VALIDATION OF RADAR BASED IDENTIFICATION OF HIGH-IMPACT WEATHER PHENOMENA BY SOCIAL MEDIA REPORTS

Dmitri Moisseev[1], Otto Hyvarinen[2], Elena Saltikoff[2] and V. Chandrasekar[3]

[1] University of Helsinki, Helsinki, Finland
[2] Finnish Meteorological Institute, Helsinki, Finland
[3] Colorado State University, Fort Collins, Colorado

ABSTRACT

Many high-impact weather phenomena like thunderstorms, tornadoes and hail are very local and short-lived. These phenomena, therefore, are sparsely represented in surface weather reports. Weather radar is an unique tool for detection of high-impact weather phenomena, because it is one of the few instruments, which can provide observations over a large area with a high temporal resolution. Nonetheless, methods to observe such phenomena with a radar are always indirect, and include limited level of confidence. Validation of radar inferences using surface based observations, therefore, is essential before the automatic methods can be used as a part of a warning process. Finding the ground truth to validate these products is challenging and getting even more difficult, as the number of visual surface observations is decreasing [1].

In the last couple of years, social media reporting became a wide spread phenomena. Furthermore, it was shown that data from social media can be used for scientific purposes, i.e. to monitor and detect influenza epidemics [2]. In such services as Facebook, Twitter, YouTube and Flickr, to name a few, people share their experiences and report day-to-day events. Weather plays a big role in people life and it is not a wonder that it finds its place in their social media reports [3].

For this study, observations of hail storms and tornado events carried out by University of Helsinki C-band weather radar and CSU-CHILL National radar facility, are compared against reports published in Flickr. Flickr is an image sharing service established in 2004 and is presently owned by Internet services provider Yahoo!. At the time of writing, more than one billion photos were available for viewing and 6% of these had geographical information included. Most of the photos were taken in northern America and western Europe. We chose to use Flickr because the amount of photos meant the probability of finding useful data was high, and, because of the application programmer's interface (API), the data could be processed automatically. Images can have extensive both automatic and user-generated meta-data, such as, a short description of the photo, information about the camera settings, the date when the photo was taken and when it was uploaded, geographical information from GPS, and freely-chosen one-word "tags" that describe the photo. The exact nature of these user-generated tags is left to the users to decide. On one hand, this makes it easier for users to add tags and encourages their use, on the other hand, homonyms and synonyms can make their utilization later harder. For our purposes, the date, the

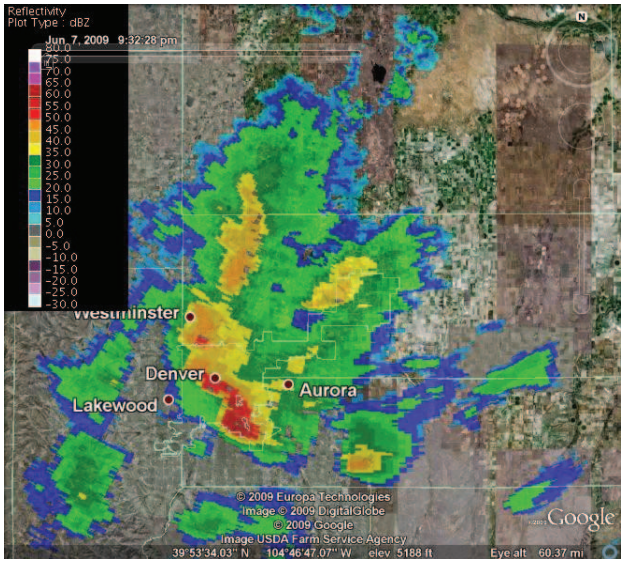
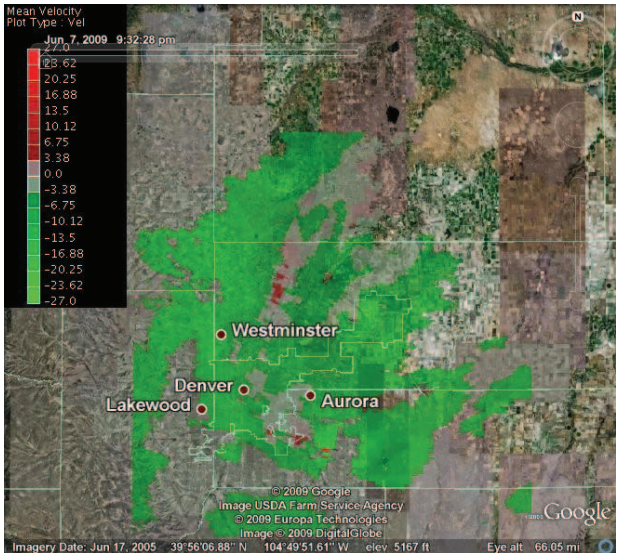


Figure1. CSU-CHILL observations taken on June 7, 2009 plotted on top of Google Earth map. The figure shows measured reflectivity field.



Same as on the left only for the radial velocity measurements.

geographical information and the tags were the most important. The tags are usually less ambiguous than the short description.

To estimate the usability of the Flickr reports for validation of radar inferences of tornadic event occurrence, a search of photos uploaded in 2008 with tags "Colorado" and "Tornado" was made. In total, there were 842 images that satisfied the search criteria. Among those reports, 424 images (50.3%) are corresponding to non-weather related things, such as an aeroplane or a rollerskating team. From the remaining reports, 226 photos (27%) showed tornado damage and 157 (22.7%) showed tornado funnels. Of the 157 tornado images one could identify 35 independent reports, i.e. same person took in average 4.5 images of the same event.

