Hailstorms over Switzerland
Verification of Radar-Based Hail Detection Algorithms with Crowd-Sourced App Data and Hail Sensor Data

Master's Thesis

Faculty of Science
University of Bern

Pascal-Andreas Noti
2016

Supervisor:

Prof. Dr. Olivia Romppainen-Martius
Institute of Geography and Oeschger Centre for Climate Change Research

Advisors:

Dr. Andrey Martynov
Institute of Geography and Oeschger Centre for Climate Change Research

Dr. Alessandro Hering
Federal Office of Meteorology and Climatology MeteoSwiss

Acknowledgement

I am grateful to Prof. Dr. Olivia Romppainen-Martius, my supervisor and the head of the Climate Impact Group. Her excellent guidance, expertise and support contributed much to the realisation of this study. Her suggestion to work on hail and thunderstorms was a great pleasure.

I am grateful to Dr. Andrey Martynov, my advisor at the University of Bern, for his great support with the data processing and for expertise. I learned much from his deep knowledge about programming, models and thunderstorms. He was always available for my questions. Thanks Andrey!

I am grateful to Dr. Alessandro Hering, my advisor from MeteoSwiss, for the data provision and for his great suggestions and expertise.

My deepest gratitude goes to my parents and my brother for encouraging me in all my pursuits, inspiring me to follow my dreams and supporting me unfailingly in all situations throughout all these years.

Many thanks to the Federal Office of Meteorology and Climatology MeteoSwiss, the Mobiliar and the inNET Monitoring AG for providing the data. Many thanks to all members of the Climate Impact Group for the great time, to Matthias Künzler from Mobiliar for providing data. Many thanks as well to all others who helped me writing this thesis. This journey would not have been possible without the kind support and help of my family, professor and advisors. Blessings upon you all.

Abstract

Hailstorms have the potential to cause severe damage to infrastructures, vehicles and agriculture. The losses resulting from hailstreaks can sum up to an amount which makes them one of the costliest natural hazards in Central Europe. Therefore, monitoring hail and the estimate of the hailstone diameter are essential tasks for weather forecasts, warnings and insurance assessments.

Crowd-sourcing is a low-cost way to allocate large datasets and becomes more popular in Meteorology. This Master's thesis focuses on the verification of radar-based hail detection algorithms using crowd-sourced data and hail sensor measurements. The Federal Office of Meteorology and Climatology MeteoSwiss and Mobiliar developed crowd-sourcing smart-phone applications, which collect hail reports from users. Additionally, hail measurements from 10 hail sensor stations located in regions with high hail occurrence was used for the analysis. Neighbourhood methods were introduced to face the chaotic nature of the user's reporting procedures. A median and an inverse distance weighting was computed for all neighbourhoods. The best match for each hail report in temporal and spatial windows was taken. Additionally, a categorical verification was applied for selected communities separately.

The crowd-sourced data and the hail sensor measurements correlate with the temporal and spatial occurrence and with the intensity of the radar-derived hail estimates. The upper quartile of the hailstone diameter derived by the detection algorithms correlate well with the reported hailstones sizes. The statistics of the larger hailstones suffered from the small number of reports. 27 % of all hail reports could be matched to hail detections of the radar-based algorithms. The hailstone drift and errors in the reporting procedure lead to many false reports. As a final remark, the temporal and spatial characteristics of the 5 minute radar-based hail detections can be captured in densely populated areas by the crowd-sourced data.

Contents

List of Abbreviations

° C	degree Celsius
%	percentage
a.s.l.	above sea level
AFWA	Air Force Weather Agency
CAPE	Convective Available Potential Energy
CH	Switzerland
CIN	Convective INhibition
COSMO	COnsortium for Small-scale MOdelling
cm	centimetre
CSI	Critical Success Index
dBZ	decibel relative to Z
E	East
e.g.	exempli gratia - for example
ECMWF	European Centre for Medium-Range Weather Forecasts
ESWD	European Severe Weather Database
et. al.	et alii (masculine), et aliae (feminine) or et alia (neuter)
ET45	Echo Top 45 dBZ
ET50	Echo Top 50 dBZ
EUCLID	EUropean Cooperation for LIghtning Detection
F	false alarm
FAR	False Alarm Rate
GPS	Global Positioning System
H	hit
H0	freezing level height
hPa	hectopascal
HS	Hail Size
ID	Identification
IDW	Inverse Distance Weighting
km	kilometre
M	missed
m	metre
MESHS	Maximum Expected Severe Hailstone Size

mm	millimetre
mPING	mobile Precipitation Identification Near the Ground
N	no-event
N	North
N	number
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Predictions
NL	Netherlands
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
Pa	pascal
POD	Probability Of Detection
POH	Probability Of Hail
SMS	short message - Short Message Service
TRT	Thunderstorm Radar Tracking
WOW	Weather Observation Website
WRF	Weather Research and Forcasting Model
W	Watt
UK	United Kingdom
US	United States of America
UTC	Universal Time Coordinated

List of Symbols

Symbol	Description	Symbol	Description
	No hail report from MeteoSwiss crowd-sourcing		No hail report from Mobiliar crowd-sourcing
	Coffee bean report from MeteoSwiss crowd-sourcing		Coffee bean report from Mobiliar crowd-sourcing
	1 Swiss franc coin report from MeteoSwiss crowd-sourcing		Chestnut report from Mobiliar crowd-sourcing
	5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Apricot report from Mobiliar crowd-sourcing
	Larger than 5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Large apricot report from Mobiliar crowd-sourcing
	Matched no hail reports from MeteoSwiss crowd-sourcing		Matched no hail reports from Mobiliar crowd-sourcing
	Matched coffee bean report from MeteoSwiss crowd-sourcing		Matched coffee bean report from Mobiliar crowd-sourcing
	Matched 1 Swiss franc coin report from MeteoSwiss crowd-sourcing		Matched chestnut report from Mobiliar crowd-sourcing
	Matched 5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Matched apricot report from Mobiliar crowd-sourcing
	Matched larger than 5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Matched large apricot report from Mobiliar crowd-sourcing
	Mismatched no hail reports from MeteoSwiss crowd-sourcing		Mismatched no hail reports from Mobiliar crowd-sourcing
	Mismatched coffee bean report from MeteoSwiss crowd-sourcing		Mismatched coffee bean report from Mobiliar crowd-sourcing
	Mismatched 1 Swiss franc coin report from MeteoSwiss crowd-sourcing		Mismatched chestnut report from Mobiliar crowd-sourcing
	Mismatched 5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Mismatched apricot report from Mobiliar crowd-sourcing
	Mismatched larger than 5 Swiss franc coin report from MeteoSwiss crowd-sourcing		Mismatched large apricot report from Mobiliar crowd-sourcing
	Matched user feedback of hail warnings		Measurements of hail from a hail sensor
	Mismatched user feedback of hail warnings		No hail measured by a hail sensor
	Nearest match approach		Best match approach
	Inverse distance weighting	$\tilde{x}$	median
	Reports from MeteoSwiss crowd-sourcing		Reports from Mobiliar crowd-sourcing
	Hail senor data from inNET		User confirmations to hail warnings

1   Introduction

On the 23th of July 2009 severe thunderstorms entered Switzerland from the southwest and crossed the Swiss Plateau and the Prealps in northeasterly direction. The storms produced hail which caused the costliest hail damage in the Swiss history. The damage on crops was appraised up to 35 millions, on buildings up to 160 millions, on vehicles up to 380 millions and on elementary damage pool up to 50 millions (*Torriani-Braga*, 2009; *Vetterli*, 2016). The total damage amount amounted to more than 600 million Swiss Franks.

Hailstorms occur mainly during June, July and August, when the more latent heat of condensation is available. Generally, hailstorms are a rare weather phenomena. The spatial extent of hail cells is very limited and their life time last from a few minutes to few hours. However, hailstones have a high kinetic energy and high impulse. The high potential energy makes them able to cause severe damages on infrastructure, vehicles and crops. For such reasons, hailstorms are one of the costliest natural hazards in Central Europe. Measuring the spatial and temporal extension and diameter of hail is crucial for evaluating damage claims, for forecasting and for warning systems. (*Schuster et al.*, 2005; *Kunz et al.*, 2009; *Kunz and Puskeiler*, 2010; *Berthet et al.*, 2013; *Mohr and Kunz*, 2013; *Kunz and Kugel*, 2015; *Nisi et al.*, 2016)

1.1   Hail Detection

Weather services, insurance, agricultural, aviation sector and the society have a strong interest in reliable forecasts and warning systems of events with such high damage potentials. However, nowcasting, the prediction of weather developments up to 6 hours ahead, is highly depending on an correct detection of thunderstorms and hailstones in order to predict hailstorms adequately (*Hering et al.*, 2005; *Li and Lai*, 2004; *Bonelli and Marcacci*, 2008; *Kober and Tafferner*, 2009). The detection and nowcasting of convective storms is quite tricky over complex topography due to spatially and temporally variable parameters, which influence the storm positively or negatively (*Mecklenburg et al.*, 2000; *Hering et al.*, 2004; *Rotach et al.*, 2009; *Mandapaka et al.*, 2012; *Nisi et al.*, 2014). Moreover, thunderstorms often develop and decay unpredictably due to their dependence on complex and rapidly evolving convective processes and their interaction with the terrain. On the other hand, the orographic forcing leads to regular development of thunderstorms in certain regions (*Foresti et al.*, 2011).

Ground-based observations were so far insufficient to measure hailstones for the following reasons. Point observations of hail do not represent large areas with the strong local-scale variability and limited spatial extent of hail events. The observational network has to be at least 10 times denser then conventional weather stations in order to capture hailstreaks appropriately (*Wieringa and Holleman*, 2006). Hailpads are a method to measure hail, but they are expensive (the involved manpower) for adequate measurement of hail over large areas and are not available in many countries, including Switzerland (*Betschart and Hering*, 2012). Only weather radars cover large areas and achieve high spatial and temporal resolution (*Basara et al.*, 2007; *Cintineo et al.*, 2012; *Kunz and Kugel*, 2015). The radar-based hail detection algorithms need, however, to be verified since they only allow indirect measurements of hail. Long time series of hail observations on ground as in other countries do not exist in Switzerland.

There were a couple of studies which conducted hail verification reasearch (e.g. *Kessinger et al.* (1995); *Witt et al.* (1998); *Schiesser et al.* (1999); *Holleman* (2001); *Ortega et al.* (2006, 2009); *Wilson et al.* (2009); *Hyvärinen and Saltikoff* (2010)). Insurance loss data has been widely used for validation and verification of radar-based hail detection algorithms (e.g. *Huntrieser et al.* (1997); *Schuster et al.* (2005); *Kunz et al.* (2009); *Kunz and Puskeiler* (2010); *Betschart and Hering* (2012); *Mohr and Kunz* (2013); *Skripniková and Řezáčová* (2014); *Kunz and Kugel* (2015); *Nisi et al.* (2016)). They cover large areas and have long time series, but have several limitations (*Willemse*, 1995; *McMaster*, 1999; *Changnon*, 1999; *Webb et al.*, 2001a); Population density, object vulnerability and claim handling strongly affect the hail reports aside the meteorological features (*Hohl et al.*, 2002a; *Mohr and Kunz*, 2013). *Betschart and Hering* (2012) and *Nisi et al.* (2016) verified the two radar-based algorithms of MeteoSwiss, the Probability of Hail (POH) and the Maximum Expected Severe Hail Size (MESHS).

According to *Nisi et al.* (2016) are the next steps in the verification of radar-based hail detection algorithms the usage of hail crowd-sourcing data and new automatic hail sensor network. In the summer 2015, the inNET Monitoring AG has installed a fully automatic hail sensor system, which detects the hailstone diameters by their kinematic energy and momentum (*inNET*, 2016). The measurements of the hail sensors can be considered as ground-truth for verifying the radar-based hail detection algorithms.

1.2   Crowd-Sourcing in Meteorology

Crowd-sourcing is defined according to *Howe* (2006) as:

"The White Paper Version: Crowdsourcing is the act of taking a job tradition-
ally performed by a designated agent (usually an employee) and outsourcing it to
an undefined, generally large group of people in the form of an open call."

According to *Boudreau and Lakhani* (2013) the crowd can be considered as an innovation partner for certain types of problems, but the crowd can outperform your company as well. There is a growing network of smart internet enabled devices due to the spread of smart-phones. This rise possibility for data crowd-sourcing, but advantages and disadvantages are involved (*Havlik et al.*, 2013); Several issues have to be considered: a reflection on the motivation for users participation, human and technical limitations of smartphone-enabled volunteer networks, legal and ethical challenges, reliability and usability issues, and issues related to trust and quality of information. *Wiggins and Crowston* (2011) sees the crowd-sourcing as *"a form of research collaboration involving members of the public in scientific research projects to address real-world problems"* and as a citizen science.

Using crowd-sourced data for monitoring weather activities has become more popular and is in the focus of weather services; *Michael Illingworth et al.* (2014) considers citizen science as a feasible and low-cost solution to increase the number of British rainfall-monitoring stations. *Hyvärinen and Saltikoff* (2010) collected ground-truth data in social media (text and photos) for the verification of hail detection algorithms. Crowd-sourced data is useful if they are available where no meteorological data is available. Crowd-sourced data can sometimes be unreliable. There were couple of other studies using crowd-sourced data for different meteorological purposes:

Betschart and Hering (2012) collected crowd-sourced hail reports on the European Severe Weather Database (ESWD hereafter). *Imran et al.* (2013) developed an algorithm, which extracts information from deasaster related messages in social media with machine learning methods. *Grasso et al.* (2016) identified weather events with Twitter messages. *Knapp et al.* (2016) detected and classified tropical cyclone storm types with crowd-sourced data. First weather services (e.g. UK Met Office, National Oceanic and Atmospheric Administration, NOAA hereafter, and Royal Netherlands Meteorological Institute) have integrated crowd-sourced data into the operational weather monitoring and forecasts. The UK Met Office launched a Weather Observations Website for collecting weather data measured by amateurs or organiations who have own automatic weather stations in 2011

(*Muller et al.*, 2015). The Royal Netherlands Meteorological Institute followed with a similar website (WOW-NL, https://wow.knmi.nl/) in april 2015 (*Koole and Siegmund*, 2016). Smart-phone Apps such as PressureNet and WeatherSignal can collect air pressure measurements of phones for forecast models (*Kim et al.*, 2015; *Muller et al.*, 2015; *Sumner*, 2015; *Hanson*, 2016; *Hoy and Klein*, 2016). The mobile Precipitation Identification Near the Ground project (mPING) of the National Severe Storms Laboratory (under NOAA) intends to improve dual-polarization algorithms with the help of reports from laypeople (*Elmore et al.*, 2014). The U.S. National Weather Service (under NOAA) has integrated social media posts and mobile phone photos into the operational protocol, weather information for the public and monitoring weather activities (*NWS*, 2015). *Longmore et al.* (2015) established the conceptual framework, architectural design and pathways to integrate mobile phone photos and display concept for operational severe weather monitoring.

Only a few studies have used hail reports as a basis of verification methods for validating radar-based hail detection algorithms (*Delobbe et al.*, 2005). *Betschart and Hering* (2012) conducted a study for the verification of the hail detection algorithms POH, MESHS and HAIL for the years 2009, 2010, and 2011. They collected data from the Internet with information about hailstone size and location for the year 2011. However, the allocation of hail reports has been time-consuming and not automated so far. The quality of the reports varies widely and the dependency on the population density is strong. Nevertheless, there is a high potential of automatically collected crowd-sourced data. The automatically compiled crowd-sourced data provides large amounts of hail reports which have not been achieved before. Furthermore, the temporal and spatial resolution is much higher compared to other data, such as insurance loss data.

The Federal Office of Meteorology and Climatology MeteoSwiss (MeteoSwiss hereafter) has been collecting hail reports (reported observations of hailstones by a person) in a crowd-sourcing project through smart-phone App since June 2015. The MeteoSwiss App provides the users with latest weather forecasts, current measurements and natural hazard warnings. The users have the opportunity to report hail events according to the hailstone size. Furthermore, MeteoSwiss launched a new hail warning system in collaboration with the Mobiliar insurance company.

