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

JARI-PETTERI TUOVINEN AND HARRI HOHTI

Finnish Meteorological Institute, Helsinki, Finland

DAVID M. SCHULTZ

Centre for Atmospheric Science, School of Earth and Environmental Sciences, and
Centre for Crisis Studies and Mitigation, University of Manchester, Manchester, United Kingdom

(Manuscript received 11 July 2019, in final form 23 January 2020)

ABSTRACT

Collecting hail reports to build a climatology is challenging in a sparsely populated country such as Finland. To expand an existing database, a new approach involving daily verification of a radar- and numerical weather prediction–based hail detection algorithm was trialed during late May–August for the 10-yr period, 2008–17. If the algorithm suggested a high likelihood of hail from each identified convective cell in specified locations, then an email survey was sent to people and businesses in these locations. Telephone calls were also used occasionally. Starting from 2010, the experiment was expanded to include trained storm spotters performing the surveys (project called TATSI). All the received hail reports were documented (severe or ≥2 cm, and nonsevere, excluding graupel), giving a more complete depiction of hail occurrence in Finland. In combination with reports from the general public, news, and social media, our hail survey resulted in a 292% increase in recorded severe hail days and a 414% increase in observed severe hail cases compared to a climatological study (1930–2006). More than 2200 email surveys were sent, and responses to these surveys accounted for 53% of Finland's severe hail cases during 2008–17. Most of the 2200 emails were sent into rural locations with low population density. These additional hail reports allowed problems with the initial radar-based hail detection algorithm to be identified, leading to the introduction of a new hail index in 2009 with improved detection and nowcasting of severe hail. This study shows a way to collect hail reports in a sparsely populated country to mitigate underreporting and population biases.

1. Introduction

A database of severe weather events is essential to study the climatology of severe weather and to improve the accuracy of forecasts and warnings (e.g., Blair et al. 2017; Púčik et al. 2019). With such a database, the climatology of such events and the regions at risk for severe weather can be depicted (e.g., Kelly et al. 1985; Allen and Tippett 2015; Blair et al. 2017; Púčik et al. 2019). The climatology can also be used to understand the physical processes behind the events (e.g., Groenemeijer and van Delden 2007). Finally, a database of weather events can be used to verify automated or human-generated forecasts and warnings and hence lead to improvements (Tuovinen et al. 2015). Consequently, such databases are valuable to national hydrometeorological services.

Particular difficulties exist, however, in building a database in a sparsely populated country. For example, the average population density of Finland is 16 people km^{-2}, equivalent to the population density of Colorado (both including densely populated metropolitan regions such as Helsinki and Denver), whereas Lapland in northern Finland has a population density of only 2 people km^{-2}, equivalent to the population density of Wyoming. The small spatial and temporal scales of severe weather events mean that many severe weather events may go unobserved, and climatologies of severe weather may be biased near cities. Finally, few quantitative measures of intensity of the severe weather events may be available to compare the observed event to the forecasts.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WAF-D-19-0142.s1.

Corresponding author: Jari-Petteri Tuovinen, jari.tuovinen@fmi.fi

DOI: 10.1175/WAF-D-19-0142.1

Given these limitations, multiple approaches to collect observed severe weather events are needed (e.g., Changnon 1969, 1999). Newspaper articles and insurance records from property losses can be used to build a database of severe weather events (e.g., Changnon 1977; Kahraman et al. 2016). Detailed records may be available in areas that are heavily agricultural or urbanized, but they may be limited outside of the growing season, especially in Finland (60°–70°N) where the growing season may be only a few months long. Tuovinen et al. (2009) used six different approaches to collect severe hail reports in Finland (Table 1). [In this article, *severe hail* is defined as hail greater than or equal to 1.91 cm (0.75 in.) in diameter, in accordance with the criterion in the United States before 5 January 2010 and current criterion in Europe.] Cases were collected using microfilmed newspapers and newspapers' Internet databases. Other approaches included reports from a storm spotter network, reports through the Finnish Meteorological Institute's (FMI's) hail-reporting form (the general public), and different social media sources (e.g., Twitter, Instagram, Facebook, blogs). Adding these new approaches has increased the number of severe hail reports substantially since the early 2000s (Tables 1 and 2).

Unofficial observations from voluntary observer networks can also be used. The advantage of a voluntary observer group is that the observers are easy to train and the quality of observations is higher than among the general public, but the network is usually sparse and covers only a small area. The European Severe Storms Laboratory developed the web-based European Severe Weather Database in 2004 as a repository for anyone to upload severe weather reports (Dotzek et al. 2009; Groenemeijer et al. 2017). From its inception in 2009 through December 2018, 39 537 severe hail reports have been collected (Púčik et al. 2019), helping to expand forecast verification and research opportunities for severe convective storms in Europe.

Other modern approaches to building a severe weather database can be employed, with smartphones providing opportunities. Hail photos are useful if the time and location match with plausible events. For example, a severe hailstorm in the United Kingdom on 1 July 2015 was studied in detail with the help of private automatic weather stations owned by members of the public (Clark et al. 2018). After data filtering and comparison with the official Met Office's automatic weather stations, information about the storm structure and development was obtained with greater detail than was possible from traditional operational datasets. In another example, Hyvärinen and Saltikoff (2010) compared hail photos in Finland from the photo-sharing service Flickr to output from algorithms for hail detection from dual-polarimetric radar, showing the possible utility of such nontraditional online data. Finally, crowd-sourcing applications, like mobile weather apps, are the newest approach allowing anyone to send their real-time weather observations (e.g., hail) from smartphones (Elmore et al. 2014; Barras et al. 2019). For example, FMI's weather app has been tested since July 2017, but was not used in this study. The preliminary results from FMI's app are promising.

For the purposes of this project, we concentrate on severe hail. We aim to increase the number of severe hail reports through an email survey approach. We also aim to show the effectiveness of using an operational hail detection algorithm on radar data and numerical weather prediction (NWP) model output when collecting reports of severe hail from sparsely populated areas. We use the output from the hail algorithm and the email surveys to verify and improve the hail algorithm's performance.

