Supplemental Material

Weather and Forecasting

Enlarging the Severe-Hail Database in Finland by Using a Radar-Based Hail-Detection Algorithm and E-Mail Surveys to Limit Underreporting and Population Biases

https:doi.org/10.1175/WAF-D-19-0142.1

SUPPLEMENT TO

Enlarging the Severe-Hail Database in Finland by using a Radar-Based Hail-Detection Algorithm and E-mail Surveys to Limit Underreporting and Population Biases

JARI-PETTERI TUOVINEN, HARRI HOHTI, AND DAVID M. SCHULTZ

Survey methods and organization

This supplement describes the procedures for reporting hail in Finland since 2008. First, we describe the method for identifying possible people who could report the existence of the hail and their contact information. Then, we describe how survey emails were constructed (e.g., the framework of pre-filled survey draft and what information was asked) and what special requirements the surveys contained (e.g. bilingual setup and backup phone calls). Then, we document how replies were handled and placed in the database. Finally, we discuss how the project was organized and managed.

1. Finding potential e-mail recipients

Once a possible location for the hail was determined from the algorithm, we needed to identify potential observers of the hail. We used three map tools to get the location of the possible hail as exact as possible (within roughly a km) as the hail algorithm only overlays main roads, cities, lakes, and rivers. Fonecta (https://www.fonecta.fi/kartat) allows users to easily overlay the location of businesses and their contact information on the map. The same kind of result was given by the GT Drivers' Road Atlas of Finland (Genimap 2003) and Kansalaisen Karttapaikka, operated by the National Land Survey of Finland (http://kansalaisen.karttapaikka.fi/kartanhaku/osoitehaku.html?lang=en). The exact latitude–

longitude coordinates of a possible observer's location were determined from the Fonecta Web site. We tried to exclude individual residents as recipients of our hail survey because it would have been time consuming to match an individual's e-mail and street address together. In contrast, local businesses, libraries, village associations, local emergency personnel, hotels, and summer cottage renters were all seen as good recipients for many reasons. First, people tend to talk about unusual events in or near the library. Second, small villages that are near the brink of extinction because of rural depopulation are pleased to help as they can show the usefulness of their village. Third, emergency officials are usually the first to hear about any storm damage in the region. Fourth, different accommodation providers are spread widely around the country, and their owners or visitors are usually a good source of information. Fifth, workers for these businesses tend to live close by and, under ideal circumstances, ask their co-workers if they have heard or seen anything unusual. The result of these efforts led to finding recipients from low-density locations who helped decrease underreporting because other sources (e.g., social media) normally identified hail events from urban areas rather quickly.

Most businesses that were contacted had their own Web sites with contact information. Google searches could also reveal the right contact information. For each event, we tried to have at least two or three different recipients in the area in case one or two e-mails went unanswered. As people's memory will fade over time, we tried to send hail surveys within five days after the event. After five days, the number of human errors started to increase (e.g., wrong observation date, not recalling the time the hail occurred). Naturally, some e-mails yielded no information (e.g., broken e-mail addresses, a proper recipient from the area could not be found, the suspected hail event took place during the night or the weekend). As recipients were mostly businesses or state-run offices working office hours on weekdays, the weekends likely have had more cases that went undetected. Emails about hail cases occurring on Friday evenings were usually responded to on Monday at the earliest, which increased the possibility of unanswered e-mails. Indeed, emails sent on Fridays and Saturdays went unanswered or produced reports of no hail most often (286 of 1504, or 19%, of emails were unanswered on Fridays, and 256, or 17%, on Saturdays; Fig. S1). The average percentage of unanswered emails from Mondays through Thursdays was 12%. In addition, Sundays had the most days with hail observed (58 days or 19.1%) followed by Friday and Saturday (46 days, or 15.1%, and 44 days, or 14.5%), among all 304 hail days during the project. Days early in the week had the fewest hail observed (Mondays 35 days, or 11.5%, and Tuesdays 37 days, or 12.2%).

