mode requires substantial investments of time, effort, and money to ensure the cross-cultural comparability of surveys. In addition, the temporal information contained in Twitter data makes it possible to monitor public opinions in real-time, thereby treating Twitter data as a distributed network of human sensors. The inter-connectedness of Twitter data makes it a perfect source for studying the network distribution of messages, i.e., "who says what to whom." As Ravikant and Rifkin (2012) observed, the value of Twitter data is increased by graphing the connections among its users: the interactions on Twitter are predominantly interest-related, as opposed to, for example, Facebook, another highly popular social network, where users are connected on a friend-related basis. Representing such features of Twitter as lists allows users to be classified into groups that exhibit various degrees of influence on public opinion, as demonstrated by Wu et al. (2011), with the potential to quantify these influences. The internal structure of the Twitter message is convenient for a much more "finely granulated" content analysis than that of presented in this study. Tweets can be considered extremely focused messages of 140 characters; the very length of such messages ensures that all content peripheral to the core message has already been shed. Thus, each tweet is a naturally occurring unit of content that is particularly suited for computer-assisted content analysis. The automated counts of the most frequent words, be they hash tags or "meaningful" words in the body of the message, can be employed with confidence that the law of large numbers will reveal the essential issues in public communication. Further, sentiment analysis algorithms (Pang and Lee, 2008) will allow researchers to monitor public attitudes toward climate change, an area that is primarily assessed with traditional surveys at present. We anticipate that mining social networks will become a major source of data on the public discourse on climate change, similar to what has occurred in many other fields of research.

5. Conclusion

Following Lasswell et al. (1952), we aimed to describe "with optimum objectivity, precision, and generality, what is said on a given subject in a given place at a given time" about climate change on Twitter in five main languages through a period of one year. We analyzed weekly and daily patterns of climate change discourse, major events that affect tweeting on climate change, frequently referenced web resources, and the most authoritative Twitter users. From a theoretical perspective, we found that the classic two-step model of communication still holds, as evidenced by high concentration of the discussion around a relatively small number of influential newsmakers. It has been claimed that recent developments in the Internet based technologies, focused on facilitating interpersonal collaboration and information sharing dubbed Web 2.0, brought new hopes of power decentralization through engaging general public into discussions on the most important issues of the day. In our study, however, we did not find evidence to support this vision. While the amount of messages related to climate change is staggering, the flow of information is highly centralized, with few media outlets, celebrities, and prominent bloggers leading the debate. From the applied research perspective, this study explored functionality of microblogging data for climate change communication research and showed that Twitter microblogging data is, in fact, conducive to communication research due to its availability, "people as sensors" geo-distributional nature, internal structure, and attached auxiliary information. Methodologically, we demonstrated that multiple issues related to message collection, geo-referencing, and interpretation, can be resolved with little manual intervention, allowing processing of very large volumes of data. We also suggested several technical solutions to data extraction and adjustments to overcome the limitations of Twitter's search interface.

Acknowledgements

The support for this study was provided by NASA Applied Sciences Program Grant #NNX10AH20G and 2013 UND Summer Professorship.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gloenvcha.2014.02.008.

References

Achrekar, H., Gandhe, A., Lazarus, R., Yu, S.-H., Liu, B., 2011. Predicting flu trends using Twitter data. In: 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Shanghai, China, 10–15 April, pp. 702–707.
Bastos, M.T., Zago, G., 2013. Tweeting news articles: readership and news sections in Europe and the Americas. SAGE Open 3 (3), 1–18.
Bennett, W.L., Iyengar, S., 2008. A new era of minimal effects? The changing foundations of political communication. Journal of Communication 58 (4), 707–731.
Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science 2, 1–8.
Bosch, T., 2012. Blogging and tweeting climate change in South Africa. Ecquid Novi: African Journalism Studies 33 (1), 44–53.
Boyd, D., Crawford, K., 2012. Critical questions for big data. Information, Communication and Society 15 (5), 662–679.
Brulle, R.J., Carmichael, J., Jenkins, J.C., 2012. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the US, 2002–2010. Climatic Change, http://dx.doi.org/10.1007/s10584-10012-10403-y.
Bruns, A., Burgess, J.E., 2011. # ausvotes: How Twitter covered the 2010 Australian federal election. Communication, Politics and Culture 44 (2), 37–56.
Cheng, Z., Caverlee, J., Lee, K., 2010. You are where you tweet: A content-based approach to geo-locating Twitter users. In: Proceedings of the Nineteenth ACM Conference on Information and Knowledge Management (CIKM'10), Toronto, Ontario.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008