In section 2, we introduce the main tools of this study: the hail detection algorithm, the hail detection criteria, and the execution of the survey. More details about the survey project's methods (e.g., finding and contacting a possible reporter in the hail area, adding information to the database, and the organization and management of the project) are described in the in the online supplemental material. Section 3 illustrates how the hail statistics from this approach (2008–17) compare to the hail statistics from a previous climatological study (1930–2006; Tuovinen et al. 2009). We also discuss season-to-season differences between cloud-to-ground lightning, summers' warmth, and hail. Section 4 discusses potential sources of bias in the hail reports, especially population bias and our survey's ability to limit underreporting. Section 5 concludes this study.

2. The hail detection and hail reporting process

This section describes the processes for hail detection and reporting in Finland since 2008. First, we describe the algorithm that combines output from radar and NWP model to indicate the chance of hail occurrence. Then, we describe the survey's method of communicating with those in areas suspected of receiving hail and compare our approach to similar studies. Finally, we discuss how the algorithm was improved over time due to ongoing verification efforts.

a. Hail detection algorithm

In 2004–07, an experiment was conducted using an algorithm derived from radar data and NWP model output to detect the occurrence of hail during summer 2006 (Holleman 2001, 2003). The probability of hail

TABLE 1. Different approaches used to collect severe hail cases. The most recent method of people sending weather observations via their mobile devices is currently in a testing phase, and these data are not used in this hail project.

Source	Description	Timeline
Newspapers (microfilms)	Numerous national newspapers browsed through the microfilms of National Library	Summer months between 1930 and 1990
Newspapers (Internet database)	Key words used to search hail news or articles from numerous newspapers	Between 1993 and 2008
Storm spotters hail database	A group of storm spotters has provided hail and other interesting weather observations	Since 2002
Newspaper viewers' and other social media photos, blogs, etc.	Photos of hail have been obtained by searching through web sites that host user's imagery and are clearly consistent with the radar and hail algorithm	Since 2007
Online hail reporting form	Anyone can fill in their own hail observation via this form on FMI's web page (http://www.fmi.fi/palaute/rae.html)	Since 2007
Hail algorithm and email contacts	Weather radar–based algorithm has been monitored throughout summers; confirmation of possible hail signal via email contacts close to the suspected area	Since 2008
FMI's observation mobile app	Anyone with the app on a mobile device can send weather-related observations that can be plotted on a map in real time	Since July 2017

(POH), originally proposed by Waldvogel et al. (1979), was tested in an operational setting by the warning forecasters at FMI. Holleman's (2001) equation for POH is as follows:

$$\mathrm{POH}(\Delta h) = 0.319 + 0.133\Delta h, \qquad (1)$$

where

$$\Delta h = \mathrm{H45} - \mathrm{H0},$$

and Δh denotes the altitude difference between the maximum of the H45 (45-dBZ echo) and the model H0 (0°C isotherm) for each cell.

The email surveys were activated if the POH reached 80%–100% for at least 15 min within a convective cell, a criterion found to be a relatively good indicator of marginally severe (1.5–2 cm) or severe hail cases. After verifying hail events from the first few summers, however, POH showed problems. POH overestimated the actual hail occurrence during warm conditions (i.e., melting level well above long-time average for summer), whereas POH underestimated the actual hail occurrence during cold conditions (i.e., melting level well below the long-time average for summer). Overestimation of hail occurrence by POH was also noticed by Betschart and Hering (2012) in Switzerland. The average maximum hail size is also smaller in high latitudes, likely a result of the weaker instability and resulting updrafts [e.g., Figs. 4e,f in Pistotnik et al. (2016)]. Furthermore, the algorithm was unable to discriminate between convective cells producing 1- and 3-in. diameter hail, something warning forecasters needed. This point was also considered to be important when changes to the hail algorithm were conducted.

After following the progress of the potential hail cell and if the criteria discussed above were fulfilled, then the next step was pinpointing the suspected area and searching for possible receivers. This process is described in the supplemental material.

b. The survey's main approach for contacting recipients

After finding suitable recipients from the suspected hail area, the next step was to contact them and ask them about any possible hail. Email surveys were the primary tool of gathering information from the detected hail area and communicating with recipients in our project. During summers 2008 and 2009, the first author was sending individual emails to locations suspected of having received severe hail, as an

TABLE 2. Statistics of periods of hail seasons combining all sources before the start of the hail survey project.

Period	Avg No.of hail days per season	Avg No. of hail cases per season	Avg No. of severe hail days per season	Avg No. of severe hail cases per season	Total No. of severe hail cases
1930–2006	—	—	5[a]	10[a]	240
2004–07	22[b]	65[b]	7[b]	13[b]	—

[a] 77-yr average (Tuovinen et al. 2009).
[b] 4-yr average.

experiment to test the viability of this email approach. To broaden the participation and to ease the workload on the first author, in the beginning of spring 2010, 30 trained storm spotters were invited to participate in the experiment.

A comparable way to increase hail observations by using a radar-based hail detection algorithm took place at the NOAA/National Severe Storms Laboratory in Norman, Oklahoma, during the Severe Hazards Analysis and Verification Experiment (SHAVE), which started in the early 2000s (Ortega et al. 2009; Wilson et al. 2009; Meyer et al. 2010). The goals of SHAVE were to develop a high-resolution database of severe weather reports alongside the coarser *Storm Data* database of the National Weather Service, to evaluate the skill of radar-derived parameters and to improve the forecasting of severe hail events. This mostly student-led organization made phone calls to businesses or residences in the suspected hail area. The project started within a few states in the Central Plains of the United States, but later enlarged the area and included other severe weather phenomena (Smith et al. 2006; Ortega et al. 2009). More recently, several crowd-sourcing studies have begun collecting data on hail. Elmore et al. (2014), Barras et al. (2019), and Friedrich et al. (2019) were all focused on mobile apps and social media sources. Such passive crowd-sourcing tools can be a great addition during hail outbreaks in urban areas, but their performance for isolated convective cells in remote, sparsely populated areas (e.g., much of Finland) may be limited. This limited performance can be addressed through our active email-based approach that specifically targets individuals, communities, and businesses in the area of suspected hail.