1.3 Objectives

The main objective of this Master's thesis is the verification of radar-based hail detection algorithm, the evaluation of measurements from hail sensors and the evaluation of the crowd-sourced data from MeteoSwiss and Mobiliar. A prototype hail detection algorithm Hail Size (HS hereafter), which estimates the hailstone diameter, is evaluated by this study. Moreover, this project will attempt to answer the following questions:

- Is the crowd-sourced data from MeteoSwiss and Mobiliar reliable and usable for the verification of radar-based hail detection algorithms?

- Can the crowd-sourced data be used for the verification of radar-based hail detection algorithms?

- Do the reported hailstone sizes from the crowd-sourced data and measurements from hail sensors correlate with the radar-derived MESHS and HS (5 minute resolution)?

- Does the hail detection by the POH correspond with reports from crowd-sourced data, hail sensors and user feedback to hail warnings in selected Swiss communities (daily resolution)?

The next chapter describes theoretical backgrounds about hail, radar, the radar hail detection algorithms, hailstone drift and neighbourhood methods. The third chapter is about the data used in this thesis. In the fourth chapter, the neighbourhood and verification methods are explained. Afterwards the main are presented. The sixth chapter deals with the discussion of the results and the comparison with findings of other studies. This study is summarised and conclusions are formulated in the seventh chapter. The last chapter gives an outlook to this study.

2   Background

2.1   Generation of Hail

Hail is defined as a solid, frozen hydrometeor with a diameter larger than 5 mm. Smaller frozen hydrometeors are called graupel. Accoring to *Holleman* (2001) there are winter and summer hail. Winter hail (mostly under 20 mm) is produced in fast liftings caused by cold fronts and when the freezing level is close to the surface. Therefore, winter hail falls on much larger area than summer hail. Nevertheless, the diameter and damage potential of winter hail is smaller. Summer hail is generated by strong updrafts of organised convective systems such as multicell or supercell thunderstorms. The strong upward motion of air with high water vapour content leads to much oversaturation. There are generally few freezing nuclei compared to the condensation nuclei. Consequently, ice particles are hardly generated in strong updraft when temperature are below 0° C. The water vapour becomes liquid and remains unfrozen as supercooled water. The supercooled water can have temperatures down to -40° C. The primary growth process of hail is the collection of the supercooled cloud droplets and raindrops by falling down against the upward motion. The supercooled water freezes on contact with ice particles (opaque layer). If the hailstone temperature is above the freezing level, the supercooled water can enter the hailstone, fill the gaps and freeze then (clear layer). Thunderstorm are often tilted, which allows the hailstone to fall in another updraft region. Consequently, the residence and growth time in the thunderstorm can be extended by making multiple cycles, which lead then to huge hailstones. (*Heymsfield et al.*, 1980; *Pflaum*, 1980; *Nelson*, 1983; *Prodi et al.*, 1986; *Nelson*, 1987; *Angsheng*, 1993; *Pruppacher and Klett*, 1997; *Levi et al.*, 1999; *Pruppacher and Klett*, 2010; *Markowski and Richardson*, 2011; *List*, 2014)

A series of favourable weather conditions have to come together for severe storm formation: A very important factor for every thunderstorms plays the amount of Convective Available Potential Energy (CAPE hereafter), which acts like the fuel of the thermodynamic engine (*Groenemeijer and Van Delden*, 2007; *Grzych et al.*, 2007). A low laps-rate in the mid-troposphere is additional strong instability factor, which increases the buoyancy of the warm rising air and therefore the CAPE (*Droegemeier et al.*, 1993; *King*, 2016). The Convectivie INhibition (CIN hereafter) counteracts the CAPE by negative buoyancy (*Davies*, 2004; *Grzych et al.*, 2007).The CIN (in form of an inversion) creates a barrier for rising air parcels. A too high CIN would create a too strong negative buoyancy which is insurmountable and prevents thunderstorm formation. If there is no CIN, the CAPE would

be constantly dissipated by the rise of air masses. Therefore, the CIN should be low, but it is needed to build up CAPE.

Strong vertical wind shear between ground and around 6 km height accelerate the updrafts further (*Weisman and Klemp*, 1982; *Doswell and Evans*, 2003; *Kaltenboeck and Steinheimer*, 2015). Strong horizontal wind generates horizontal vorticity, which gets tilted vertically by rising air masses. Hence, positive and negative vertical vorticity is generated (*Markowski and Richardson*, 2011). The positive vertical vorticity speeds up the updraft and the negative vorticity the downdraft. Is the storm-relative mean flow normal to the horizontal vorticity, is this called crosswise vorticity. The streamwise vorticity is the case when the storm-relative mean flow and the horizontal vorticity are parallel. The streamwise vorticity brings the updraft and tilted vorticity region together, which musters the storm energy and sustains the longevity of the storm. The streamwise and crosswise vorticity exist in vertical unidirectional wind shear. The term storm-relative helicity is used in the case of a vertical turning wind shear. The storm-relative helicity is a measure of the angel between the direction of fluid motion and the vorticity of the fluid. A clockwise turning wind shear leads to stronger updrafts than streamwise vorticity and is known to supply essential energy to supercell thunderstorms. The CAPE, the CIN, the wind shear and the storm-relative helicity can be measured in nearby proximity soundings prior to the storm (*Brooks*, 2009). Thunderstorms can be triggered by mesoscale and large scale dynamics such as cold-fronts (*Holleman*, 2001; *Schemm et al.*, 2016).

2.2   Radar Equations

Radar is an acronym for RAdio Detecting And Ranging. The position, size, angle and velocity of objects in the atmosphere can be determined, which allows to detect civil and military aircraft, guided missiles, flying animals and weather related particles. The transmitter of the object-detection system generates short electromagnetic pulses with wave length ranging from approximately 1 to 10 cm. The antenna measures the returning echo/signal. In weather applicatinos, returning echo is due to the scattering of the electromagnetic pulses by hydrometeors. Their properties such as dielectric constant, number of hydrometeors and diameter define the scattering. (*Liljequist and Konrad*, 1994; *Seinfeld and Pandis*, 2016)

The average power received from a volume filled with hydrometeors is defined as (*Marshall et al.*, 1947; *Austin*, 1987; *Uijlenhoet and Pomeroy*, 2001):

$$P_r = P_0 \frac{\pi^4 Ah}{8r^2\lambda^4} \frac{(m^2-1)^2}{(m^2+2)^2} Z \quad [W] \tag{2.1}$$

$$P_r = C \frac{|K|^2}{r^2} Z \quad [W] \tag{2.2}$$

,where P_0 [W] is the power of the electromagnetic pulse, A [m^2] is the effective area of the antenna, h is defined as the length of the wave, r is the range [km], λ [mm] stands for the wavelength, C stands for the radar constant and Z [mm^6m^{-3}] expresses the radar reflectivity factor. The dielectric constant is determined as $|K| = (m^2-1)(m^2+2)^{-1}$, where m is the complex scatter index of refraction. The dielectric constant for water is approximately 0.93 (*Battan*, 1973) and for ice 0.208 (*Smith*, 1984).

Various studies investigated the relation between Z and diameter (D) of rain droplets, snow flakes and hailstones (*Spilhaus*, 1948; *Marshall and Palmer*, 1948; *Wexler*, 1948; *Marshall et al.*, 1952; *Marshall and Gunn*, 1952; *Sekhon and Srivastava*, 1970; *Markowitz*, 1976; *Smith*, 1984; *Aydin et al.*, 1986; *Sassen*, 1987; *Uijlenhoet and Pomeroy*, 2001; *Raupach and Berne*, 2016). Using the wavelength of weather radars, the reflectivity factor of rain, snow and hail is controlled by the Rayleigth scattering. The reflectivity factor of a hydrometeor mixture based on Rayleigh scattering is defined as:

$$Z = \int\limits_0^\infty N(D)D^6 \, \mathrm{d}D \quad [mm^6m^{-3}] \tag{2.3}$$

,where D [mm] is the equivolume particle diameter and N(D) [$m^{-3}mm^{-1}$] expresses the number of particles in a size interval. The radar reflectivity is mostly used in the logarithmic scale and gets transformed with the base of Z_0= 1 mm^6m^{-3} to [dBZ] as in the following:

$$dBZ \propto 10 * log_{10}(\frac{Z}{Z_0}) \tag{2.4}$$

2.3   Swiss Radar Network

The MeteoSwiss radar network consists today of five volumetric C-band Doppler radars. The radar stations are located in la Dôle (Jura Mountains), on the Albis (Zürich), on the Monte Lema (Ticino), on the Pointe de la Plaine Morte (Valais) and on Weissfluhgipfel (Grisons). The C-band Doppler radar network has been upgraded with dual-polarisation capability since the June 2011 (*Joss et al.*, 1998).

The Swiss weather radar network has been extended with two new stations in the inner alpine regions since summer 2015 (*Germann et al.*, 2015). The spatial resolution of the Cartesian products is 1 x 1 km^2.

A full 360° scan is computed every 5 minutes with the CCS4 software. The 5 minutes measurements can be aggregated to composites of daily or hourly maximum respective mean values of the radar-based hail detection algorithms for every grid point (*Germann et al.*, 2006; *Hering et al.*, 2004). Each radar station reaches a maximum range of 246 km.

The surrounding volume of the radar station is scanned in 20 different elevation steps by the antenna (see figure 2.1). The bottom pulse is emitted at an angle of 0° relative to the horizon. The largest angle for the beam is 40°. The whole scan process is accomplished in two cycles of 2.5 minutes. In the first cycle the antenna scans every second elevation angle (blue shaded). In the second cycle the antenna runs through the other elevation angles (yellow shaded). The radial velocities of the hydrometeors are derived from the phase shifts (from pulse to pulse). The clutter is automatically detected and removed. The first 2.5 minute cycle gets updated with the second in order to create a complete volume product. The products from each radar station are merged in order to generate a composite products.(*Germann et al.*, 2015)

Two radar-based hail-detection algorithms, the Probability of Hail (POH) and the Maximum Expected Severe Hailsize (MESHS), are derived from the 45 dBZ and 50 dBZ echo top height products (ET45 and ET50 hereafter). ET45 (ET50) corresponds to the highest altitude at which at least 45 dBZ (50 dBZ) is detected (*Donaldson*, 1961). Diagnosis of severe convection (*Held*, 1978; *Waldvogel et al.*, 1979; *Witt et al.*, 1998) and radar-based thunderstorm nowcasting systems (*Dixon and Wiener*, 1993; *Johnson et al.*, 1998; *Hering et al.*, 2008) are based on such echo top heights.

Beam shielding, strong ground clutter, beam overshooting and bright-band contamination lead to multifaceted error structures over complex terrain by radar measurements (*Joss et al.*, 1998; *Holleman et al.*, 2006; *Germann et al.*, 2006; *Villarini and Krajewski*, 2010; *Mandapaka et al.*, 2012; *Meischner*, 2013; *Kaltenboeck and Steinheimer*, 2015). Fortunately, the echo top height products are only minimally affected by those parameters, because in case of hailstorms the ET45 and ET50 are typically at elevations far above the ground.
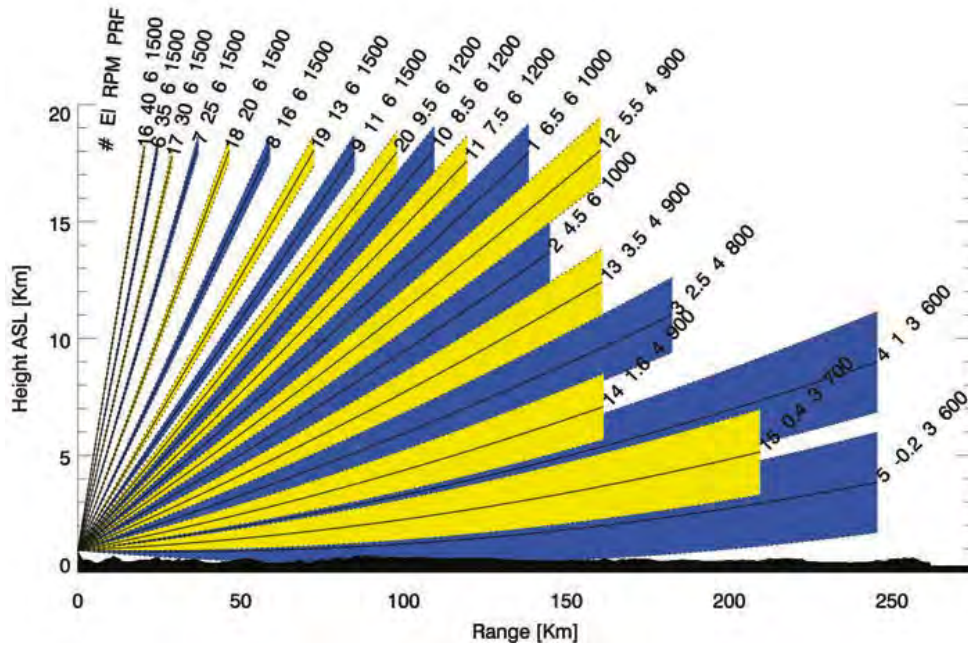
Figure 2.1: The Swiss radar scan strategy in a vertical cross section. The blue shaded areas represent the scanned volumes of the during the first scan cycle of 2.5 minutes. The yellow shaded areas represent the scanned volumes of the during the second scan cycle of 2.5 minutes. The lines in the middle of the shaded areas are the center of the outgoing electromagnetic pulse. The numbers on top and on the right of the figure stand for the angle of outgoing pulse. (*Germann et al.*, 2015)

2.4   Regional NWP Model COSMO-CH

MeteoSwiss operates a modified high-resolution model from the COnsortium for Small-scale MOdelling (hereafter COSMO). The COSMO-CH analysis (http://cosmo-model.org) provides the freezing height (hereafter H0), which is used to compute the POH and MESHS. An interpolation is needed where the H0 value is missing due to the higher elvation of the Alps. COSMOS-7 with a resolution of 6.6 x 6.6 km^2 is used by MeteoSwiss since 2003. The European Centre for Medium-Range Weather Forecasts (ECMWF hereafter) provides the global forecasts for the boundary conditions for European regional forecast models such as the COSMOS-7. MeteoSwiss has been operating COSMO-2 at a spatial resolution of is 2.2 x 2.2 km^2 and temporal resolution of 1 hour since 2008. COSMO-2 is embedded inside COSMO-7. Both COSMO models get the initial state of atmosphere by an own assimilation system (*Leuenberger and Rossa*, 2007), which collects data from measurement stations, radiosondes, composite of all Swiss radars, commercial planes and so on. (*Doms et al.*, 2015; *Germann et al.*, 2016)

11

2.5 Probability of Hail (POH)

Parameters such as the vertical distance between ET45 and H0 has been extensively studied by *Waldvogel et al.* (1979). Hail may grow through riming in deep convective storms only at certain zone, which corresponds the height difference between ET45 and H0. *Waldvogel et al.* (1979) showed that the vertical distance between ET 45 and H0 is a good estimator of presence of hail on ground in Switzerland (see equation 2.5). This finding was experimentally verified by later studies (*Foote and Knight*, 1979). The freezing level H0 is highly depending on latent energy from melting and evaporation processes. The hail detection algorithms neglect such small-scale processes in the calibration of the parametrization of the soundings (*Waldvogel et al.*, 1979; *Treloar*, 1998; *Joe et al.*, 2004).