2. The content of the inquiry e-mail

We designed a standard pre-filled draft message to send to people who might possibly have seen hail (Fig. S2). The main message of the e-mail had to be short and clear. The recipient needed to easily identify that the sender was indeed personnel from this project—either from FMI or the storm-spotters' group, the Ursa Astronomical Association in Helsinki (i.e., the sender's e-mail address ending in fmi.fi or ursa.fi)—and that the subject title of the e-mail was interesting enough to open it (e.g., "Was there hail/severe hail in northern Helsinki on 7 July between 1800 and 1830 local time?"). The better the message was planned, the greater the chance was to get a response. When a project volunteer sent out a hail inquiry, a note was automatically added to the end of each e-mail stating that more information about the project could be obtained from a contact person at FMI.

Every e-mail contained a short introduction (e.g., who the sender was and why he or she was contacting them), a few words about the project (e.g., hail survey, testing the hail algorithm, future importance on warning for severe hail), and some information about the event (e.g., the time window when hail most probably occurred, the coverage of the area by mentioning some place names, from what direction the suspected hail cell approached). The essence of the e-mail was the following: "Was there hail, and, if so, were the maximum observed hailstones at least 20 mm in diameter in your location between xxxx and xxxx local time?" (Fig. S2). We also wanted the recipients to send their photos of the hailstones, if they indicated they had some in their reply. Finally, we stated the importance of responding to this e-mail.

Finland is a country with two official languages: Finnish and Swedish. According to the Finnish Population Register Center, around 5.4% of Finland is Swedish-speaking and they mostly populate coastal regions. Because we also wanted to get reports from Swedish-speaking areas, the draft message was also translated into Swedish. Because many storm spotters were uncomfortable with sending e-mails in Swedish (their non-native language), the first author contacted these areas when needed. During 2008–2017, a total of 62 e-mails in Swedish were sent, of which 61% received replies. During the busiest year 2010, 19 Swedish-language e-mails were sent, and 10 received replies, but none were sent in 2012 and 2013. More detailed statistics of survey emails appear in Table 3.

To avoid getting vague or general answers and to get the specific information we were looking for, we designed the e-mails to prompt its recipient for specific information without giving away too much information from the algorithm. First, we tried to help the recipient remember the event by describing the movement of the cell(s) (e.g., from what direction or place the cell(s) came from). Second, the recipient was told the approximate time when the hail algorithm showed high probabilities over or close to the contacted area (e.g., 1700–1730 LT). These two steps were done with caution because we did not want to get reports that simply verified the hail algorithm. Instead, by giving some first-guess time frame (usually 15–45 min in length), we were able to avoid receiving general, but largely unhelpful, expressions such as "after midday," and "later in the evening."

Sometimes unexpected exceptions occurred. For example, on 28 July 2009, we sent numerous hail surveys into central Finland (Fig. 2). One unlucky recipient replied to our e-mail five weeks after the event because the recipient's modem was damaged by a lightning strike during the hailstorm and, because of the rural location, was fixed weeks later. This event produced hailstones around 3 cm in diameter near the recipient's location, but also produced baseball-sized hail (7.5 cm or 2.9 inches in diameter) a few hours later. Of all 1109 survey emails that received replies, 810 (73%) received replies within three days from sending and 1065 (96%) received replies within a week.

Occasionally, we needed to take other measures as our e-mail approach was unsuccessful. Either the suspected area was unpopulated or the only e-mail recipient did not reply to the survey. In these situations, the telephone was the last option. A phone call was made to a citizen for whom a phone number, but no e-mail address, could be found from the Internet. All telephone contacts (29 calls) turned out to be successful as these citizens were able to confirm the occurrence of severe hail by themselves or by someone they knew (e.g., neighbor). The telephone was not used at all in 2012, 2015, and 2017, the survey's least-productive hail seasons with the fewest severe-hail days.