Our study has the closest resemblance to SHAVE, but, instead of using telephone calls, email was our primary tool for the following reasons. First, although 97% of Finns own a mobile phone, this device is not typically fixed to a specific location (Kuusela et al. 2007). Not knowing where the receiver answers the call can make mobile phones unsuitable for surveys. Second, people are reluctant to answer calls from an unknown number because quite often these mystery callers may be salespeople. Had we used phone surveys, our callers could have easily been misidentified by receivers as salespeople. Third, emails remain in mailboxes until moved, and answering email takes slightly more time than a phone call. The written form of an email and the extra time needed to compose an email rather than talk on the phone tend to make people think more, write with more care in a rational order, and include more details, although the extra time needed may reduce the willingness to participate. Last, hundreds of phone calls cost more than the same number of emails. In any case, project volunteers would have been reluctant to make these calls and disrupt people's day. For these reasons, writing emails was an easier way to communicate, albeit more time consuming.

We also highlight the importance of contacting community organizations, as this factor also differentiates the present study from previous ones. Instead of contacting just individuals, we contacted small community organizations, and especially village associations, which are real hubs of rural areas. Sharing information with each other can be quick and can reach almost everyone in the community. This type of contact can yield much greater amount of information of the event, and in the best scenario, several observations from different parts of the community.

c. Revising the hail algorithm in 2009 and 2017

In 2009, improvements to the algorithm were made, and the probabilistic approach to predicting hail occurrence was rejected. To address these weaknesses, a new hail index (HHI) was created so that the strongest hail cells could be more easily identified:

$$\mathrm{HHI} = \mathrm{POH}(\Delta h)/10. \tag{2}$$

The HHI index ranges from 0 to 15, is dimensionless, and is still based on Δh, the height difference between the 45-dBZ echo top and the 0°C isotherm. Values greater than 10 represent stronger hail cells with a higher likelihood for severe hail (e.g., a value of 10 is equivalent to a height difference of 5.5 km, whereas a value of 15 is equivalent to a height difference of 9 km). Despite these changes, HHI, as its predecessor POH, had problems in abnormally warm or cold conditions.

Warm periods reduced the performance of the hail algorithm as the height of the 0°C isotherm reached well above the long-term average. For instance, several above-normal warm periods were experienced in July 2010, which was the warmest July on record in many locations (around 5°C above the monthly average); we noticed that for several days all HHI output values less than or equal to 10 came back as nonhail reports (hail suspected but only rain observed). Lengthy periods of warmth were experienced again in 2014, but the hail survey instructions had already been updated by then. Since 2012, the updated instructions stated that the height of the 0°C isotherm should be checked first, and, if it was close to 4 km, hail survey emails should not be sent before HHI $\geq$ 10 was reached within a cell for at least 15 min. This effort had a clear impact on hail reporting. After the change in instructions, only 30% of returned emails (29 reports) in 2014 reported no hail compared to 45% (55 reports) in 2010 before the change (Table 3).

Problems with the POH algorithm were identified during below-normal temperatures (i.e., anomalously low 0°C isotherm heights), too. Hail was observed with low POH values (10%–30%) when the 0°C isotherm stayed well below (at least 1500 m below the mean of 3000 m) the long-time average levels during cold periods. Even missed cases (small hail but no POH output probability) occurred, indicating limited melting of hail.

We further improved HHI in 2017. When the height of the 0°C isotherm was above normal (e.g., 500 m above the mean of around 3000 m), index values were increased by +1 (at least 3.5 km) or +2 (at least 4 km), as a consequence of more hail melting potential. When the height of the 0°C isotherm was below normal (e.g., 1300 m below the mean of around 3000 m), index values were instead lowered by −1 (lower than 1.7 km) or −2 (lower than 1.2 km), respectively. This tuned hail index (THI) should work better than HHI in abnormal circumstances.

Tables 3 and 4 have different numbers of severe hail cases for two reasons. First, Table 3 includes only hail survey statistics (contacts with email or telephone), whereas Table 4 is a summary over all possible hail sources combined (including media, social media, etc.). Second, we followed the same procedure used in Tuovinen et al. (2009) when verifying one case from another and filing to the database (15-min and 20-km separation for severe hail). For example, there were 122 severe hail observations in 2010, but only 91 after following the verification procedure (e.g., combining severe hail cases from multiple sources and removing duplicate observations). Most of the hail-producing cells in Finland moved over sparsely populated areas and therefore our focus was obtaining reports from the area in which each hail cell had the maximum value output from the hail algorithm. The verification of reports is performed with care to ensure the maximum reported hail size is archived to the database. This procedure decreased the number of severe hail cases by about 10%, but also lowered the impact of irrational reporting of hail size over major cities. People in the same neighborhood can report very different hail sizes occurring from the same hailstorm. This behavior occurs because of various methods defining hail size (e.g., some measure hail with ruler, many compare hailstones to coins, others estimate size from indoors) leading to inconclusive data reporting.

3. Statistics and factors contributing to hail presence

In this section, we focus on statistics from the hail survey project, review the seasonal aspects of cloud-to-ground

TABLE 3. Statistics of email survey responses during the 2008–2017 hail survey project. Table includes the number of volunteers taking part each year and a count of the types and numbers of communications (number of sent and replied emails and phone calls) between project members and recipients. Similarly, information of hail size is divided into four categories (no hail, hail unknown, small hail, and severe hail). The sum and annual average values are highlighted using bold and bold italic text, respectively.