The Probability of Hail ((*Witt et al.*, 1998)) is computed out of the height distances between ET45 (the maximum height which 45 dBZ is measured at) and H0 (see equation 2.5). The height distances smaller than 1.65 km correspond to POH values of 0 %, whereas height distances greater than 5.5 km represent POH values of 100 % (see equations 2.6 and 2.7). The POH values give no information about the hailstone size and are only a measure for presence of hail. Weather services use the POH algorithm operationally for their forecasts (*Witt et al.*, 1998; *Holleman*, 2001; *Salek et al.*, 2004; *Betschart and Hering*, 2012; *Puskeiler*, 2013; *Kunz and Kugel*, 2015). MeteoSwiss uses an updated version suggested by *Foote et al.* (2005) operationally since 2008.

$$\triangle z = ET45 - H0 \tag{2.5}$$

$$\triangle z > 1.65 \; km \Rightarrow POH \; of \; 0 \; \% \tag{2.6}$$

$$\triangle z > 5.5 \; km \Rightarrow POH \; of \; 100 \; \% \tag{2.7}$$

2.6 Maximum Expected Severe Hail Size (MESHS)

The Maximum Expected Severe Hail Size is a radar-based hail detection algorithm for estimating the largest hailstone diameter which can be expected on the ground. *Treloar* (1998) studied the relation between several radar parameters and upper air features. The relation of ET50 (the maximum height which 50 dBZ is measured at) and H0 can be used for estimating the MESHS (see Figure 2.2). The MESHS detects maximum hailstone diameters larger than 2 cm. This hailstone size is

approximately a minimum critical threshold value for causing damages (*Nisi et al.*, 2016). The MESHS was developed during the Sydney Forecast Demonstration project (*Joe et al.*, 2004) and implemented in the operational forecast during the Olympic Games in 2000 (*Webb et al.*, 2001b; *Wilson et al.*, 2004). The MeteoSwiss MESHS algorithm (see equation 2.8) uses the ET50:
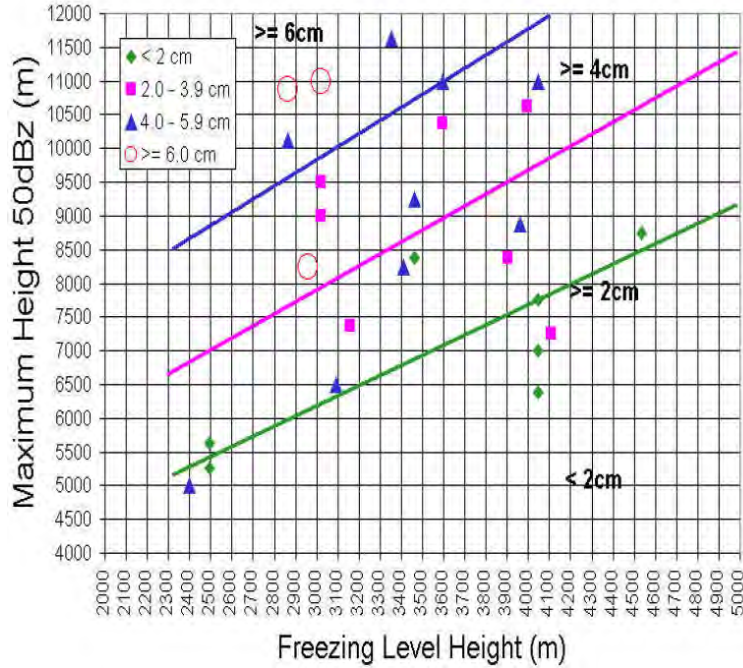

Figure 2.2: The "50 dBZ hail nomogram" calibrated for local hail climatology in the Sidney area. The soundings were measured by the Sydney Airport and the radar reflectivity by the S-band Sydney radar. (*Treloar*, 1998; *Joe et al.*, 2004)

$$\triangle z = ET50 - H0 \tag{2.8}$$

2.7 Thunderstorm Radar Tracking (TRT)

The Thunderstorm Radar Tracking (TRT hereafter) was developed and evaluated by an cooperation between MeteoSwiss and Météo-France. Cells of severe thunderstorms are detected and the track of the detected thunderstorm cells is calculated. The motion of the cells are determined by their displacements and the persistence of their evolution. The movement of the tracked cells is predicted in 5 minutes for lead-times up to 30 minutes by an extrapolation based on Langrangian persistence rules. The TRT includes a heuristic cell severity ranking algorithm to classify each storm into 5 severity classes (*Hering et al.*, 2008; *Rotach et al.*, 2009). The ranking is defined by a fuzzy-logic-like scheme that uses cell-based attributes such as the

Vertically Integrated Liquid, the ET45, the maximum cell reflectivity and the area above 55 dBZ. Uncertainties are calculated from the standard deviation of the velocity vectors from the last three time steps and get incorporated in the TRT. (*Hering et al.*, 2005, 2015)

2.8   Hailstone Drift

The horizontal wind can transport hailstones away from the radar-detected position. Especially, strong downdrafts spread out horizontally near the ground and enlarge the distances of the drifted hailstones. *Schuster et al.* (2005) investigated the relationship between building insurance loss data and the hail-derived kinetic energy. They found that hailstreaks can be displaced to the neighbouring grid cell (grid cell size of 2 km). *Kessinger et al.* (1983) analysed the horizontal wind fields of downdraught regions within thunderstorms by using the Doppler radar of the Denver Airport in Colorado. Wind fields, which diverge in all directions, are found around the location of the highest radar reflectivity core when the downdrafts get closer to the ground. The strongest wind gusts occur on the left side of the storm track (right-moving storm after storm splitting). When the downdraft get closer to the ground, it turns more horizontally and towards 180° relative to the storm direction. (*Kessinger et al.*, 1983; *Wilson et al.*, 1984; *Kessinger et al.*, 1988)

Knupp (1987) and *Schmid et al.* (1992) have analysed the main downdraft directions in convective storms. *Schiesser* (1990) developed a theoretical formula for estimating the fall distance of hailstones:

$$d = \frac{(H_r - H_g)U}{V_t} \tag{2.9}$$

,where mean horizontal wind speeds U between the ground level H_g and the lowest radar measuring level H_r within downdraught regions were calculated by *Browning and Foote* (1976). *Matson and Huggins* (1980), *Mezeix and Admirat* (1978) and *Knight and Heymsfield* (1983) measured the dependence of the terminal velocity V_t based on the mean hailstone diameter D. They concluded that hailstones with 0.5 and 1 cm radius may reach a final fall speed of 10 respectively 20 ms[-1]. According to *Browning* (1977) and *Houze et al.* (1993) the most frequent horizontal wind speeds in thunderstorm ranges between 10 and 25 ms[-1] with gusts up to 50 ms[-1].b*Hohl et al.* (2002b) calculated a 3 km horizontal drift distance for hailstone with 1 cm diameter falling from 1.5 km a.s.l. (Radar Measuring Level) to 0.4 km a.s.l. and affected by horizontal wind of 20 to 30 ms[-1].

Schuster et al. (2006) computed the theoretical fall distance d with the formula 2.9 and found distances ranging from 2 and 2.8 km. The total radar kinetic energy fields reach the highest correlations with insurance loss data when they are shifted by the distance of the hailstone drift as it has been investigated for hail damage crops (*Schiesser*, 1990), for hailpads (*Schmid et al.*, 1992) and for hail damaged buildings (*Hohl et al.*, 2002a).

2.9 Neighbourhood Methods

In order to account fo the unknown displacement of the reported hailstones, this study uses neighbourhood methods. *Casati et al.* (2008), *Gilleland et al.* (2009) and *Jolliffe and Stephenson* (2012) reviewed and summarized new verification approaches. Many new spatial verification approaches have been developed independently in meteorology and especially in numerical weather prediction (*Zepeda-Arce et al.*, 2000; *Atger*, 2001; *Brooks et al.*, 2003; *Damrath*, 2004; *Turner et al.*, 2004; *Weygandt et al.*, 2004; *Marsigli et al.*, 2005; *Theis et al.*, 2005; *Rezacova et al.*, 2007; *Segawa and Honda*, 2007; *Roberts and Lean*, 2008; *Cullen and Brown*, 2009; *Gilleland et al.*, 2009).

The neighbourhood methods (fuzzy methods) validate forecasts within spatial and/or temporal neighbourhoods (*Ebert*, 2008, 2009). These methods are called "fuzzy", because they may include wrong values which might not be related to the forcasted values. This can affect the verification negatively. On the other hand, the model may produce values which are temporal, spatially displaced, but they are the equivalent values to the observations. *Lorenz* (1969) already described the so-called "double-penalty effect". The double-penalty effect consists of negative effects of displacement errors and closeness. Neighbourhood verifications often contain several spatial and temporal windows surrounding each pixel. The verification is not sensitive to a circular or square neighbourhood shape (*Jolliffe and Stephenson*, 2012). The measure for the verification can be the mean difference, binary category or different simple statistics.

The "skilful scale" is the optimal spatial scale at which an acceptable level of performance is reached (*Mittermaier and Roberts*, 2010). The verification grid should be bigger than the spatial displacement. A too small scale leads to low aggreement of the forecast and observation since the displacement is bigger than the spatial window of the verification. A too coarse scale may contain values from nearby events and the verification would give too good scores even though the real aggreement might be bad. *Mittermaier* (2014) suggested never to use the native grid resolution for comparing the results in the verification.

High resolutions are generally problematic in meteorological verification methods; Traditional verification statistics validate fine-scale models negatively due to small scale differences which do not exist in coarser resolution forecasts (*Ebert*, 2008). Furthermore, the observations can contain errors of sampling, measurement and representativeness, which are much stronger at fine-scale resolution and badly influence the verification (*Tustison et al.*, 2001). Generally, it is nearly impossible for a forecast model to predict weather at high-resolution due to the chaotic nature of the atmosphere. Spatial verification techniques have to be developed further in order to show the benefits of near-convection-resolving model forecasts (*Mass et al.*, 2002; *Schwartz et al.*, 2009; *Mittermaier et al.*, 2013; *Smith et al.*, 2013; *Mittermaier*, 2014).

3 Data

This Master's Thesis used the MESHS, POH and HS radar-based hail detection algorithms for the analysis. The Max Echo, MESHS and POH were allocated from the MeteoSwiss. The HS was computed out of the MESHS and POH and is described in the 4th chapter. This chapter describes the other datasets, the crowd-sourced report, user feedback to hail warnings, WRF data and lightning data in the following. The WRF data was used in the case studies to support the findings. The lightning data and the Max Echo helps to identify sever thunderstorm in the case studies. Peaks of lightning activities can indicate the start of the hail generation.

3.1 MeteoSwiss Crowd-Sourced Data

MeteoSwiss have been collecting the crowd-sourced hail reports since the beginning of June 2015 by their smart-phone App. This Master's thesis used the data collected from June, July and August 2015 and April, May, June, July, August 2016. The interface of the App allows to choose among 5 options which should indicate the size of the observed hailstone. The user have the option to report "No hail", "Coffee bean", "1 frank coin", "5 frank coin" or "Greater than 5 frank coin". The position and time of the report is derived by the mobile network localisation and the network time. The users of the MeteoSwiss App can manually change the actual location and/or current time of the hail event. Each report contains, apart from the size, location and time of the hail event, some anonymous information about the user's phone such as language, operating system and software ID. 27'031 hail reports have been collected during the 8 months in 2015 and 2016.

3.2 Mobiliar Crowd-Sourced Data

Mobiliar started as well a crowd-sourced collection of hail reports through a smart-phone App at the beginning of June 2015. The Master's thesis used data from June, July and August 2015 and April, May, June, July, August 2016. The user can report the absence of hail as "No hail", a hail event as "Coffee bean", "Chestnut", "Apricot", "Big apricot" or as "I do not know". The position and time of the report is derived by the mobile network localisation and the network time. Both coordinates are rounded to the next half kilometre steps. During the 8 months in 2015 and 2016 3'393 reports were sent. An anonymous individual software ID is added to each report.

3.3   inNET Hail Sensor Data

The inNET developed an automatic hail sensor network which measures hailstone diameters and their damage potential (*inNET*, 2016) in Switzerland. *Löffler-Mang et al.* (2011) had the idea to detect individual hailstones with small piezo-electric microphones inside a PVC body with a octagonal shape. The octagonal shape avoids weak signals from isolated corners as for instance if the basic plate would be squared. One microphone is located in middle of the top plate and the other in the middle of the bottom octagonal plate. The vibrations caused by the hailstones generate voltage in the pieco-electric microphone. The signals from the microphone get converted into momentum and kinetic energy, which can be converted into diameters of the hailstones by using an appropriate calibration. The data is immediately transmitted to cloud-based servers, which allows a real-time measurement of hail. A first test network of 10 hail sensor stations got installed in the June 2015. The hail sensors have been tested in hail hotspot regions of Switzerland, such as the Napf-region. The stations are installed on flat roofs of public buildings or at a MeteoSwiss weather station. This study computes the average of the diameter derived from the kinetic energy and the diameter derived from the momentum.

3.4   User Feedback to Hail Warnings

MeteoSwiss designed a fully automatic hail warning system, which is based on the operational, multi-sensor nowcasting system TRT and the POH. The MeteoSwiss hail warning system considers the POH additionally to the TRT (see figure 3.1). A threshold value of POH > 60 % is used for defining the presence of hail (binary field). The uncertainty, size and core of the hail cell are estimated by an ellipse. If such an ellipse moves over an defined warning region of MeteoSwiss, a flash warning SMS is sent to all costumers and a selected group of volunteers within that region. The volunteers confirm if there was hail. A confirmation gets only accepted, if the costumer sents the SMS inside the post code region of his physical address. A reminding SMS is sent, if there was no response within 30 minutes.

3.5   WRF Model Data

The Weather Research and Forecasting (WRF hereafter) model version 3.6.1 designed by the National Centers of Environmental Prediction and the National Center for Atmospheric Research (NCEP/NCAR hereafter) was implemented in the case study. The domain of the simulation reaches an extension from 42.72°
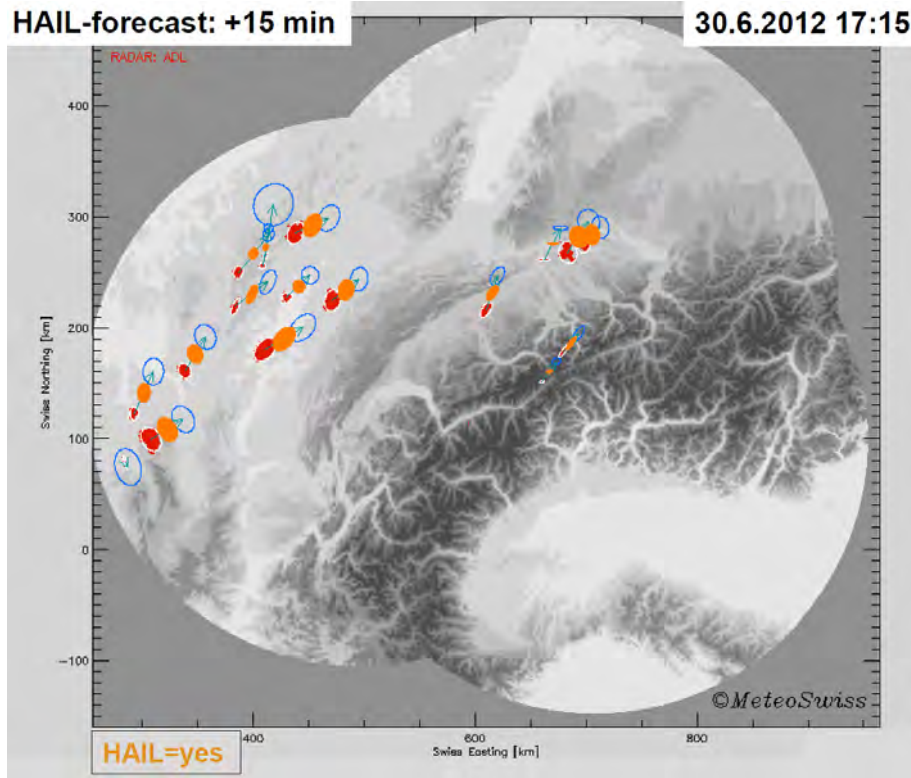
Figure 3.1: The hail forecast of the MeteoSwiss hail warning system from the 30th of June 2012 at 17:15 UTC. Red ellipses show detected hail cells based on measurements at time-lag 0. Orange ellipses represent the forecasted cells for 17:30 UTC. Blue bordered ellipses are the forecasted hail cells for 17:45 UTC.

to 49.91° N and from 4.14° to 12.09° E. An spatial interpolation was applied in order to get a horizontal resolution of 1 x 1 km². The resolution of the model itself is 2.14 x 1.35 km². The model has 35 vertical levels with a top level of 50 hPa. The initial state is given by the ECMWF analysis data. The boundary conditions were updated with forcing data from the ECMWF analysis every 6 hours. The model contains the Noah land-surface model, but no cumulus parametrization. The new Thompson scheme (*Thompson et al.*, 2008) and the Morrison double-moment scheme (*Morrison et al.*, 2009) were used to define the microphysical processes in the model. The Air Force Weather Agency (AFWA) Diagnostics is implemented to compute hail diameters. The AFWA package for WRF (*Brimelow et al.*, 2006; *Creighton et al.*, 2014) uses a HAILCAST-1D hail model that produces the AFWA-HAIL NEWMEAN as hail diameter output. Changing the microphysical conditions can lead to different results even though the initial conditions, configuration and boundary conditions remain the same (*Jankov et al.*, 2005; *Otkin and Greenwald*, 2008; *Jankov et al.*, 2009; *Mercader et al.*, 2010; *Rajeevan et al.*,

2010; *Awan et al.*, 2011; *Cossu and Hocke*, 2014; *Halder et al.*, 2015; *Que et al.*, 2016).