3. Receiving and archiving e-mail replies

After getting hail surveys back from recipients, the case was documented and placed in the database. Also, any additional questions were e-mailed to the recipient. The tone of the replies was generally very positive and polite. It was nice to get replies such as "How on Earth did you know it was hailing here!?", "I must say science has advanced quickly lately!", or "Well, I am very surprised about this question because it really did hail here." There were also numerous recipients who greatly appreciated contact with them and were keen to send their hail observations in the future without notice. Some recipients even contacted us after a hailfall even without our e-mail survey. For example, one recipient replied, "You asked about hail earlier in the summer, and I replied then. Yesterday evening, thunderstorms crossed our area again, and we received hail of 0.6 inches." The person went further with a follow-up question: "Was this thunderstorm cloud and hail seen by the weather radar?"

Naturally, we shared the hail tool's view from the time with the recipient and explained what hail-index values meant.

The contents of these answered e-mails were mostly informative. Around 6% of the recipients kept their reply short: "It was hailing here at the time you described, and the hail size was normal". This type of survey answer was stored into the database as hail size unknown during the suggested time window and resulted in 71 cases (Table 3). This inability to further query recipients who provided minimal information was one of the disadvantages of the e-mail surveys. However, our main goal was to gather severe-hail reports, and nearly all the short replies involved suspected cases of non-severe hail. If the recipient told us that he or she was not around during the time of the suspected hailfall, the reply was classified as non-informative, although these types of answers were rare with only 11 cases. Another more common and annoying disadvantage of e-mail surveys was non-existent e-mail addresses, which led to the project team having to find another recipient.

After any kind of hail-survey reply, the report was stored into the database of the hail project. Especially during the project phase with volunteers, each area (see section 4 for further discussion of these areas) kept record of all their sent and received e-mails in a Microsoft Excel form, but the final report needed to be filed within the storm-spotters' Web-based system (http://www.ursa.fi/ursa/jaostot/myrskybongaus/havainnot) where reports were finalized and transformed into e-mail hail reports. The contact person of the project was asked to fill in the maximum value of hail-algorithm output from the contact location by looking through the hail-algorithm archive, along with all the other information the reply email included. After this step, all members of the project and FMI forecasters were able to see the final event report via email. All e-mails, sent and received, were kept and saved at the project's Gmail inbox. The name of the person replying was kept anonymous and could only be seen by logging into the system. Normally someone from the area (see section 4 for further discussion of these areas) who was responsible during the week filed all received replies into the system by early the next week. If the reported hail size was larger than 2 cm, a report to the European Severe Weather Database was also generated simultaneously and sent automatically to their database. This procedure was in place during the volunteer phase, 2010–2015. For simplicity, the first author kept his own track of cases in Excel files and personal work e-mails during other seasons while working alone.

4. Organization and management

The first training meeting was held in April 2010 at FMI where 12 interested volunteers attended, were told about the goals of this project, and were shown an example of how the hail-survey process progressed step by step, including filling out the hail-reporting form and reporting received hail reports. In 2011 and 2012, interest grew and 14 volunteers participated. During 2010–2015, 23 different storm spotters took part in the project, several of whom participated in multiple years. At this point, the project was called TATSI, after the Finnish translation for the hail algorithm's enhanced tracking and verifying scheme. Each participant got access to the project Web site (Fig. S3), which included radar data, output from the hail-detection algorithm, 15-day archive data, and a few other helpful tools. During subsequent years, several northern Europe satellite loops, radar cross sections, and surface observations were added to the project Web site.

The storm-spotters formed groups, and each group was responsible for their own portion of Finland. During 2008, 2009, and 2015–2017, all of Finland served as a single area. In 2010–2012, the country was divided into four areas (Fig. S4). In 2013 and 2014, three areas (southern, central, and northern) were used instead. If a hail cell crossed the border of two areas, both areas were activated. The leader of each group was responsible for making work shifts for their area. Typically, two people were active in the hail survey, and the other two had a week off. This routine was changed on a weekly basis. Each group would have a primary contact person and an assistant assigned each week to track cells. The first author provided guidance and assistance, and he moderated over all areas of the project throughout the season.