	No. of participants	No. of sent emails	No. of replied emails	No. of unreplied emails	Ratio of sent/returned emails (%)	No. of phone calls	No. of reports of hail not obs	No. of reports of hail size unknown	No. of reports of small hail obs	No. of reports of severe hail obs
2008	1	125	95	30	76	2	12	8	50	25
2009	1	109	78	31	71.6	2	13	7	41	17
2010	12	842	377	465	44.8	10	171	22	128	56
2011	14	368	150	218	40.8	4	39	8	65	38
2012	14	154	93	61	60.4	0	47	5	34	7
2013	12	172	86	86	50.0	2	23	5	34	24
2014	8	186	96	90	51.6	3	29	6	36	25
2015	5	91	44	47	48.4	0	14	3	22	5
2016	1	104	49	55	47.1	6	12	3	20	14
2017	1	93	42	51	45.2	0	10	4	18	10
Sum		**2244**	**1110**	**1134**		**29**	**370**	**71**	**448**	**221**
Annual avg		***224.4***	***111.0***	***113.4***	***49.5***	***2.9***	***37.0***	***7.1***	***44.8***	***22.1***

TABLE 4. Annual hail and cloud-to-ground lightning statistics of Finland for 2008–17. Numbers include all possible sources of hail (also hail survey project). These hail statistics are reported after each season to the media and public. Bold text highlights the sum and annual average values.

	No. of hail cases obtained	No. of severe hail cases	No. of hail days	No. of severe hail days	Total No. of cloud-to-ground flashes
2008	182	44	42	20	63 333
2009	154	28	40	10	53 561
2010	391	91	47	22	167 713
2011	304	40	52	17	180 167
2012	235	11	49	6	78 000
2013	291	52	47	15	118 490
2014	461	86	58	28	201 230
2015	303	12	52	5	29 810
2016	433	33	54	15	113 300
2017	415	17	33	8	30 300
Sum	**3169**	**414**	**474**	**146**	**1 035 904**
Annual avg	**316.9**	**41.4**	**47.4**	**14.6**	**103 590**

lightning and very warm days, and explore the role of the height of the melting level and attempt to link these factors to hail occurrence in Finland.

a. *Statistics of hail survey*

The average number of emails sent each year was 224, ranging from 91 to 842 (Table 3). Altogether, 2244 emails were sent in 2008–17, of which 1110 (49.5%) received replies. Given that the typical email response rate in surveys is around 30% (Sheehan 2001), our results are above average. Overall, 740 email replies out of 1110 (67%) indicated the presence of hail and 669 (60%) indicated hail size. This ratio of confirmed hail is high, underlining the high-quality guidance received from the algorithm. Although 20% of replies were confirmed as severe hail, this was a relatively high number considering the shortcomings discovered during the survey (e.g., hail algorithm's poor performance during abnormally warm conditions). Also, receiving a vast amount of small-hail observations (448 out of 1110; 40%) helped maintaining this aspect of severe weather statistics from Finland. Even the 370 replies out of 1110 (33%) that indicated no hail was observed were important for verification because they revealed weaknesses of the hail algorithm, leading to improvements. If the statistics of Tables 3 and 4 are compared (survey versus all approaches), our survey gathered 23% of all hail observations and 53% of all severe hail observations in Finland during 2008–17.

The number of days with recorded hail or severe hail in Finland has increased since implementation of our new approach. During 1930–2006 without the hail algorithm's help, 5 severe hail days and 10 severe hail cases were recorded on average each season over Finland (Tuovinen et al. 2009). Help from other sources (e.g., online hail reporting form, social media photos) increased these numbers during 2004–07 to 7 recorded severe hail days and 13 observed severe hail cases just before the hail survey project started (Table 2). Since 2008, our project along with other approaches (Table 1) has tripled the average number of recorded severe hail days (14.6) and quadrupled the average number of observed severe hail cases (41.4) compared to the 1930–2006 statistics (Table 4). The cumulative number of observed severe hail cases have increased 79% compared to the climatological statistics (i.e., 2007–17 had 430 severe hail cases). There is no comparison for the number of hail days from the climatological statistics, but statistics show a 10-yr average of 47.4 days, which is more than triple the average of recorded severe hail days. For comparison, in the United States, the annual average number of days with hail damage to property is 123 (Changnon et al. 2009). Allen and Tippett (2015) showed that from the 1960s and 1970s to the 2000s the annual number of days of at least 0.75-in. hail reported somewhere in the United States has grown from around 150 to closer to 240.

b. *Association between cloud-to-ground lightning, summers' warmth, and hail*

We next consider the interannual variability (e.g., frequency and seasonal variation) of convective storms. As a proxy for the frequency of convective storms, we use the total number of cloud-to-ground lightning flashes over Finland during the calendar year (around 97% occur between May and September). Severe convective storms tend to produce more lightning than nonsevere ones (e.g., Changnon 1992; Carey and Rutledge 1998; Williams 2001). The current lightning detectors have been operating since 1998 and therefore all study years can be treated equally.

An annual average of 118 400 cloud-to-ground flashes occurs in Finland, derived from a most recent 20-yr average (1998–2017; Tuomi and Mäkelä 2008; Mäkelä et al. 2017).

The number of flashes varied during our study from a minimum of 29 800 in 2015 to a maximum of 201 200 in 2014 (Table 4). The 2015 minimum value is the lowest observed during the lightning-measurement era in Finland (1960–2017). This almost sevenfold difference between the maximum and minimum values indicates vast year-to-year variations.

Fewer than 30% of severe hail cases were received during the lowest five lightning seasons (2008, 2009, 2012, 2015, and 2017) where the average cloud-to-ground lighting amount was 57% below the 20-yr average (Tables 3 and 4). The average number of recorded severe hail days (9.8) was also 33% lower compared to the whole study (14.6). Clearly, these seasons were less supportive of deep moist convection and severe hail. These five summers were all cooler and shorter than average, concurrently lacking any warm periods (5-yr average of 23.4 very warm days of maximum temperature reaching at least 25.1°C or 77.2°F compared to the long-time average of 37 days; https://ilmatieteenlaitos.fi/helletilastot). Cold periods, especially during maximum daytime heating in late June and early July, can effectively hinder the peak period of severe hail in Finland.