3.6   Lightning Data

In 2001 Austria, France, Italy, Norway and Slovenia founded the European Co-operation for LIgthning Detection (EUCLID hereafter). The EUCLID networks detects lightnings and provides the data to the forecasting offices. Electromagnetic sensors measure the signals emitted by the lightning return stroke. The time information is given by GPS satellites. Nowadays, other European countries including Switzerland joined EUCLID. MeteoSwiss supplied the lightning data to this study. The lightning dataset contains information about location, time, number of lightnings, type of lightning (cloud-to-ground lightning or cloud-to-cloud lightning), current intensity and current polarity. (*Diendorfer*, 2002; *Schulz et al.*, 2014a,b)

4 Methods

The hailstone drift has benn reviewed in chapter 2.8. Individual hailstones can be drifted horizontally 2 to 2.8 km with the wind when falling from level of radar detection to the ground (*Schuster et al.* (2006)). The horizontal displacement of hailstone by drifting has to be considered in the verification. Since the hailstone falls a considerable time length from the radar detection level to the ground. The verification can miss such displaced hailstones when a high resolution analysis is applied (5 minute and 1 km resolution). The verification methods have to be flexible enough to integrate values which are temporally and spatially shifted. Nevertheless, the verification methods have to be conservative in the sense that not all values can be integrated.

Verifcation approaches for validating radar-based hail detection differ slightly from the convential verification designs in meteorology; So far the verifcations of hail detection used a daily resolution and the verfication grid is just centred around the grid cell, which contains the hail report. *Hohl et al.* (2002b), *Schuster et al.* (2006) and *Betschart and Hering* (2012) used a 3 x 3 verification grid with a grid size of 2 x 2 km^2. *Kessinger et al.* (1995),*Witt et al.* (1998), *Holleman* (2001), *Wilson et al.* (2009) and *Hyvärinen and Saltikoff* (2010) used spatial grid up to 30 km. *Kunz and Kugel* (2015) used a single grid with a grid size of 10 x 10 km^2.

4.1 Neighbourhood Concept and Design

This Master's thesis takes a new approach: The goal is to use the 5 minute resolution of the radar data and not to use the daily aggregation as the other verification studies (the radar-based hail detection algorithms) did. On the hand, the spatial uncertainties and the drift of the hailstone are encountered by spatial windows, the neighbourhoods. A neighbourhood contains grid cells respectively pixels which are within a defined radius around a hail report and within a specific time period. On the other hand, the temporal uncertainties and temporal shifts are faced by temporal windows. Temporal series of neighbourhoods make up the temporal windows. In here, 9 temporal windows and two different radii of the neighbourhoods are used for each hail report. 18 neighbourhoods ($9 * 2 = 18$) centred at the location of each hail report gain the pixel values from the hail detection algorithms (see figure 4.1).

This neighbourhood approach focuses on hailstones larger than 2 cm. It is assumed here that hailstone larger than 20 mm are not transported further than 4 as smaller hailstone might do. Large hailstones fall with higher speeds and are affected less by the horizontal wind due to the their higher ratio of mass volume to surface area, if the hailstones are spherically shaped. Therefore, a neighbourhood with 2 km radius is used. The 2 km radius was chosen following similar minimum grid sizes from the other verification studies (*Schuster et al.*, 2006; *Hohl et al.*, 2002b; *Betschart and Hering*, 2012). A 4 km radius of the neighbourhood was chosen as the maximum spatial tolerance (the double of 2 km) and encounters additional spatial uncertainties which are due to movements of the App users. The 4 km neighbourhood is in the limit of the maximum drift distance of smaller hailstones as well. The smallest hailstones and graupel may get transported further and not be detected in the neighbourhoods. The distances were calculated as length between the middle point of the 1 x 1 km^2 radar pixel and the reporting position.

To account for the temporal uncertainties 9 temporal windows (each with the 2 spatial windows) were set. The maximum temporal tolerance is 45 minutes (not fully centred around the reporting time). This study assumed that a higher accuracy of the reports is expected if the temporal tolerance is short. Since the user can customise the time and place, a short time period before and after the reporting time has to be analysed. People may forget quickly about the precise size of the hailstone and an exact time of the hailstone event. This study wants to select hail reports of higher quality by a short time tolerance. Therefore, a 20 minute long time period was chosen for the time tolerance before and after the reporting time. It should be possible to report within 20 minutes after the hail event. This study did not investigate the optimal time period. It has to be first proved that crowd-sourced App data can be used for verification, especially at this high resolution. The temporal windows have the same resolution as the radar data (5 minute). The "central" time period is determined by the 5 minute period which the reporting time falls in. The reporting time is always rounded up to the next 5 minute step to be consistent with the 5 minute measuring procedure of the radar stations (ceiling time hereafter). The other 8 temporal windows are distributed before and after the ceiling time.
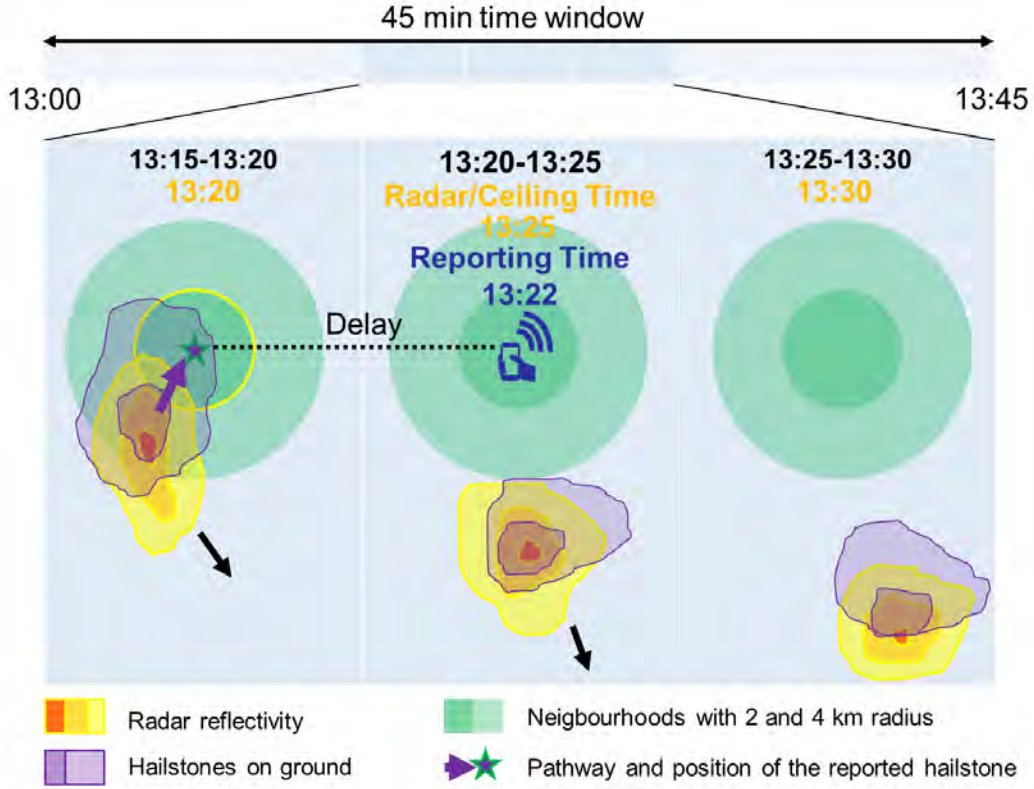
Figure 4.1: Neighbourhood design with its temporal and spatial windows. For illustration an abstract thunderstorm detected by radar moves from north-west to south-easterly direction. The median of the radar data inside of 18 neighbourhoods is computed for every single report. The 18 neighbourhoods consist of 2 spatial windows(2 and 4 km) x 9 time windows (9 x 5 minute time periods). 8 before and 8 after the time period which contains the hail report

.

4.2   Spatial Determination of the Hail Detections Values

The reported hailstone size is compared to a sinlge radar value for each hail report. Therefore, a single value has to represent each neighbourhood. Just taking the most simlar hail detection pixel value is wrong. The user can make mistakes in reporting the size, the time and the location of the hailstone. Moreover, the hail detection algorithm can provide wrong diameters. The MESHS is the estimation of the largest hailstone within a pixel. The user cannot see all hailstones in his or her neighbourhood. It is not guaranteed that the largest hailstones is reported by the users. Consquently, all pixel values within the neighbourhood are considered. The median is neutral to extrem values and might be a good choice if the user is wrong about the location and the reported hailstone size. An interpolation with the Inverse Distance Weighting (IDW hereafter) can be good choice if the user reports size and location of the hail correctly. A median respectively an IDW of

all pixels larger than 0 mm respectively 0 % is computed for each neighbourhood. This Master's Thesis follows the idea of ignoring zeroes like similar to *Betschart and Hering* (2012).

The interpolation with IDW allows to adjust the spatial weighting according the reported size. Larger hailstones are less affected by the drift than smaller hailstones. Larger power parameters of the spatial weighting was choosen for larger hailstones. The IDW interpolation function was calculated according to *Shepard* (1968):

$$IDW = \frac{\sum\limits_{i=1}^{n} \frac{x_i}{d_i^p}}{\sum\limits_{i=1}^{n} \frac{1}{d_i^p}} \tag{4.1}$$

, d is the distance between the reporting position and the pixel. x_i is denoted as the value of a single pixel. p represents the power parameter (exponent) and is set for the reportable sizes as the following: The MeteoSwiss App reports get the factors 2 for coffee bean, 3 for 1 Swiss frank coin, 4 for 5 Swiss frank coin and 5 for greater than 5 Swiss frank coin. The power parameter for the hail sensor measurements are calculated by dividing the measured diameter by 10.

Time has the priority over space. Therefore, all 2 km neighbourhoods are looked through for matches before testing all 4 km neighbourhoods. If there are no matches, the 4 km neighbourhoods are scanned through for detecting matches. Furthermore, only medians respectively inverse weighted distances larger than 0 mm respectively 0 % were considered as matches. The following two approaches for prioritizing median respectively inverse distance weighted of the temporal windows were conceptualised:

4.3 Nearest Match Approach

The nearest match approach assumes that the hail report is sent as soon as possible after the hail event. The hail event is expected to occur before the report is sent and very close to the reporting time. In first run the algorithm looks for matches in all 2 km neighbourhoods. It is first tested if there is a match in the central period. If this is not the case, the 5 minute before the central one is scanned for matches. Is this not the case, the 5 minute period after the central one is checked. Is this again not the case, the algorithm looks for matches in the 5 minute period, which 10 minute before the central period. This scanning process goes for and

back until it finds a match or it reaches the last 5 minute period (20 minutes after the central period). Then it would start analysing all 4 km neighbourhoods with the same procedure as described before. If the algorithm finds a match with the Nearest match approach, the median and the inverse weighted distance of the neighbourhood with the prefix "nearest" are transferred to dataset containing the crowed-sourced reports respectively hail sensor measurements. The time of the neighbourhood with the match is called "nearest radar time" and transferred as well. The nearest radar time is expected to be the time period when the hailstone was measured by the radar.

4.4   Best Match Approach

The best match approach assumes that reported size is very similar to the real size of the hailstone. Therefore, it is looked for the smallest size difference between the reported size and the size measured by the radar. The approach neglect the melting of the hail and trust more in the user's (correct) estimation of the hailstone size. For calculating the size difference, diameter values were added to the reportable sizes in the App. The following diameter were set for the reportable sizes of MeteoSwiss App: 0.1 mm for "No hail", 8 mm for "Coffee bean", 23 mm for "1 Swiss franc coin", 31 mm "5 Swiss franc coin" and 42 mm for "Greater than 5 franc coin". The following diameter were defined for the Mobiliar App: 0.1 mm for "No hail", 8 mm for "Coffee bean", 15 mm for "Chestnut", 40 mm for "Apricot" and 60 mm for "Big apricot". For computing the difference between the hail sensor measurements and the median respectively the inverse weighted distance, the average of the diameter derived from the kinetic energy and momentum was taken. The time with a match is called "best match radar time".

4.5   Categorical Verification

Categorical verification is often used in meteorology to verify radar-based hail detection algorithms (*Huntrieser et al.*, 1997; *Delobbe et al.*, 2003; *Brimelow et al.*, 2006; *Kunz and Puskeiler*, 2010; *Betschart and Hering*, 2012; *Skripniková and Řezáčová*, 2014; *Kunz and Kugel*, 2015). The radar data is converted to binary data according to the thresholds $POH > 60 \%$. The Probability Of Detection POD, False Alarm Rate FAR and Critical Success Index CSI are verification scores, which are calculated from a 2 x 2 contingency table (*Donaldson*, 1961; *Holleman*, 2001). The contingency table represents the joint distribution of the detections and observations. The term "observation" means a hail or no hail event reported

Table 4.1: 2x2 contingency table for POH verification according to (*Betschart and Hering*, 2012)

	Hail observation	No hail observation
Radar detection	H	F
No radar detection	M	N

from the ground. The term "detection" describes the measurement of hail or no hail from weather radars.

The combination of a detection and an observation of a hail event is defined as *hit* or variable *H*. The *false alarm* or variable *F* corresponds the combiniation of a detection and no observation on the ground. The combination of no detection and an observation is indicated as variable *M* or *missed*. The variable *N* or *no-event* represents no detection and no observation of hail events. The sum of H and F (M and N) is named as positive (negative) detection. Similarly, the sum of H and M (F and N) are called positive (negative) observation. (*Wilks*, 2006)

The Probability Of Detection (POD) is defined as the ratio of hits (H) divided by the positive observations (H and M). A POD value of 1 corresponds to a 100 % accurate detection of hail events, wheras a POD value of 0 % means non of the hail events were detected by the radar but by ground obsevations. A high POD and a low FAR means a good result of a verification (*Holleman*, 2001). The POD is defined in the following equation (4.2):

$$POD = \frac{H}{H + M} \tag{4.2}$$

The False Alarm Rate (see equation 4.3) is defined as the ratio of the no hail detections (false alarm or „F") divided by the positive detections (F+M). A False Alarm Rate value of 1 means that all radar detections have no corresponding observation on the ground. Low FAR values correspond to high consistence of detections and observations. Several studies doing hail verification indicated that the FAR can be problematic since it does not allow to really distinguish between „wrong hail detection" or „no obsevation" because nobody respectively no measurement system on ground reported hail (*Delobbe et al.*, 2003; *Delobbe and Holleman*, 2006; *Saltikoff et al.*, 2010; *Betschart and Hering*, 2012; *Kunz and Kugel*, 2015).