What effect did this changing number of participants have on the statistics of the e-mail surveys and the project management? One hypothesis is that the first author, who was a paid employee of FMI and has followed the performance of the hail algorithm since 2006, would tend to have higher success in finding suitable recipients and receiving e-mail answers than the others. During 2008 and 2009, there was high return rate of survey e-mails that subsequently declined after 2010 (Table 3). However, two important factors needed to be addressed. First, the amount of sent e-mail multiplied in 2010 and 2011 with more volunteers involved, meaning more potential hail cells were covered on these active hail summers (22 and 17 severe-hail days versus 10 in 2009). Second, an increase in non-hail reports was not all bad. Instead, those reports helped us to reveal problems with the hail algorithm and led to a change in the survey's guidance. In 2012, the return rate increased again, but mostly due to it being a poor hail season (only six severe-hail days) and only a few hail cells reached the criteria for inclusion in the study (Table 3), allowing the volunteers more time to track down hail reports. The return rate stayed low during 2015–2017, likely a result of the increasing popularity of other sources of hail information.

During June 2013, FMI took steps towards launching an open-data philosophy by releasing weather data for use of the general public free of charge (Honkola et al. 2013). During 2014–2017, other datasets (e.g., real-time, historical, lightning observations, radar archive) were released. From our perspective, the data releases had a negative impact on the project. Many storm-spotters have special skills in the field of computers and coding, enabling them to construct their own special tools and algorithms. TATSI's benefit to participants was access to all the tools provided in the project site. So when data was available through other means, TATSI could not draw the attention of many of the spotters after 2013. The nature of volunteer (i.e., non-paying) project, volunteers' lack of time during the summer holiday season, and for some lack of interest in hail (e.g., more interest in tornadoes) were other reasons that they stopped participating. By 2014, eight spotters were still involved in the project, but in 2015 that number was down to five; thereafter, the author continued alone. Despite a changing number of participants, all days when the hail algorithm anticipated a good chance of severe hail were properly handled via e-mail surveys. During very busy periods (a couple days per summer), we prioritized the strongest hail cells according to the hail algorithm's output. However, this happened only a couple days per summer. By requesting severe-hail observations from the general public throughout each season and summarizing hail indices at the end of each season, we achieved greater recognition for severe hail and its underreporting, reflected specifically in the newspapers' Internet sites which reported hailstorms more frequently as the media began to recognize that this field of expertise was being covered by FMI.