In contrast, the number of cloud-to-ground flashes was 42%–70% above the 20-yr average during the 2010, 2011, and 2014 seasons, in which our hail survey gathered 119 out of 221 severe hail reports (53.8%; Table 3). Similarly, 217 out of 414 severe hail reports (52%; Table 4) occurred during these seasons nationwide. Even the number of recorded severe hail days (22.3) was 53% above average. These three summers were all warm and long-lasting with many warm periods (3-yr average of 51 very warm days). These warm periods, occasionally associated with unstable air masses, increased the presence of lighting-rich deep moist convection and led to higher cloud-to-ground lightning activity (3-yr average of 183 000 strikes compared to the 10-yr average of 104 000 strikes; Table 4) in Finland.

Our hail survey statistics imply that lightning-rich warmer-than-average summers tend to be associated with an increased likelihood of severe hail. Although there seems to be an association between severe hail days and total cloud-to-ground lightning in Finland (correlation coefficient of 0.80), further evidence and statistics (e.g., role of different convective modes with lightning) is needed. After all, hail and lightning have somewhat different ingredients associated with them (e.g., Johns and Doswell 1992; Carey and Rutledge 1998). Even so, the number of hail days seems to stay at a certain high level more systematically and have less variance (between 33 and 58 days; Table 4) from one season to another regardless of summers' warmth, length, or lightning amounts.

4. Potential sources of bias

Next, we list different biases in detection and reporting methods that affect the survey's reliability.

a. *Reporting biases*

Human reporting raises a number of potential sources of bias. All contacted receivers were taken as trustworthy, and no dishonesty was detected. Verification of all received cases ensured this, but also revealed fallibility of some reports. First, a total of 20 phony reports were sent via FMI's online hail reporting form. All of them were easily identified as fictitious because senders used unknown email addresses or provided nonsense information. Bogus reports also exclusively reported very large hail sizes (8–10 cm or 3.2–4.0 in., or larger). Second, hail reports that were returned several weeks after the event were often a day or two off from the actual event date. Third, people may over or underestimate maximum hail size or fail to capture the storm's largest hailstones (Blair and Leighton 2012; Blair et al. 2017), but this was not possible to determine in Finland. Fourth, verifying events that the algorithm missed (i.e., strong convective storms that did not produce hail) was more difficult than verifying severe hail events, unless we received unsolicited reports. Because of our natural curiosity about the most intense storms, verification of suspected cases of severe hail likely was pursued with slightly more vigor than suspected cases of smaller hail. However, missed events were crucial for improving the hail algorithm. Last, our prefilled survey emails with specific information about the possibility of hail in the area (e.g., telling the recipient that our radar-based algorithm indicated high likelihood of hail in their area) might have led the recipient to provide the answer that we were looking for. However, we felt that this minimum level of information was required to obtain the level of information that was required to verify the hail algorithm.

b. *Doppler radar detection biases*

Radar-based measurements have their own issues that affect detectability. There are multiple possible reasons for reflectivity falling under the threshold of the hail index. First, attenuation occurs when the radar beam needs to travel through several convective cells or through an intense hail shaft. Second, beam widening can lead to smoothing of small and intense cell cores. Third, attenuation by the radar dome's wetting during a heavy downpour can reduce the reflectivity by up to 12 dBZ (e.g., Bechini et al. 2010). Fourth, some of the missed detection events by the algorithm can be related to the lowest radar beam overshooting the cells'

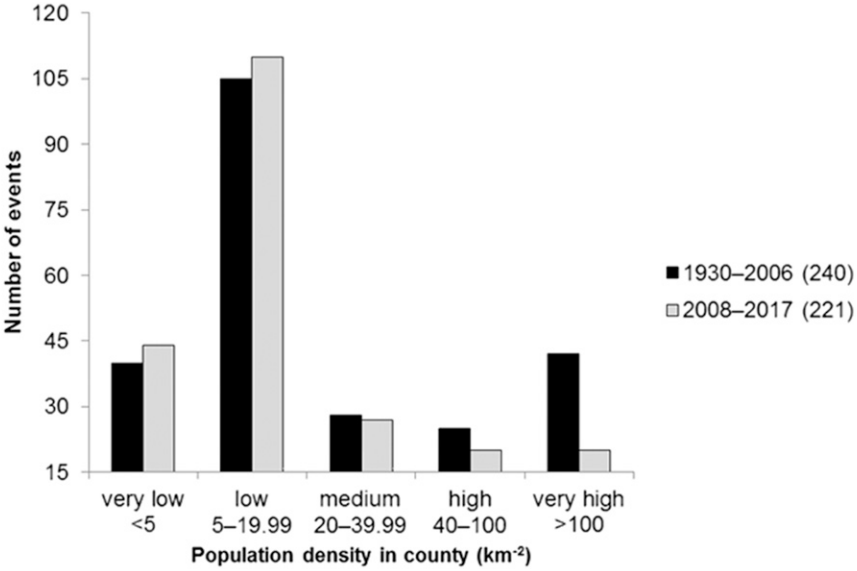

FIG. 1. Number of severe hail cases occurring in counties binned by population density (people km^{-2}) during two hail datasets: severe hail climatology 1930–2006 (black bars with a total of 240 cases, Tuovinen et al. 2009) and hail survey project 2008–17 (gray bars with a total of 221 cases). Counties are divided into five different bins depending on population density. The number of counties valid in 2008 was used in both statistics (415 counties).

maximum reflectivity at long distances or hail occurring with reflectivity < 45 dBZ. Fifth, the distance from the radar becomes problematic when the hail occurs very close to the radar or farther than 150 km away (e.g., Delobbe et al. 2005).