The Critical Success Index (CSI, see equation 4.4) is defined as the ratio of hail detections (hit or H) divided by the sum of positive observations and no hail detections (H+F+M). A Critical Success Index value of 1 corresponds to the best

score, whereas the value of 0 to the worst score. A high POD and low FAR value
is equal to a high CSI.

$$FAR = \frac{F}{H + F} \tag{4.3}$$

$$CSI = \frac{H}{H + M + F} \tag{4.4}$$

Nisi et al. (2016) has done a categorical verification for the 25 most populated
in Switzerland in order to have a high potential of hail reporting (claiming dam-
age, insurance loss data). Since this study works only with data over 8 months,
there would be few communities with hail occurrence. Therefore, more communi-
ties have to be considered. The categorical verification in this study is done for
Swiss communities, which contain at least 30 % settlement area (see figure 5.21).
A selection according to population density neglects the spatial distribution of the
population within populated communities. The information about the commu-
nities (settlement share) were downloaded from the website of Federal Statistical
Office (https://www.bfs.admin.ch/bfs/en/home.html).
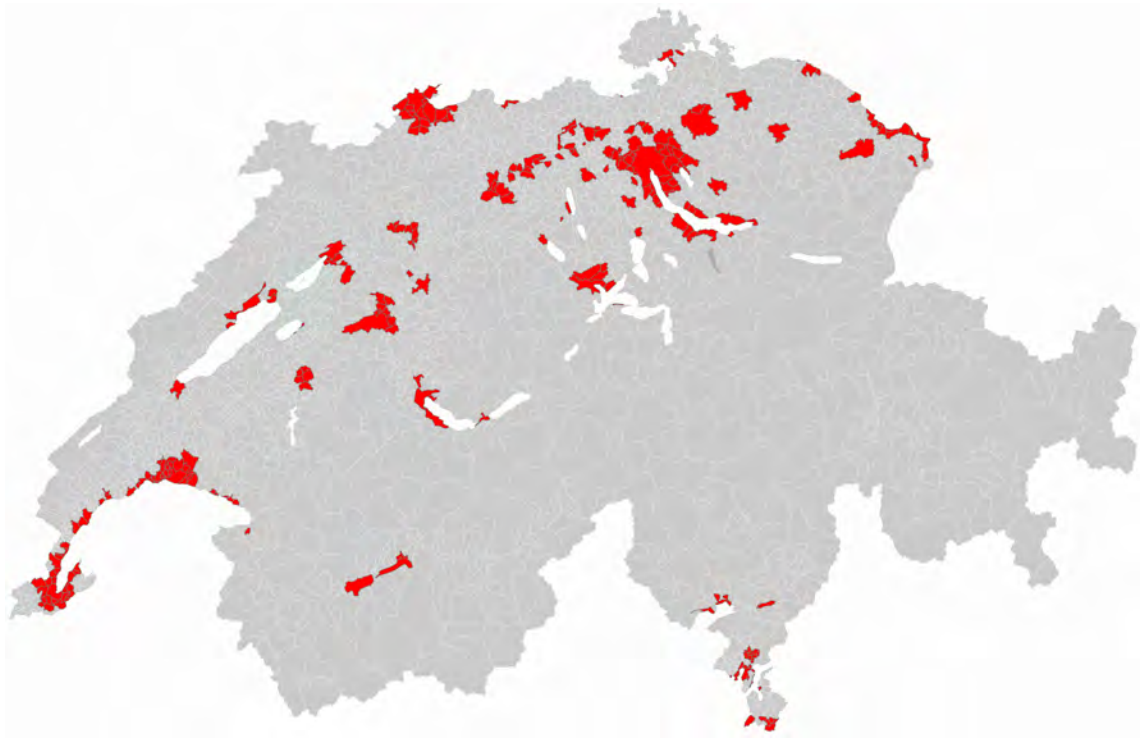
Figure 4.2: The map shows the Swiss communities with at least 30 % settlement
area (red areas).
.

The observations and the hail detections by the algorithms have to be filtered in order to be sure about a hail event. The observations have been filtered according to a threshold of 45 dBZ (median of the neighbourhood) in advance. There would be too many false reports among the crowd-sourced data if no filter were applied. False reporting can be sent due to testing, hate demonstration, disregards of the temporal and spatial importance and mistakes while customising the time and location. Additionally, only the radar detections were considered, which fall in the time period between 07:00 Universal Time Coordinated (UTC hereafter, 09:00 local time) and 19:00 UTC (21:00 local time). The ground observations are not filtered for the diurnal cycle. This is an adaptation to people's diurnal activities. The most people sleep over night, which would lead to many unreported hail events. These spatial and temporal restrictions guarantees a high potential for reporting / observing hail throughout the communal area. Furthermore, the user are better spread over the communal area during the day than during night. Furthermore, only POH pixel equal or larger than 80 % are counted as radar detections. The reporting location and the radar pixels position (midpoints) have to be inside the communal borders in order to be noted and transferred to the verification contingency table. A contingency table is computed for each community. The overall contingency table is calculated by the summation of the "communal" contingency tables. The overall POD, FAR and CSI is then computed from the overall contingency table.

4.6   Hail Size (HS)

The MESHS algorithm is a measure for hailstones larger than 2 cm. Smaller hailstones are more frequent and better to measure as hail studies based on hailpads have shown (*Changnon Jr*, 1971; *Fraile et al.*, 1992; *Giaiotti et al.*, 2001; *Fraile et al.*, 2003; *Sánchez et al.*, 2009). An extension of the MESHS would not only allow the detect the most frequent hailstone sizes, the detection of smaller hailstones allows to estimate the frequency distribution curve of all hailstone sizes in Switzerland. Furthermore, the most crowd-sourced hail reports in this study are the smaller hailstone classes such as the coffee bean and 1 frank coin size classes. The POH is developed as an estimator of the presence of hail on ground regardless the size. *Delobbe et al.* (2005) and *Saltikoff et al.* (2010) showed that all hailstones larger than 2 cm were measured when the POH was at least 80 % (validated on crop damage data). According to *Löffler-Mang et al.* (2011) smaller hailstones, soft hail and graupel were measured at lower POH values (20 to 50 %) by the hail

sensors. The POH might to be a good estimate for the extension of the MESHS.

The Hail Size algorithm (HS) is an extension the MESHS with a transformed POH. The idea came from Alessandro Hering from MeteoSwiss. The POH is converted from [%] scale into [mm] scale. The transformed POH is added where MESHS is lower than 20 mm. The HS is considered as an alpha prototype of a hail detection algorithm, which measure the full range of hailstones. The range of POH values between 0 and 60 % are linearly transformed to a range from 0 to 10 mm. The other values from 60 to 100 % are converted to a range from 10 to 20 mm. The 80 % threshold for the 2 cm was not chosen because it is the minimum POH value for hailstones of 2 cm. It does not mean that 2 cm hailstone are observed at POH value higher than 80 %.

The figure 4.3 shows the correlation of POH and HS computed out of the radar data on the 6th and 7th of June 2015. 31'582 out of 116'517'983 pixels contained POH larger than 0 % and are used for the correlation of HS and POH. The relation implies trimodal regimes (shown by the different coloured regressions in figure 4.3). Two of them are created by the intended computation of the HS for the values smaller than 2 cm. The third is probably a non-linear relation between POH and MESHS. A linearity with high heteroscedastic variance can be as well the case. There is a gap between MESHS (the values equal and larger than 20 mm) and the POH-HS regression below.
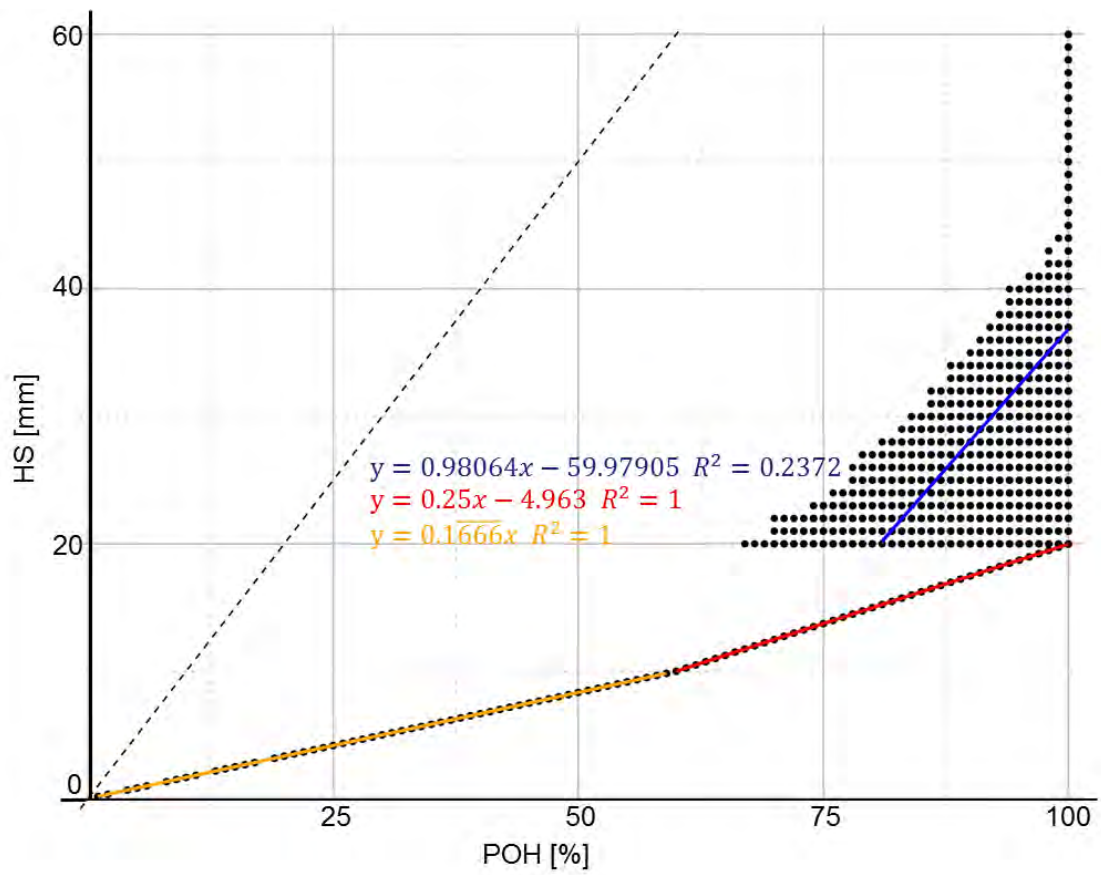
Figure 4.3: Relation of POH and HS on the 6th and 7th of June 2015. Radar pixel values of POH are plotted against HS pixels. The dashed line shows a theoretical perfect correlation.

5   Results

5.1   Case Study: 6th of June 2015

On the 6th of June 2015 several thunderstorms formed over the Prealps in the east of Thun around 15:00 UTC (17:00 local time). The wind in the higher troposphere blew northerly over the Prealpine region. Whereas the alpine pumping transported warm humid air towards the Alps within the boundary layer. At 15:50 UTC the lightning activity reached a first peak and hail was detected by the radars (in the major and split cells). The hailstorm moved towards the city of Thun and turned then in the direction of Bern. On the way to Bern the multicellular hailstorm reached the largest extension and intensity. Before the city of Bern the severe thunderstorm lost its intensity. Then the storm got again intense and started to turn westwards while a thunderstorm cell was emerging over the city of Bern. At 18:50 UTC the thunderstorm collapsed fully.

Two time series of 5 minute HS sequences with the nearest match approach and the best match match approach are mapped in the figure 5.1 and figure 5.2. The hail reports can be assigned to different time periods which depends on the approach. For instance there is a 5 frank coin report in the Northwest of the storm at 16:35 UTC when the nearest match approach is used. A single pixel value of 20 mm lies in the 2 km neighbourhood, which is counted as a match. In the 4 km neighbourhood would be pixels with higher values of nearly 50 mm available, which seem to be the hailstones reported by the user. The best match approach assigns the 5 frank coin to the 16:45 UTC time period. The 2 km neighbourhood has now a median which is closer to the aimed diameter of 31 mm.