Cheong, M., Lee, V., 2010. Dissecting Twitter: A review on current microblogging research and lessons from related fields. In: Memon, N., Alhajj, R. (Eds.), From Sociology to Computing in Social Networks. Springer, Vienna, pp. 343–362.
Costolo, A., 2012. Internet Advertising Bureau's Engage Conference Speech. http://news.cnet.com/8301-1023_3-57541566-93/report-twitter-hits-half-a-billion-tweets-a-day (last accessed 21.10.13).
Eisenstein, J., O'Connor, B., Smith, N.A., Xing, E.P., 2010. A latent variable model for geographic lexical variation. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, MA.
Gao, H., Barbier, G., 2012. Harnessing the crowdsourcing power of social media for disaster relief. IEEE Intelligent Systems 26 (3), 10–14.
Gelernter, J., Mushegian, N., 2011. Geo-parsing messages from microtext. Transactions in GIS 15 (6), 753–773.
Geller, A., Kaplan, D., Lasswell, H.D., 1942. An experimental comparison of four ways of coding editorial content. Journalism Quarterly 19, 362–370.
Graham, M., Hale, S.A., Gaffney, D., 2013. Where in the World are you? Geolocation and Language Identification in Twitter. Professional Geographer. (in press) In: http://arxiv.org/abs/1308.0683 (last accessed 21.10.13).
Heim, K., 2013. Tweeting the news: Adding Twitter and social media to journalism. In: Conference of the International Communication Association, London, UK, 17–21 June.
Hindman, M., 2008. The Myth of Digital Democracy. Princeton University Press.
Howe, P.D., Markowitz, E.M., Lee, T.M., Ko, C.-Y., Leiserowitz, A., 2012. Global perceptions of local temperature change. Nature Climate Change 3, 352–356.
Katz, E., 1957. The two-step flow of communication: An up-to-date report on an hypothesis. Public Opinion Quarterly 21 (1), 61–78.
Krippendorff, K., 2004. Content Analysis: An Introduction to its Methodology. Sage Publications, Thousand Oaks, CA.
Lasswell, H.D., 1948. The structure and function of communication in society. In: Bryson, L. (Ed.), The Communication of Ideas. University of Illinois Press, Urbana, IL, pp. 117–130.
Lasswell, H.D., Lerner, D., Pool, I.d.S., 1952. The Comparative Study of Symbols. Stanford University Press, Stanford, CA.
Lazarsfeld, P.F., Berelson, B., Gaudet, H., 1948. The People's Choice. Columbia University Press, New York.
Leiserowitz, A., 2012. Extreme Weather and Climate Change in the American Mind. http://www.climatechangecommunication.org/sites/default/files/reports/Extreme-Weather-Public-Opinion-April-2013_0.pdf (last accessed 21.10.13).
Molodtsova, T., Kirilenko, A., Stepchenkova, S., 2013. Utilizing the social media data to validate "climate change" indices. In: American Geophysical Union Fall Meeting, San Francisco, CA, 9–13 December.
Pang, B., Lee, L., 2008. Opinion mining and sentiment analysis. Found. Trends Inf. Retr. 2 (1–2), 1–135.
Ravikant, N., Rifkin, A., 2012. Why Twitter is Massively Undervalued Compared to Facebook, Tech Crunch. http://techcrunch.com/2010/10/16/why-twitter-is-massively-undervalued-compared-to-facebook/ (last accessed 21.10.13).
Ross, C., Terras, M., Warwick, C., Welsh, A., 2011. Enabled backchannel: conference Twitter use by digital humanists. Journal of Documentation 67 (2), 214–237.
Russell, M.A., 2011. 21 Recipes for Mining Twitter. O'Reilly Media, Sebastopol, CA.
Sakaki, T., Okazaki, M., Matsuo, Y., 2013. Tweet analysis for real-time event detection and earthquake reporting system development. IEEE Transactions on Knowledge and Data Engineering 25 (4), 919–931.
Satchwell, C., 2012. "Carbon literacy practices": textual footprints between school and home in children's construction of knowledge about climate change. Local Environment 18 (3), 289–304.
Segerberg, A., Bennett, W.L., 2011. Social media and the organization of collective action: using Twitter to explore the ecologies of two climate change protests. The Communication Review 14 (3), 197–215.
Semiocast, 2012. Twitter Reaches Half a Billion Accounts. More than 140 Millions in the U.S. . http://semiocast. com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US (last accessed 21.10.13).
Stempel, G.H., 1961. The prestige press covers the 1960 presidential campaign. Journalism and Mass Communication Quarterly 38 (2), 157–163.
Stone, B., 2009. What's Happening? . https://blog.twitter.com/2009/whats-happening (last accessed 05.06.13).
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: Proceedings of the ACL 2012 System Demonstrations, Jeju Island, Korea.
Williams, S.A., Terras, M.M., Warwick, C., 2013. What do people study when they study Twitter? Classifying Twitter related academic papers. Journal of Documentation 69 (3), 384–410.
Wu, S., Hofman, J.M., Mason, W.A., Watts, D.J., 2011. Who says what to whom on Twitter. In: Proceedings of the 20th International conference on World Wide Web, Hyderabad, India, March 28–April 1, 2011.

Please cite this article in press as: Kirilenko, A.P., Stepchenkova, S.O., Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change (2014), http://dx.doi.org/10.1016/j.gloenvcha.2014.02.008