For more detailed study, we have selected one case that took place on June 7, 2009. According to the National Weather Service, 5 confirmed tornadoes were spawned across the Colorado front range during the afternoon of June 7th, 2009. CSU-CHILL radar was carrying out observations at the time. This research was facilitated by the use of Virtual CHILL (VCHILL) technology [4]. VCHILL is an example of a virtual environment that has successfully broadened research participation by providing remote access to radar data. As a result the Finnish co-authors of the paper had an easy access to the radar observations that were carried out by CSU-CHILL National Radar Facility.

In Table 1 details of selected Flickr reports are shown. We have performed a full text search using "tornado" and "Colorado" keywords and restricting dates to be after June 6 and before June 8, 2009. In total 207 reports that satisfy the search criteria were found. Only reports of one person for one tornado event were geotagged.

Nonetheless, as can be seen in Table 1, other reports provide valuable information. Most of the reports have references to cities, counties or other local geographical locations that allows to identify approximate locations of tornado events.

In Figure 1 CSU-CHILL observations collected at 19.32 UTC on June 7, 2009 are shown. In the radial velocity figure, one can observe at least 4 well defined circulation patterns. Two in Aurora, one in Denver and one north-east of Westminster. Those observations are in good agreement with the Flickr reports (Table 1). One of the challenges in radar identification of tornado events is to identify tornado touchdowns. Dual-polarization radar observations show a promise in providing such information [5]. A combination of social media reports and radar observations can be used to refine those identification techniques.

Social Media reports show a promise in providing a valuable surface based observations. These observations display a great promise in providing additional source for validation of radar based identifications of high-impact weather events. At the moment, Flickr photographs do not provide an overall statistical reference yet. The limitations of the current reports are partially caused by the lack geotagged observations. We expect that the number of the geotagged reports would dramatically increase in the near future. Probably, the most limiting factor for the statistical analysis of those reports is presence of relatively large amount of clutter, namely reports that are unrelated to the weather phenomena. This constrain, however, can be partially mitigated by presence of a group of volunteers who are interested in those phenomena and who preselects the reports. A good example of such a group, is "Severe Weather" group on Flickr.

REFERENCES

[1] Fiebrich, C.A., "History of surface weather observations in the United States," Earth-Science Reviews, Vol. 93, 77-84, 2009.

[2] Ginsberg, J., M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," Nature, Vol. 457, 1012-1014, 2009.

[3] Hyvärinen, O., and E. Saltikoff, “Social media as a source of meteorological observations,” Submitted to Mothly Weather Review, 2009.

[4] Chandrasekar, V., Y.-G. Cho, D. Brunkow and A. Jayasumana, "Virtual CSU-CHILL Radar: The VCHILL," J. Atmos. Oceanic Technol., Vol. 22, pp. 979–987, 2005.

[5] Kumjian, M.R. and Ryzhkov A.V., "Polarimetric signatures in supercell thunderstorms," J. Appl. Meteor. Climatol., Vol. 47, 7, 1940-1961, July 2008.

Table 1: Details of selected Flickr reports for June 7, 2009 tornado events that took place in Colorado.

Flickr user description	Taken on (local time)	Taken on (UTC)	Geo-tag	Posted	Tags	Flickr user-name
Aurora Colorado - Saddle Rock	2009:06:07 13:09:57 (MDT)	19:09:57	NA	June 8, 2009 at 7.34PM MST	Tornado, aurora, colorado, saddle rock, storm, denver, 2009	stevebeam
Tornado Denver, Colorado	2009:06:07	NA	NA	NA	Nikon D80, Nikkon 80-200mm, Denver, Colorado, Funnel Cloud, Weather, Tornado	Jomyke
Colorado tornado, far away in Aurora	2009:06:07 13:19:44 (MDT)	19:19:44	NA	June 8, 2009 at 12.19PM MST	tornado, hail, storm, weather, colorado, lafayette, indian peaks golf course, golf, golf course	Weffie
Elbert County, CO Tornado 6/7/2009 Viewed from 1 mile south of Byers, Colorado at 2:37 p.m. MST	2009:06:07 14:39:23 (MST)	21:39:23	NA	June 8, 2009 at 1.03AM MST	Tornado, Colorado, funnel	StormFarnik
Funnel Cloud - Colorado June 2009	2009:06:07 12:51:33 (MDT)	19:51:33	NA	June 8, 2009 at 7.59PM MST	Kirkachu, Hail, Tornado, Funnel, Cloud, Storm, Colorado, June	WendyWith Dolls
Aurora, Colorado - June 7, 2009 (tornado about to touch down)	2009:06:07 13:53:25 (MDT)	20:53:25	NA	June 7, 2009 at 5.36PM MST	Happykap, june 2009 tornado, aurora, colorado, June 7 2009	
Tornado Damage. A tornado this afternoon blew this truck off of I-70 near Watkins and also damaged a billboard on the opposite side of the Interstate when it crossed the highway	2009:06:07 15:37:19 (MDT)	22:37:19	NA	June 7, 2009 at 11.23PM MST	Tornado, Severe Weather, Rollover, Truck, Navajo, Storm, Interstate, Highway I-70, State Patrol, Colorado State Patrol, CSP, Colorado Adams County, Canon Xsi, Blake20CO	Blake20CO
Tornados-Weld County	2009:06:07 13:20:52 (MDT)	19:20:52	Location shown on the map	June 10, 2009 at 10.15AM MST	Tornados, tornado, colorado tornado, denver, denver tornado, hail, kenny ainsworth, nikon, d60, nikon d60	Kenny Eh