There is a 5 frank coin size report counted as a match at 17:00 UTC in the southeast of the storm when computed by the nearest match approach. The best match approach leads to a match at 16:55 UTC. The median of the neighbourhood is smaller at 17:00 UTC than at 16:55 UTC. A nearest match of a 1 frank coin in the southeast is assigned to 16:50 UTC. The best match approach shifts the 1 frank coin report to 16:40 UTC, when the median is closer to the aimed 23 mm diameter. The hail sensor measured an impact with 9.3 mm at 16:31 UTC. There are radar pixels in the 4 km neighbourhood with similar values, but both approaches would lead to matches at 16:50 UTC.
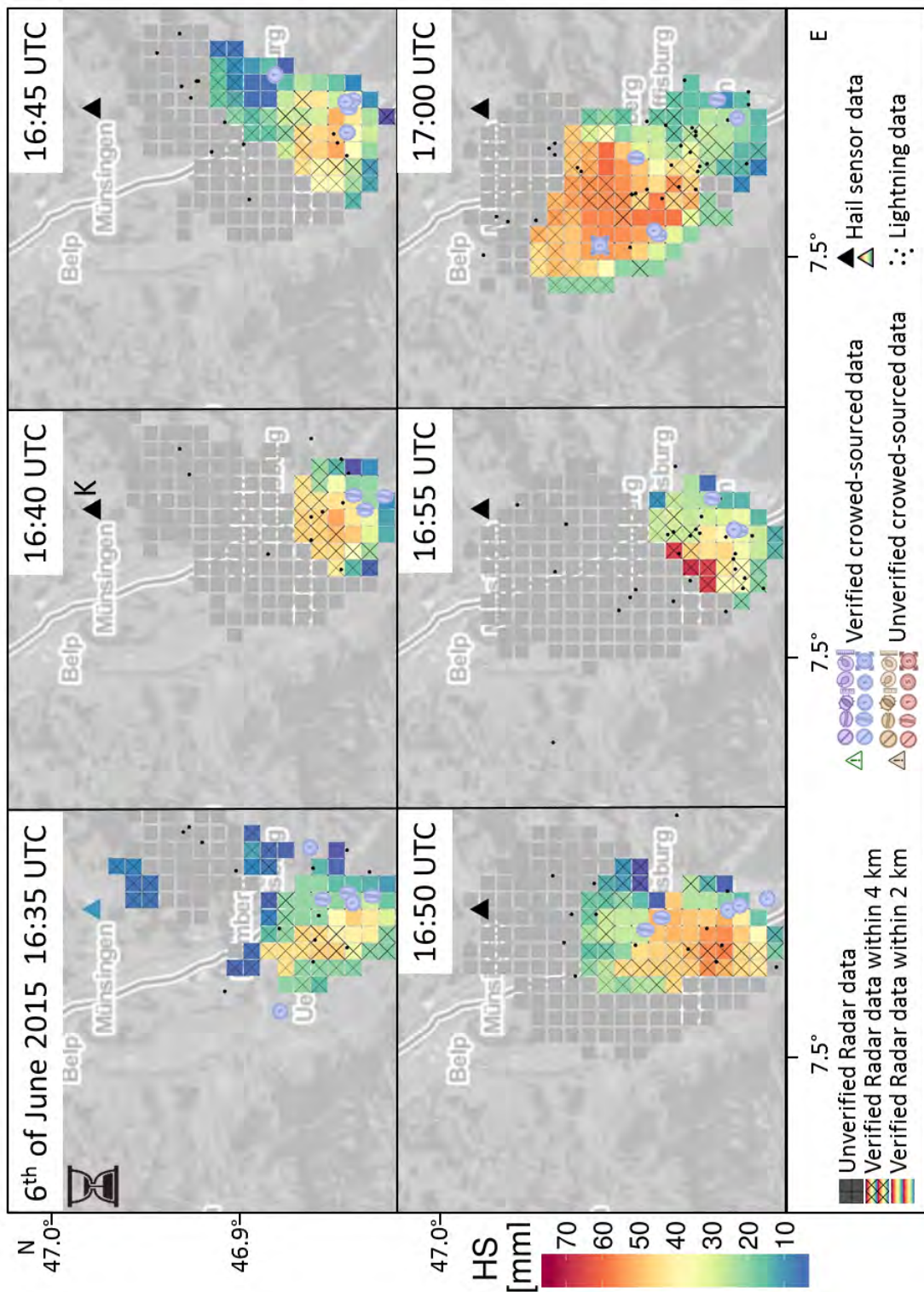
Figure 5.1: Time series of maps showing the HS, hail sensor measurements and the crowed-sourced data for the 6th of June 2015. The hail reports are linked to time derived by the nearest match approach. K represents the hail sensor station in Konolfingen. The time of the hail sensor station is rounded to the next five minutes.
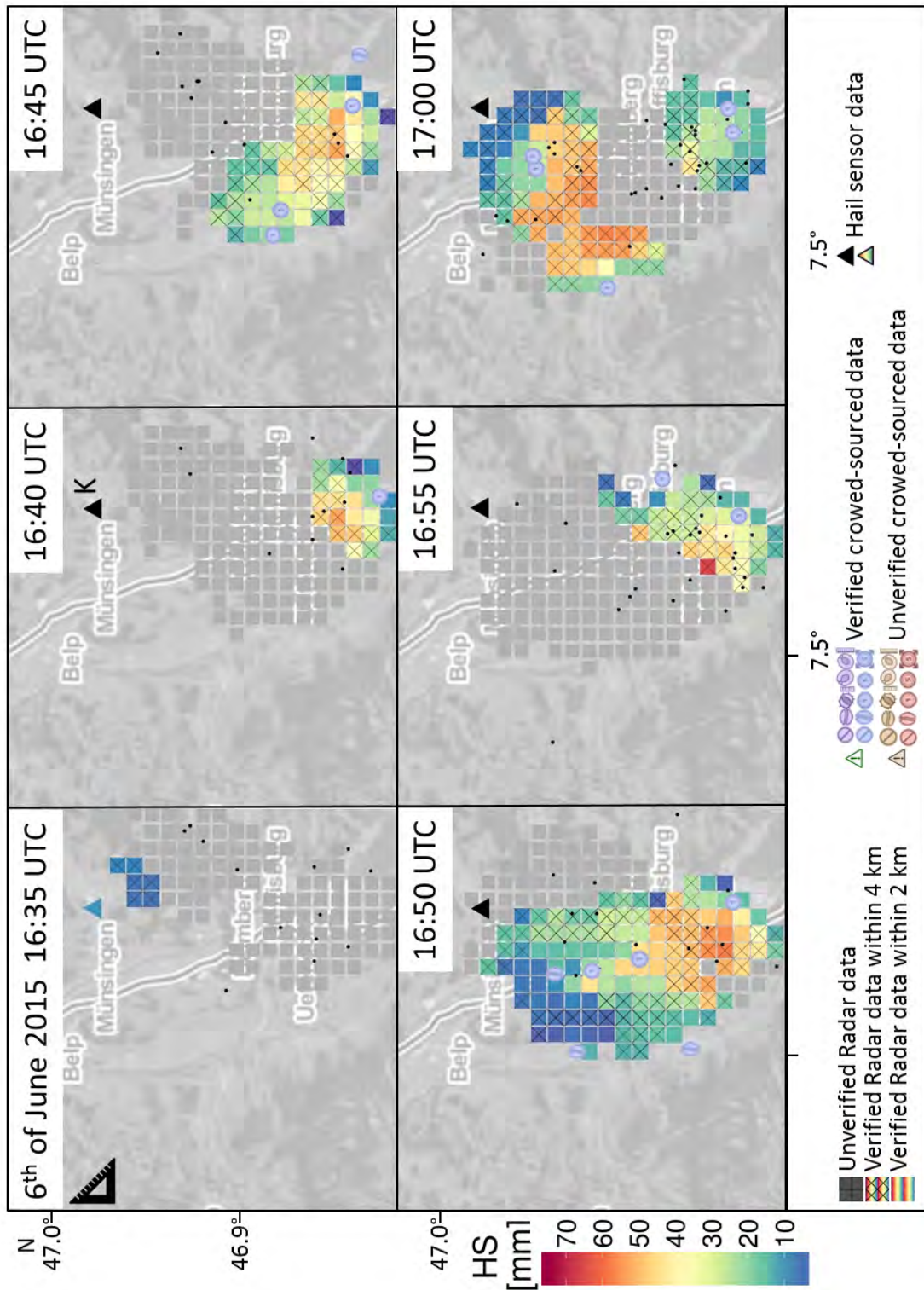
Figure 5.2: Time series of maps showing the HS, hail sensor measurements and the crowed-sourced data for the 6th of June 2015. The hail reports are linked to time derived by the best match approach. K represents the hail sensor station in Konolfingen. The time of the hail sensor station is rounded to the next five minutes.

In figure 5.3 the daily maximum values of HS, hail sensor measurements and the crowed-sourced data is plotted for the hailstorm on the 6th of June 2015. The matched large hail classes of the crowd-sourced data are located in areas with high HS with the exception of one hail report (in the north of Bern). Several 5 frank coin and larger than 5 frank reports were sent in the Gürbetal where the storm reached its highest intensity. The core of the high intensity region lies over a forest, hilly region where the population density is very low. The first peak near Thun has been reported by two 5 frank coin size observations. The third HS peak has 4 large hail size reports nearby (some in the city of Bern).

The reports with smaller reported sizes outnumber the larger ones and are spread over large areas of the area detected by the algorithms. The daily max values of the radar-based hail detection algorithms are too high for the smaller reported hailstone size classes. A five minute sequences are better for the representations of the smaller classes. The unverified reports within radar-based hail detection regions are not counted as matches, because they do not fall in the time window of 45 minutes (for quality reasons). Many warnings verified by users did warn for time periods which are outside the 45 minute time window. They can still be used as indication of hail on daily basis, but not for the verification of the hailstone size. There are unverified reports outside the radar detected hail regions. Some of the unverified reports lie in areas with dBZ values of 25 and higher (compare figure 5.3 with figure 5.4), but with low lightning activity. The hailstorm is captured from Thun to Bern by verified and unverified crowd-sourced data. In region with unverified radar HS the population density is only very low and therefore a few potential reporters are available.

The hail sensor station of Konolfingen measured once a hailstone diameter of 9.25 mm. The station is on the right-hand side of the hailstorm track at 16:31 UTC. The station Schüpfheim registered two impacts, 9.3 mm at 17:31 and 16.95 mm at 18:56, from a less intense thunderstorm moving over the Canton of Obwalden, Nidwalden and the southern part of the Canton of Lucerne. There was a hail detection in the 4 km neighbourhood. The HS did not detect any hail in the neighbourhood of the hail sensor in Entlebuch, although the hail sensor measured hail. The pixels in the neighbourhood reaches Max Echo values up to 50 dBZ. Whereas, the lightning activity is relative low.

The two WRF model outputs, the Morrison double-moment scheme and Thompson scheme, showed an earlier hail occurrence for a similar thunderstorm in the re-
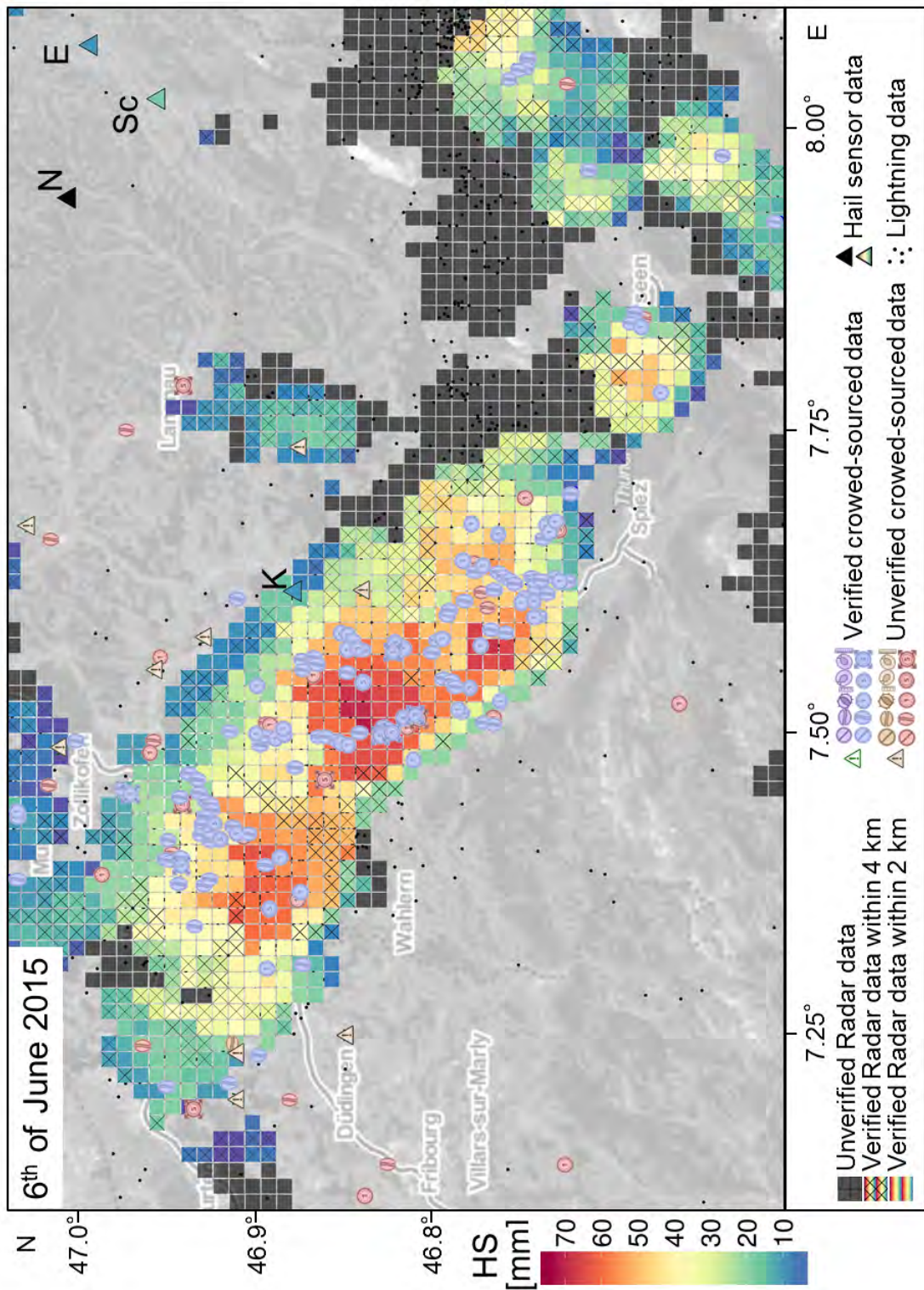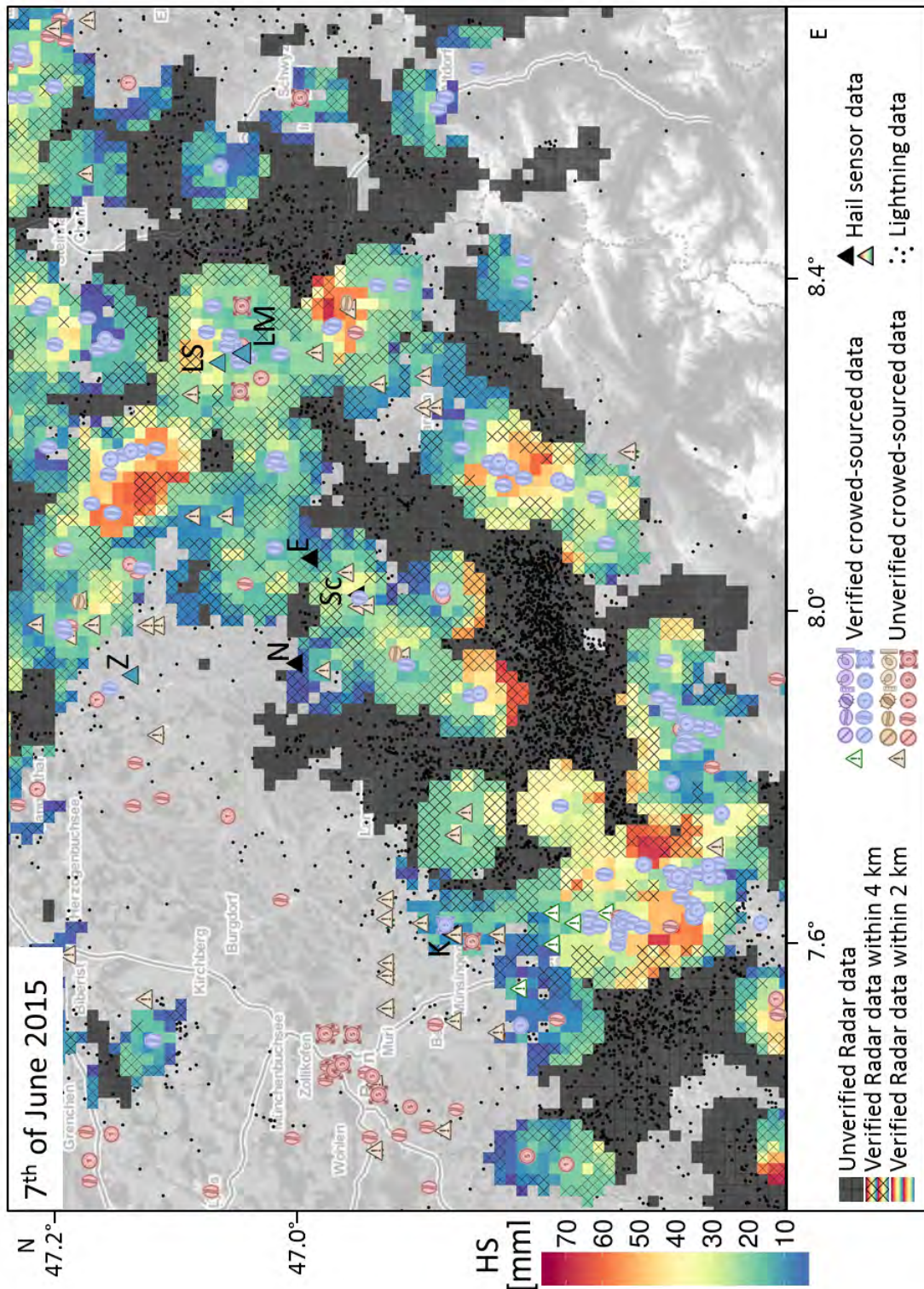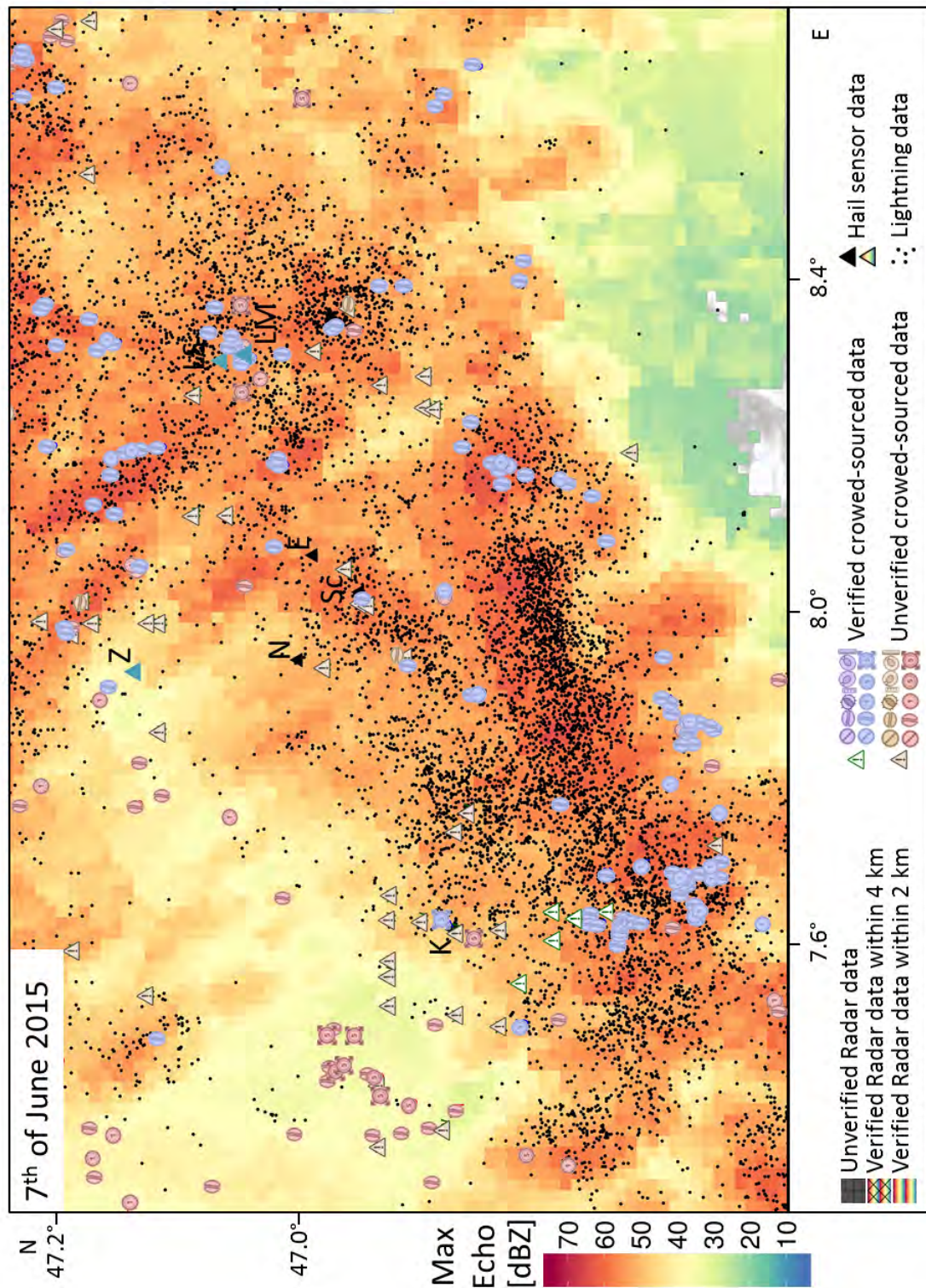
Figure 5.3: Map showing the daily max HS, hail sensor measurements and the crowed-sourced data over the region of Bern and Thun for the 6th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim and E Entlebuch.
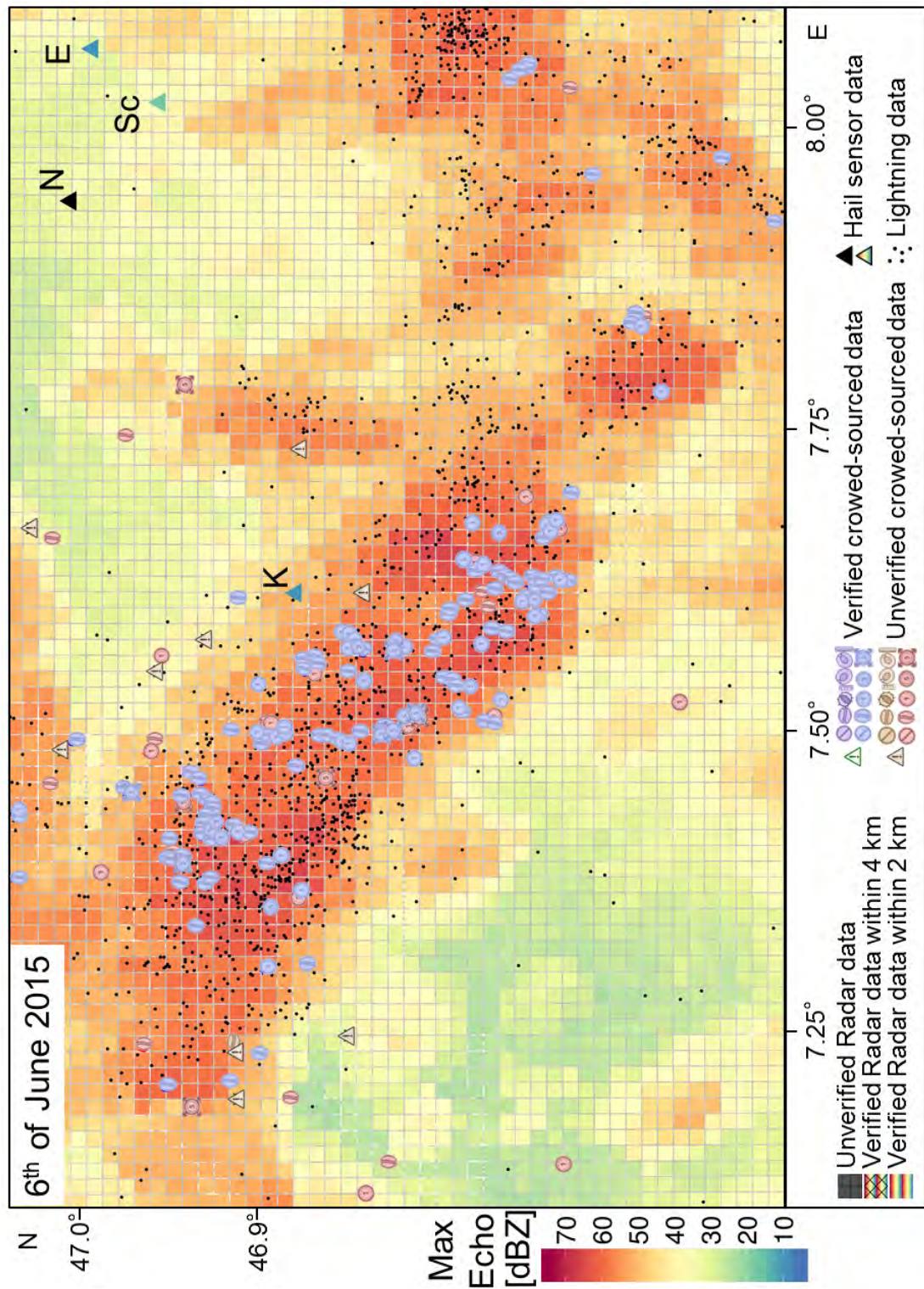
Figure 5.4: Map showing the daily max Max Echo, hail sensor measurements and the crowed-sourced data over the region of Bern and Thun for the 6th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim and E Entlebuch.