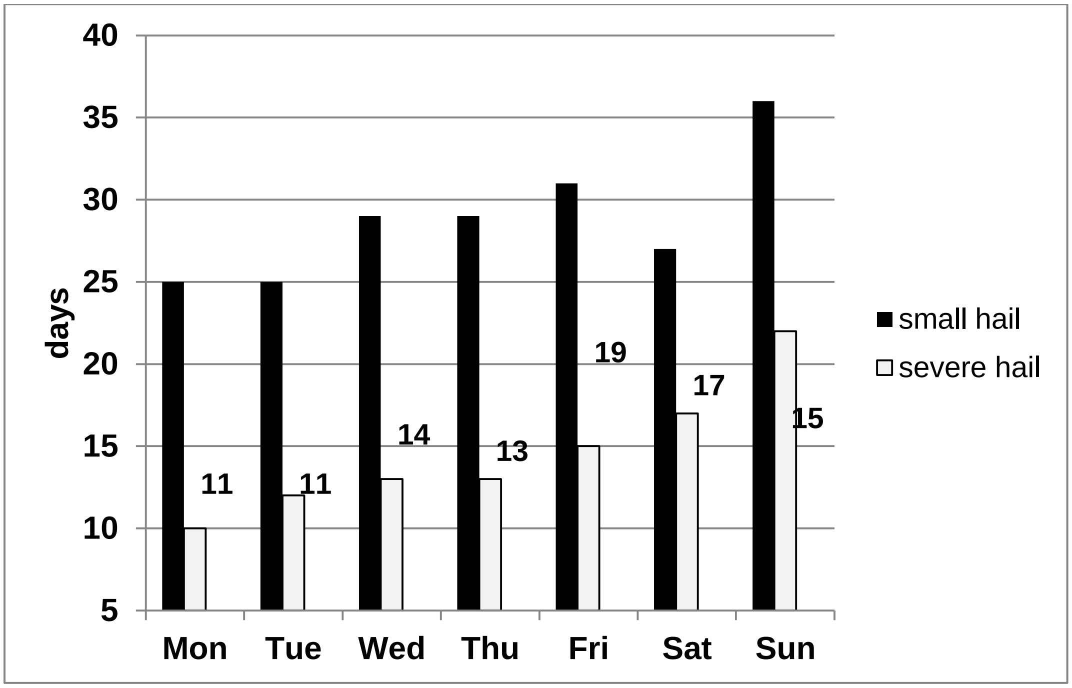

Figure S1: The number of hail and severe-hail days on each day of the week from late May to August during 2008–2017. The number over each bar indicates the percentage of unanswered or non-hail observations on each day of the week.

Hello!

I'm meteorologist Jari Tuovinen from the Finnish Meteorological Institute, and I would be keen to know if there was severe hail falling on Sunday 28 June around 1400-1415 LT in Ranta-Kosula village?

We are developing product that is based on weather radar, and this product can distinguish whether a thunderstorm produce hail on the ground or was just regular rainstorm with thunder. I'm sending this survey to places where this hail product has clearly indicated a signal about hail. During Sunday 28 June, there were numerous strong thunderstorm cells moving from north to south, and some have been observed to produce up to 5.5 cm diameter hailstones. I am mainly contacting local cottage renters, travel businesses, village associations, and libraries.

How large were these possible hailstones? We are interested mainly in about 2 cm or larger hailstones (comparable to 20 cent or 1 Euro coin). Have you heard others speaking about large hail?

I want to thank you beforehand and wish you a pleasant summer! I hope that hail did not do any damage in the area.

Best regards,
Jari Tuovinen
meteorologist
Finnish Meteorological Institute

Figure S2: An example of a hail-survey e-mail (translated into English) sent on 29 June 2009, following the previous day's high likelihood of hail determined from the hail algorithm.