c. Hail algorithm interpretation biases

Making the call to start the hail survey process was not always straightforward, especially for the storm spotters who were not so familiar with the hail algorithm. Sometimes the area highlighted by the hail algorithm grew very large, especially when radar reflectivity values exceeded well over 50 dBZ. Hailfall usually occurs close to the downdraft region of the convective cell, covering only a fraction of the cell's area on the radar display (Foote 1984; Höller et al. 1994). Therefore, pinpointing the region of hailfall within the predicted hail cell was occasionally difficult. Also, on a few occasions, strong winds aloft may have displaced the hail observation up to 7 km from the actual hail algorithm signal (E. Saltikoff 2017, personal communication). Severe hail from long-lasting hailstorms may still fall on the ground 5–10 min after the last high value of the hail index, possibly indicating that this amount of time is needed for the hail to fall from its elevation above ground (Witt et al. 1998).

d. Population bias

Our previous hail climatology for 1930–2006 showed that hail reporting suffered from population bias (Tuovinen et al. 2009). One goal of this present study was to limit population bias through a focus on nonurban areas where underreporting was greatest. Using radar-based detection tools like the hail algorithm should help to minimize population bias and reveal a more unbiased snapshot of hail occurrence (Saltikoff et al. 2010; Cintineo et al. 2012). To confirm this, we compared severe hail locations of the previous and current studies to the population density of each county (Fig. 1). We used the county geographies that were valid in 2008 for both datasets because of a steadily declining trend in the number of counties in Finland as smaller less-populated counties have merged into bigger population centers. Typically, when several small counties join together or are joined to a populous county, the outcome is a medium- or urban-sized county in population density. In the 1990s, there were more than 500 counties, but in 2008 the number was down to 415 and by 2017, only 317 counties were left.

In our hail survey, severe hail cases occurred less often in urban areas and more in low-populated areas compared to the 1930–2006 climatology, suggesting that our program to supplement the reports in sparsely populated areas was working. For example, the two lowest bins in Fig. 1 (population density of 0–19.99 km^{-2}) had 6.2% more cases than in Tuovinen et al. (2009). Naturally, with such small populations in these regions, getting any increase in the number of reports is difficult. The two highest bins (population density over 40 km^{-2})

had 40.3% fewer cases compared to the 1930–2006 climatology. Overall, the statistics indicate that our approach was partly offsetting population bias and yielding more cases from sparsely populated areas.

5. Conclusions

We introduced a new way of collecting hail reports and verifying a radar-based hail detection algorithm that is feasible in a country with a low population density but could have potential even in countries with a high population density. This study started as an experiment in 2008 and was later expanded (since 2010) to include storm spotters (as a project named TATSI). Each participant was instructed about the radar-based hail detection algorithm, the cell-tracking criteria, and the procedures for finding suitable receivers and contacting them during training by the project lead at FMI (i.e., the first author). The participants formed groups and worked in shifts from late May through August, and tracked hail cells using the radar-based hail detection algorithm. This algorithm was originally POH but was replaced by the HHI hail index since 2009. After certain criteria of high likelihood of hail was surpassed (HHI index reaching 8–10 for at least 15 min), participants looked for possible email recipients (e.g., local companies, village associations, libraries) in the location of the suspected hail and contacted them via an email survey requesting any information on hail presence and size. Finally, all replies were then stored in a database.

After verification POH showed problems in detection during anomalously warm or cold weather and the heights of the 0°C isotherm that led to over or underestimating hail presence. In several verified cases, low POH values (10%–30%) produced small hail when the melting level was low compared to the long-term average. In turn, high POH values of 80%–100% turned out to be just rain when the melting level was high compared to the long-term average. Another problem with the POH was its inability to indicate the most severe hail cells.

Therefore, the POH algorithm was modified and, after adaptations, HHI provided a more detailed view of the strongest hail cells during severe and significant hail cases, a valuable improvement among warning forecasters. Still, HHI also encountered difficulties in abnormally warm or cold environments. Further development has improved detectability and the most recent version of the hail algorithm (THI) is responding better to abnormal variations of the melting level and has been in operational use since 2017.

The success of our project can be seen through the large number of received emails, the increased number of severe hail observations compared to past, and the demonstrated ability to reduce the potential for population bias. During 2008–17, 2244 hail survey emails were sent, of which almost half (1110) were answered, generally in a positive tone. Of these 1110, 66.7% of respondents reported hail, with 60.3% indicating hail size. Altogether, 221 cases were severe hail and 448 cases small hail. Reports where no hail was observed were valuable for verifying the performance of the hail algorithm as these cases revealed problems and helped us improve hail detectability. During 2008–17, 3169 hail cases and 414 severe hail cases were received nationwide through all the approaches we employed. Therefore, the hail survey project gathered around 23% of all hail cases, but most importantly, 53% of all severe hail cases in Finland. Compared to the previous study of Tuovinen et al. (2009) during 1930–2006, the present study during 2008–17 showed a 292% increase in the number of recorded severe hail days (14.6) and 414% increase in the number of observed severe hail cases (41.4). The annual average of hail days (47.4) is high despite having short seasons for deep moist convection.

Finally, we received slightly more cases (6%) from low-density population counties and 40% less from high-population counties compared to the previous study of 1930–2006. Since 2008, our new approach has decreased population bias by reaching even the most remote areas in Finland. The power of keeping hail in the interest of public and media should not be underestimated.

Finally, we conclude with recommendations for others considering this approach. Launching a survey-based project operated via email takes some time and effort to function properly. Using email as an approach has its own pros and cons, but is likely the most viable option in locations of low-population density compared to other approaches (e.g., telephone or mobile apps). Email surveys could also be used as a secondary approach in situations where more cases from a specific area are needed and other sources are unreachable or provide insufficient details. Finding recipients from the hail-suspected area could be a laborious task, but forming an efficient way to utilize Internet searches (e.g., linking local companies and cottage renters with street names) should help this task. Instead of focusing on individual citizens as recipients, contacting communities as a whole (like local village associations) yields much more as communities usually cover the majority of the local citizens and information between members travels fast. Targeted efforts that reach out to outdoor groups and fishing or hunting lodges could also be an effective way to gain coverage in areas that these people frequent, in contrast to the typical observers. After a while, operating with a prefilled inquiry draft and organizing email communication becomes more routine and starts benefiting the project with additional reports. In our case, one person was not able to react to all potential hail signals during busy periods, which occasionally led to prioritizing and underreporting. In turn, a bigger group of people allowed more potential hail cells to be studied, but brought their own challenges to project operation and management. We estimate that around five people working closely together would have been the best outcome in Finland.