gion of Thun (see figure 5.5 and figure 5.6). The Morrison double-moment scheme leads to two small regions of hailstones which reach sizes up to 30 mm. The area with the largest hailstone diameters corresponds to the one detected by the HS/MESHS. The Thompson scheme creates hailstone up to 30 mm for larger areas, but the core with large hailstones is again close to the one of HS.
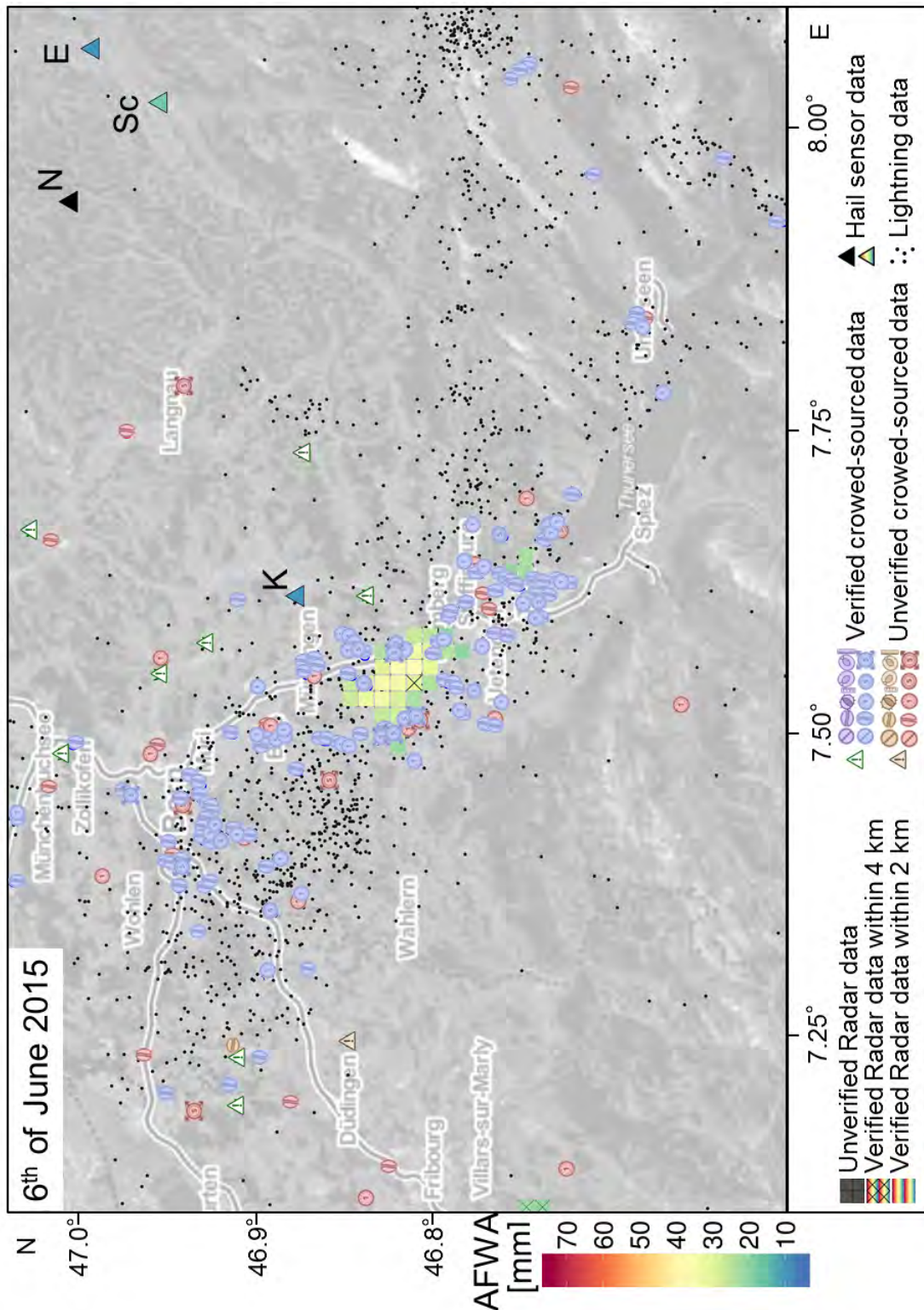
Figure 5.5: Map showing the daily hail from the Morrison double-moment scheme, hail sensor measurements and the crowed-sourced data over the region of Bern and Thun for the 6th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim and E Entlebuch.
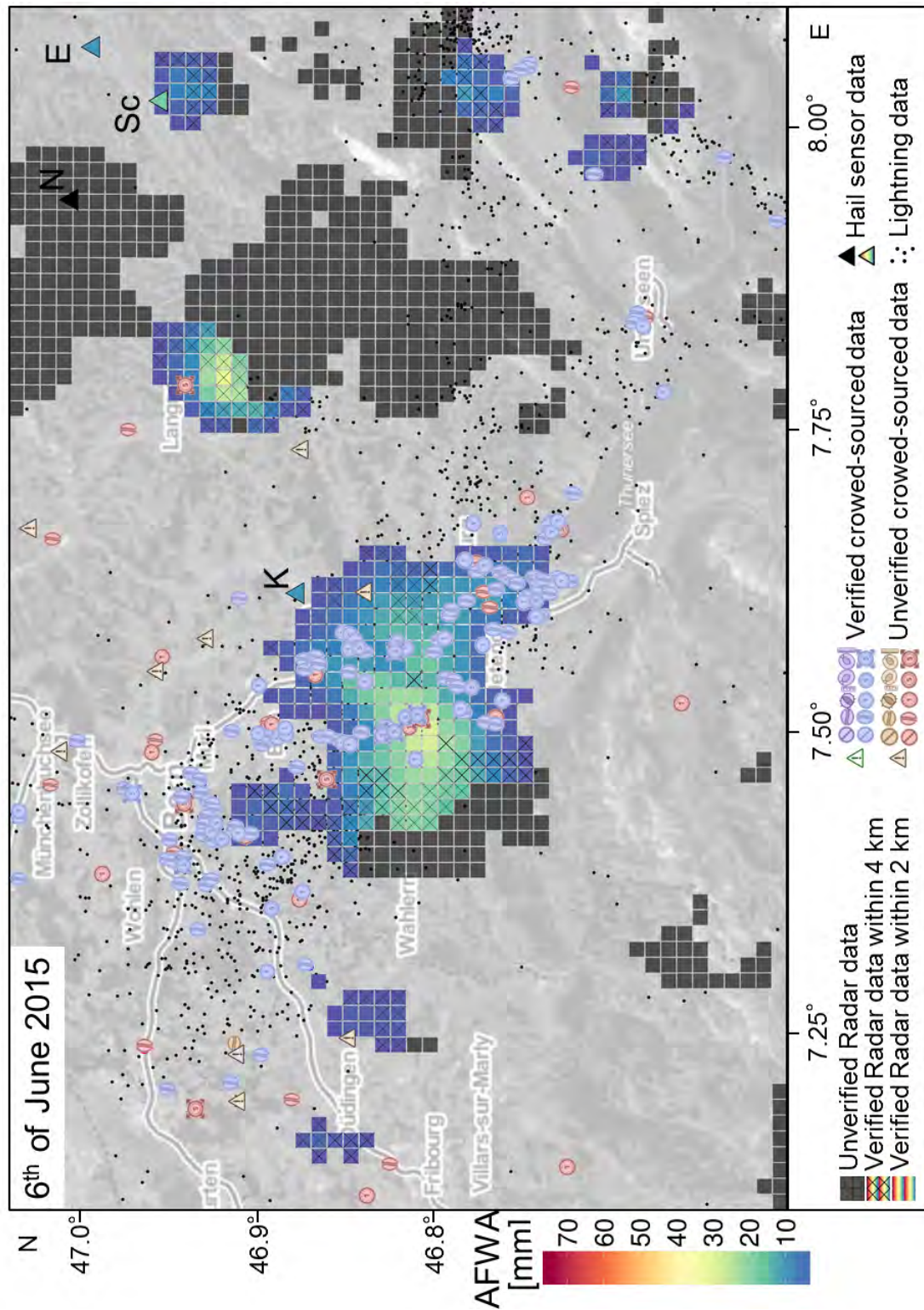
Figure 5.6: Map showing the daily hail from the Thompson scheme, hail sensor measurements and the crowed-sourced data over the region of Bern and Thun for the 6th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim and E Entlebuch.

5.2   Case Study: 7th of June 2015

On the 7th of June 2015 thunderstorms were fromed in the South of the Canton of Vaud and Canton of Fribourg. The thunderstorms moved in connection to a cold front in northeasterly direction. A thunderstorm cell formed in the northern outskirts of Thun at 14:30 UTC. The lightning activity peaked quickly after the first rainfall and the radar detected hailstones at 14:35 UTC. Further hailcells formed over Thun and its surroundings. The hailstorms remained an hour over Thun before they started to move on in northeasterly direction. At 18:10 hailstone were detected by the radar over Lucerne. The cold front from the southwest met a cold bise which blew from the northeast over the central Switzerland. The resulting occlusion forced the warm air between the cold fronts to rise. This led to stationary intense convective cells producing large hailstone and large amount of rain. Later on, the thunderstorm crossed the Lake of Constance and decayed in the night over Germany (Bavaria).