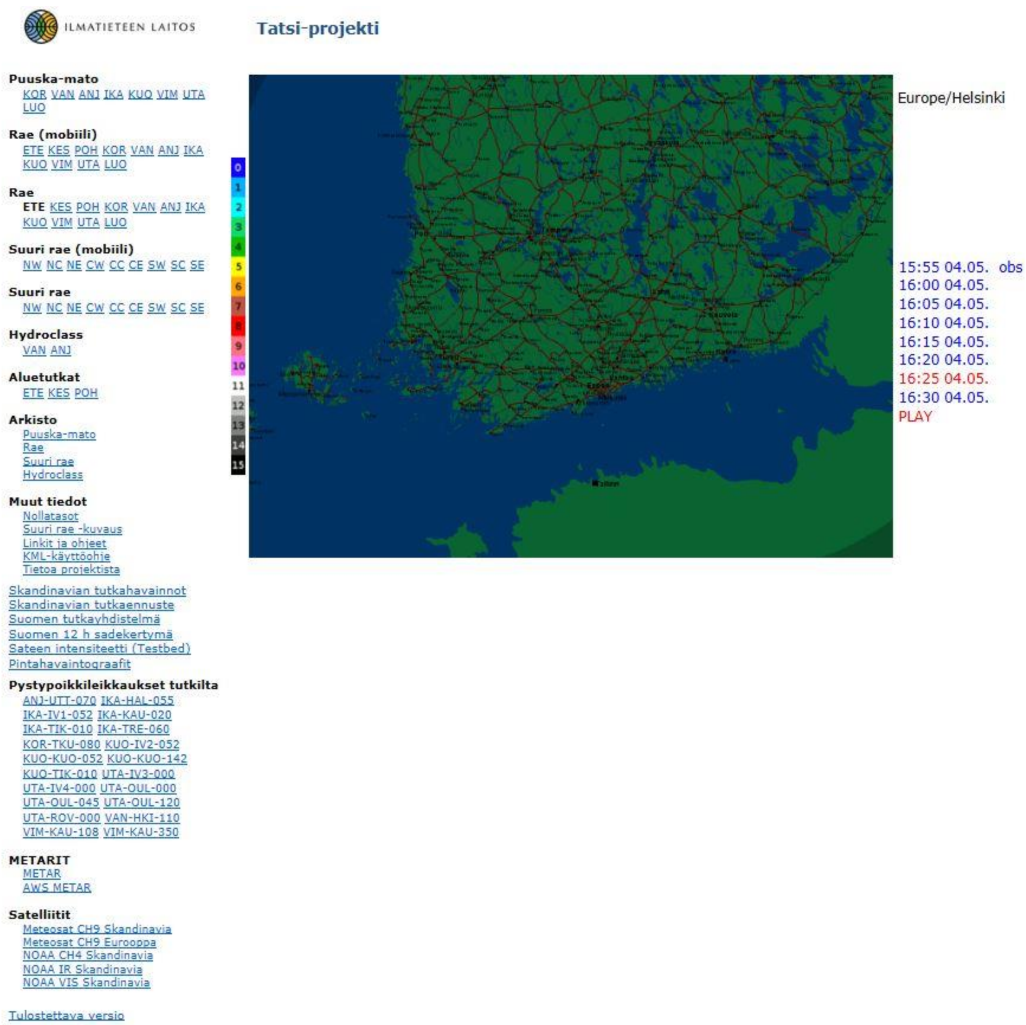

Figure S3: The TATSI survey project's Web page (http://virpo.fmi.fi/tatsi) since 2011 with operational hail algorithm HHI (*Rae* in left panel) and its 15-day archive. Multiple radar and satellite loops, radar cross-sections, and the most recent values of the 0°C isotherm height from the HIRLAM NWP model for each radar location are also available.

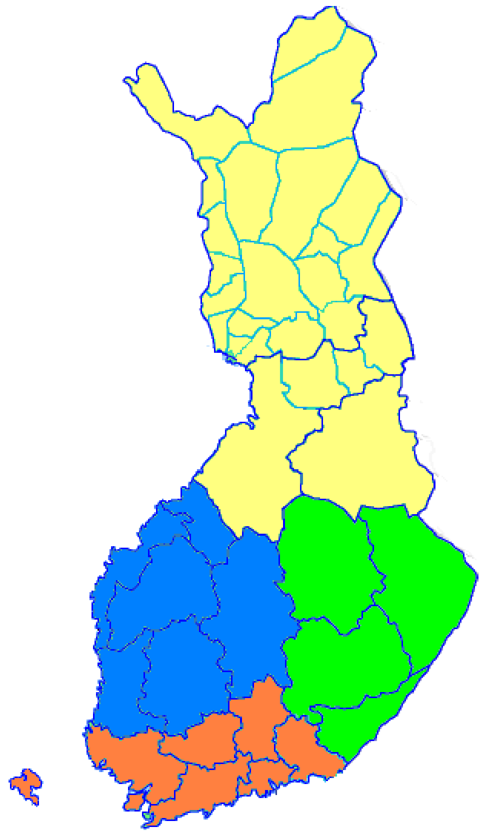

Figure S4: The area distribution of hail-survey project (TATSI) in 2010–2012. During other project years, the whole of Finland served as a single area (2008–2009 and 2015–2017) or was divided into three areas (2013–2014; southern, central and northern). Central was combined from the eastern and western parts. Each group was responsible only for cells moving in their area.