Acknowledgments. We thank all the storm spotters and collaborators who contributed to the hail email surveys during TATSI: Jari Ylioja, Esa Palmi, Panu Lahtinen, Petri Tynkkynen, Janne Kommonen, Eero Karvinen, Teemu Mäntynen, Jani Poikela, Mats Kommonen, Rami Saarikorpi, Anne Pöyhönen, Kimmo Venäläinen, Toni Maurinen, Jussi Karhunen, Ilmo Nurminen, Antti Kiiskinen, Joni Hiidensilta, Esa Tuunanen, Petri Pietikäinen, Mikko Lehtovirta, Jaakko Haapala, Ari-Juhani Punkka, and Pauli Jokinen. We thank Elena Saltikoff for help and Mikko Parviainen for computer support. We thank the anonymous reviewers for their comments that improved the manuscript. Schultz was partially funded by Vaisala Oyj (2006–10) and the Natural Environment Research Council through Grant NE/N003918/1 (2015–19).

REFERENCES

Allen, J. T., and M. K. Tippett, 2015: The characteristics of United States hail reports: 1955–2014. *Electron. J. Severe Storms Meteor.*, **10** (3), https://www.ejssm.org/ojs/index.php/ejssm/article/viewArticle/149.

Barras, H., A. Hering, A. Martynov, P. A. Noti, U. Germann, and O. Martius, 2019: Experiences with >50,000 crowd-sourced hail reports in Switzerland. *Bull. Amer. Meteor. Soc.*, **100**, 1429–1440, https://doi.org/10.1175/BAMS-D-18-0090.1.

Bechini, R., V. Chandrasekar, R. Cremonini, and S. Lim, 2010: Radome attenuation at X-band radar operations. *Proc. Sixth European Conf. on Radar in Meteorology and Hydrology*, Sibiu, Romania, ERAD, P15.1.

Betschart, M., and A. Hering, 2012: Automatic hail detection at MeteoSwiss–Verification of the radar-based hail detection algorithms POH, MESHS and HAIL. Arbeitsbericht MeteoSchweiz 238, 70 pp., https://www.meteoschweiz.admin.ch/content/dam/meteoswiss/en/Ungebundene-Seiten/Publikationen/Fachberichte/doc/ab238.pdf.

Blair, S. F., and J. W. Leighton, 2012: Creating high-resolution hail datasets using social media and post-storm ground surveys. *Electron. J. Oper. Meteor.*, **13**, 32–45, http://nwafiles.nwas.org/ej/pdf/2012-EJ3.pdf.

——, and Coauthors, 2017: High-resolution hail observations: Implications for NWS warning operations. *Wea. Forecasting*, **32**, 1101–1119, https://doi.org/10.1175/WAF-D-16-0203.1.

Carey, L. D., and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *J. Geophys. Res.*, **103**, 13 979–14 000, https://doi.org/10.1029/97JD02626.

Changnon, S. A., 1969: Hail measurement techniques for evaluating suppression projects. *J. Appl. Meteor.*, **8**, 596–603, https://doi.org/10.1175/1520-0450(1969)008<0596:HMTFES>2.0.CO;2.

——, 1977: The scales of hail. *J. Appl. Meteor.*, **16**, 626–648, https://doi.org/10.1175/1520-0450(1977)016<0626:TSOH>2.0.CO;2.

——, 1992: Temporal and spatial relations between hail and lightning. *J. Appl. Meteor.*, **31**, 587–604, https://doi.org/10.1175/1520-0450(1992)031<0587:TASRBH>2.0.CO;2.

——, 1999: Data and approaches for determining hail risk in the contiguous United States. *J. Appl. Meteor.*, **38**, 1730–1739, https://doi.org/10.1175/1520-0450(1999)038<1730:DAAFDH>2.0.CO;2.

——, D. Changnon, and S. D. Hilberg, 2009: Hailstorms across the Nation. *An Atlas about Hail and Its Damages*, Illinois State Water Survey, 92 pp.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea. Forecasting*, **27**, 1235–1248, https://doi.org/10.1175/WAF-D-11-00151.1.

Clark, M. R., J. C. D. Webb, and P. J. Kirk, 2018: Fine-scale analysis of a severe hailstorm using crowd-sourced and conventional observations. *Meteor. Appl.*, **25**, 472–492, https://doi.org/10.1002/met.1715.

Delobbe, L., I. Holleman, D. Dehenauw, and J. Neméghaire, 2005: Verification of radar-based hail detection product. Preprints, *WWRP Symp. on Nowcasting and Very Short Range Forecasting (WSN05)*, Toulouse, France, WWRP/WMO, P8.07.

Dotzek, N., P. Groenemaijer, B. Feuerstein, and A. M. Holzer, 2009: Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. *Atmos. Res.*, **93**, 575–586, https://doi.org/10.1016/j.atmosres.2008.10.020.

Elmore, K. L., Z. L. Flamig, V. Lakshmanan, B. T. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz, 2014: mPING: Crowd-sourcing weather reports for research. *Bull. Amer. Meteor. Soc.*, **95**, 1335–1342, https://doi.org/10.1175/BAMS-D-13-00014.1.

Foote, B. G., 1984: A study of hail growth utilizing observed storm conditions. *J. Climate Appl. Meteor.*, **23**, 84–101, https://doi.org/10.1175/1520-0450(1984)023<0084:ASOHGU>2.0.CO;2.

Friedrich, K., and Coauthors, 2019: CHAT: The Colorado Hail Accumulation from Thunderstorms project. *Bull. Amer. Meteor. Soc.*, **100**, 459–471, https://doi.org/10.1175/BAMS-D-16-0277.1.

Groenemeijer, P., and A. van Delden, 2007: Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. *Atmos. Res.*, **83**, 473–487, https://doi.org/10.1016/j.atmosres.2005.08.006.

——, and Coauthors, 2017: Severe convective storms in Europe: Ten years of research and education at the European Severe Storms Laboratory. *Bull. Amer. Meteor. Soc.*, **98**, 2641–2651, https://doi.org/10.1175/BAMS-D-16-0067.1.

Holleman, I., 2001: Hail detection using single-polarization radar. Scientific Rep. KNMI WR-2001-01, 72 pp, http://bibliotheek.knmi.nl/knmipubWR/WR2001-01.pdf.

——, 2003: Biennial Scientific Report 2001–2002. KNMI Meteorological Research, 73 pp., http://bibliotheek.knmi.nl/biennial-triennial/obs&modbiennial01-02.pdf.

Höller, H., V. N. Bringi, J. Hubbert, M. Hagen, and P. F. Meischner, 1994: Life cycle and precipitation formation in a hybrid-type hailstorm revealed by polarimetric and Doppler radar measurements. *J. Atmos. Sci.*, **51**, 2500–2522, https://doi.org/10.1175/1520-0469(1994)051<2500:LCAPFI>2.0.CO;2.

Hyvärinen, O., and E. Saltikoff, 2010: Social media as a source of meteorological observations. *Mon. Wea. Rev.*, **138**, 3175–3184, https://doi.org/10.1175/2010MWR3270.1.

Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612, https://doi.org/10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2.

Kahraman, A., Ş. Tilev-Tanriover, M. Kadioglu, D. M. Schultz, and P. M. Markowski, 2016: Severe hail climatology of Turkey. *Mon. Wea. Rev.*, **144**, 337–346, https://doi.org/10.1175/MWR-D-15-0337.1.

Kelly, D. L., J. T. Schaefer, and C. A. Doswell III, 1985: Climatology of nontornadic severe thunderstorm events in the United States. *Mon. Wea. Rev.*, **113**, 1997–2014, https://doi.org/10.1175/1520-0493(1985)113<1997:CONSTE>2.0.CO;2.

Kuusela, V., M. Callegaro, and V. Vehovar, 2007: The influence of mobile telephones on telephone surveys. *Advances in Telephone Survey Methodology*, J. M. Lepkowski et al., Eds., John Wiley & Sons, 87–112.

Mäkelä, A., T. Laurila, J. Haapalainen, and T. Halabi, 2017: Salamahavainnot 2016—Lightning observations in Finland, 2016. FMI publications, Tech. Rep. 2017:8, 51 pp., http://hdl.handle.net/10138/231177.

Meyer, T. C., J. M. Lamers, and K. L. Ortega, 2010: Comparing radar signatures to high resolution hail reports. *26th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 396, https://ams.confex.com/ams/90annual/techprogram/paper_163028.htm.

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: The Severe Hazards Analysis and Verification Experiment. *Bull. Amer. Meteor. Soc.*, **90**, 1519–1530, https://doi.org/10.1175/2009BAMS2815.1.

Pistotnik, G., P. Groenemeijer, and R. Sausen, 2016: Validation of convective parameters in MPI-ESM decadal hindcasts (1971–2012) against ERA-Interim reanalyses. *Meteor. Z.*, **25**, 753–766, https://doi.org/10.1127/metz/2016/0649.

Púčik, T., C. Castellano, P. Groenemeijer, T. Kühne, A. T. Rädler, B. Antonescu, and E. Faust, 2019: Large hail incidence and its economic and societal impacts across Europe. *Mon. Wea. Rev.*, **147**, 3901–3916, https://doi.org/10.1175/MWR-D-19-0204.1.

Saltikoff, E., J.-P. Tuovinen, T. Kuitunen, H. Hohti, and J. Kotro, 2010: A climatological comparison of radar and ground observations of hail in Finland. *J. Appl. Meteor. Climatol.*, **49**, 101–114, https://doi.org/10.1175/2009JAMC2116.1.

Sheehan, K. B., 2001: E-mail survey response rates: A review. *J. Comput.-Mediat. Commun.*, **6**, JCMC621, https://doi.org/10.1111/j.1083-6101.2001.tb00117.x.

Smith, T. M., K. L. Ortega, K. A. Scharfenberg, K. L. Manross, and A. Witt, 2006: The Severe Hail Verification Experiment. *23rd Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 5.3, https://ams.confex.com/ams/23SLS/techprogram/paper_115436.htm.

Tuomi, T. J., and A. Mäkelä, 2008: Thunderstorm climate of Finland 1998–2007. *Geophysica*, **44**, 67–80.

Tuovinen, J.-P., J. Rauhala, A.-J. Punkka, H. Hohti, and D. M. Schultz, 2009: Climatology of severe hail in Finland: 1930–2006. *Mon. Wea. Rev.*, **137**, 2238–2249, https://doi.org/10.1175/2008MWR2707.1.

——, ——, and D. M. Schultz, 2015: Significant hail-producing storms in Finland: Convective-storm environment and mode. *Wea. Forecasting*, **30**, 1064–1076, https://doi.org/10.1175/WAF-D-14-00159.1.

Waldvogel, A., B. Federer, and P. Grimm, 1979: Criteria for the detection of hail cells. *J. Appl. Meteor.*, **18**, 1521–1525, https://doi.org/10.1175/1520-0450(1979)018<1521:CFTDOH>2.0.CO;2.

Williams, E. R., 2001: The electrification of severe storms. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 527–561.

Wilson, C. J., K. L. Ortega, and V. Lakshmanan, 2009: Evaluating multi-radar, multi-sensor hail diagnosis with high resolution hail reports. *25th Conf. on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., P2.9, https://ams.confex.com/ams/89annual/techprogram/paper_146206.htm.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13**, 286–303, https://doi.org/10.1175/1520-0434(1998)013<0286:AEHDAF>2.0.CO;2.