In figure 5.7 shows the daily maximum HS values, hail sensor measurements and crowd-source data in the Napf-region for the 7th of June 2015. Hailstones were detected in a band from southwest to northeast. There are clusters of verified reports in Thun, Spiez, Interlaken, Giswil, Lucerne and around Sursee. Those reports indicate hailstone sizes similar to coffee bean and 1 frank coin sizes. Most of the verified crowd sourced hail reports are coffee bean and 1 frank coin sizes. A 5 frank coin size report and a larger than 5 frank coin size report are located between Bern and Thun. There are a couple of unverified reports and confirmed warnings in the city and surroundings of Bern. 7 unverified reports claim hailstone diameters of 5 frank coin and larger than 5 frank coin sized hailstones. The lightning activity was low in the area of Bern and the Max Echo with between 30 and 35 dBZ is moderate (see figure 5.8). On the way from Bern to the hail sensor station of Zell, reports of coffee bean and 1 frank coin size were sent, where the dBZ is not higher than 10 dBZ. The area containing verified crowd-sourced data has dBZ values which ranges from 45 dBZ up to values larger than 60 dBZ.

The four hail sensor stations Konolfingen, Napf, Entlebuch and Schüpfheim did not register any hail impact, even though the radar detected hail above these stations. Moreover, users did report hailstones in Konolfingen and Schüpfheim. The hail sensor station in Zell measured 16 impacts ranging from 4.9 to 10.5 mm. The first impact was registered at 18:15 UTC and the last hit occurred at 21:22 UTC. The first 12 impacts happened within 37 minutes. The hail sensor station in Lucerne Moosstrasse registered three hits at 18:35 UTC with 10.3 mm, 18:38 UTC

with 11.2 mm and 18:46 UTC with 10.6 mm. There are two hail sensor station in Lucerne Sedel next to each other (LS1 and LS2). 8 impacts ranging from 5.9 to 11.8 mm were measured by the parallel hail sensor station of Lucerne Sedel (LS2). The first hit was registered at 18:33 UTC and the last at 19:13 UTC.

Figure 5.7: Map showing the daily max HS, hail sensor measurements and the crowed-sourced data over the Napf-region for the 7th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim, E Entlebuch, LM Lucerne Moosstrasse and LS Lucerne Sedel with its two parallel hail sensor station.

Figure 5.8: Map showing the daily max Max Echo, hail sensor measurements and the crowed-sourced data over the Napf-region for the 7th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim, E Entlebuch, LM Lucerne Moosstrasse and LS Lucerne Sedel with its two parallel hail sensor station.

5.3   Case Study: The Neighbourhood and the Hail Sensor

The second hail sensor station of Lucerne Sedel (LS1) detected the first hailstone impact at 18:25:38 UTC and the last at 19:13 UTC (see figure 5.9 a) and b)). There were 20 impacts, which differs from amount of impacts at the parallel hail sensor station. The figure 5.9 a) and b) show the time on the x-axis, the measured diameters of the hail sensor station of Lucerne Sedel indicated by the colourbar and the median of the neighbourhood is indicated on the y-axis for each impact for the nearest and best match approach. In the figure 5.9 c) you can see a time series of maps with the HS pixels in the related 2 and 4 km neighbourhood of the hail sensor station in Lucerne Sedel.

The radar algorithms POH and HS detected already hail at the time 18:25 UTC. The hail sensor station measured an hail impact with 7.2 mm in the 5 minute later time period. The nearest match approach takes the median at 18:30 UTC, whereas the best match approach takes the lower median at 18:25 UTC. The measured hailstones at 18:35 UTC, 18:40 UTC, 18:45 UTC and 18:50 UTC are assigned to higher median when computed with the nearest match approach. The best match approach leads to matches with smaller median values. Actually, there are pixel with adequate HS values in the southeastern part of the 2 km neighbourhoods with the computation by nearest match approach.
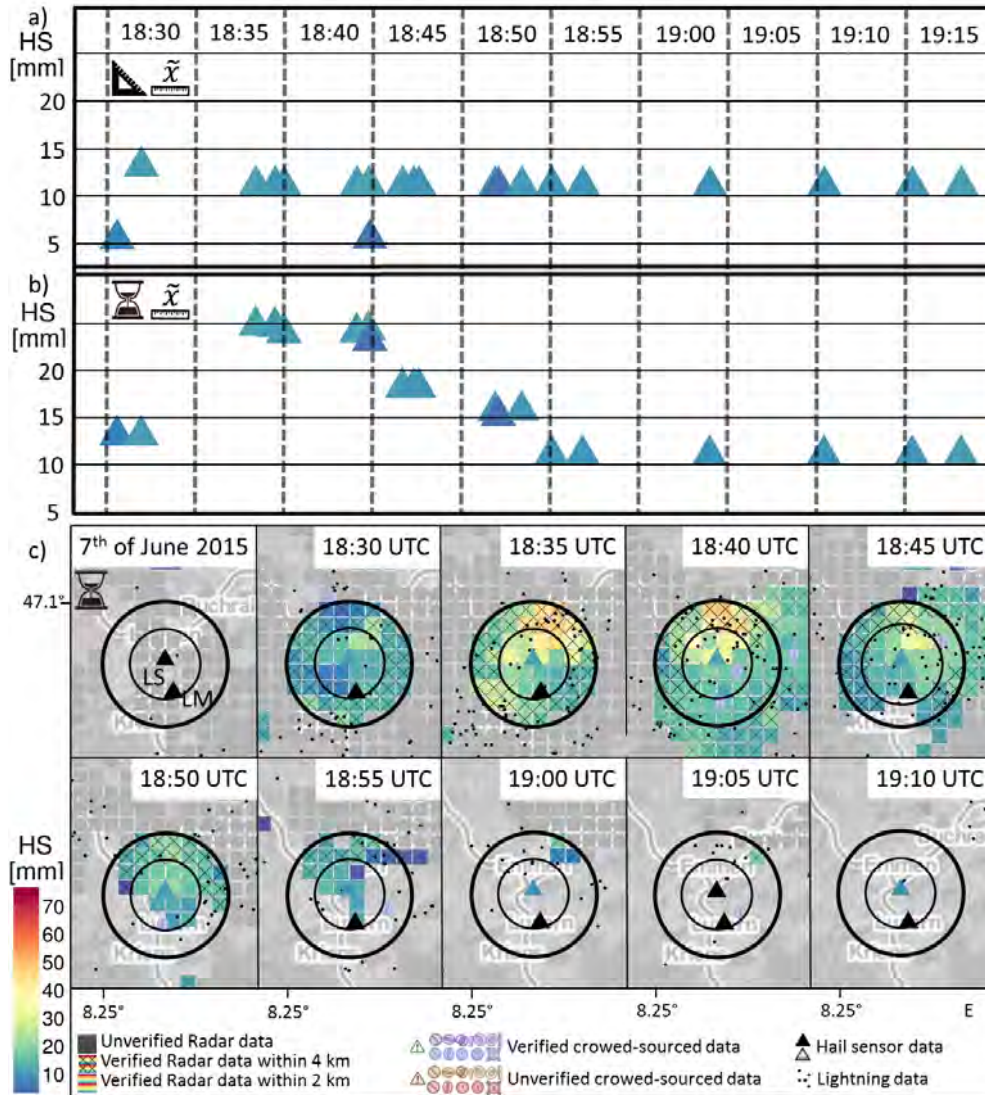
Figure 5.9: The measured hail diameter by the hail sensor station of Lucerne Sedel (LS1) on the 7th of June 2015 are computed with the best match approach in a) and with the nearest match approach in b). In the panel c), maps showing 5 minute HS values, the neighbourhood of the hail sensor station of Lucerne Sedel, the hail sensor measurements and the crowed-sourced data over the Lucerne. The following hail storm stations represented by: LM Lucerne Moosstrasse and LS Lucerne Sedel with its two parallel hail sensor station.

5.4 MESHS Verification with the MeteoSwiss App Data

The MeteoSwiss crowd-sourced data has been compared to the MESHS. 4'117 out of 26'733 could be matched to MESHS values in the neighbourhood (see table 5.1). The different reported sizes have varying fractions of matches. The 1 frank coin reaches the highest fractions of matched reports with 26.4 %. The coffee bean has the second largest fraction of matches with 23.1 %. The larger than 5 frank coin shows the lowest fraction of matched reports. A large part of the mismatches were sent when the sky was clear or when the radar did detect signals only for light rain.

Table 5.1: Matches and mismatches of the MeteoSwiss crowd-sourced data when compared to MESHS. The absolute and relative values are shown. The percentages are the ratios among the reported size classes.

	Matches #	Matches %	Mismatches #	Mismatches %	Total #	Total %
⊘ *	260	2.9 %	8'549	97.1 %	8'809	100 %
🫘	2'586	23.1 %	8'607	86.9 %	11'193	100 %
①	1'045	26.4 %	2'909	73.6 %	3'954	100 %
⑤	161	16.8 %	798	83.2 %	959	100 %
⑤	65	3.6 %	1'753	96.4 %	1'818	100 %
Total	4'117	15.4 %	22'616	84.6 %	26'733	100 %

The median of MESHS values increases from the coffee bean to the 5 frank coin size with the nearest matche and best match approach (see figure 5.10). The larger than 5 frank coin size follows the increasing median trend only with the best match approach. The MESHS median of the 5 frank coin is quite similar to the diameter of the 5 frank coin in the nearest match approach. The median of the 5 frank coin and 1 frank coin gets close to the real diameter with the best match approach. The MESHS of the upper quartile gets larger from the coffee bean to the larger than 5 frank coin size in both approaches. The MESHS of the lower quartile increases with larger reported sizes only with the best match approach. Furthermore, the probability densities (indicated by the violin plots) of the larger than 5 frank coin size class shows a bimodal distribution of MESHS (around 23 mm

and 42 mm). The maximum in the distribution around 42 mm is more prominent in the best match approach, but it occurs as well in the nearest match approach. All densities are cut on the 20 mm MESHS side due to the limitation of MESHS. Whereas the densities smoothly decrease on the upper MESHS side. Moreover, the probabiliy densities of the no hail, coffee bean and 1 frank coin categories seem to be compressed next to 20 mm threshold.
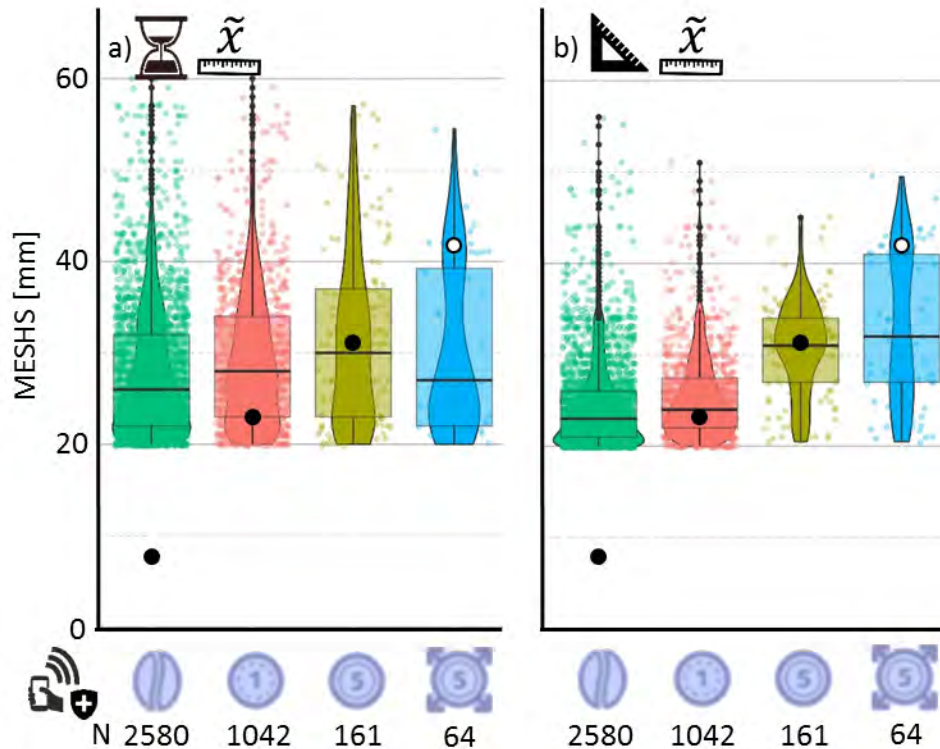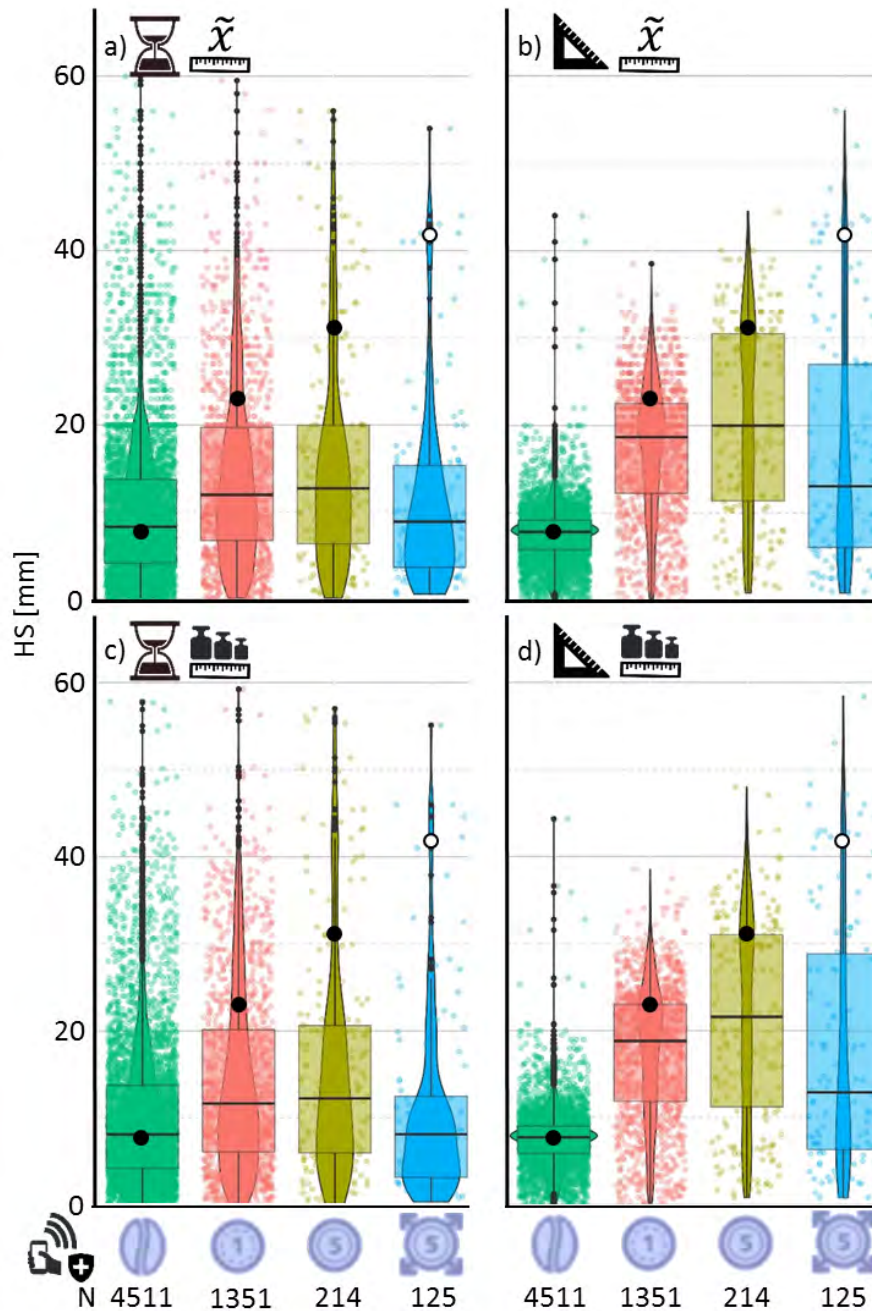

Figure 5.10: Boxplot of the MESHS median (top) respectively IDW (bottom) by the reported sizes of the MeteoSwiss crowd-sourced data (only the matches are shown). In the panel a), the median of the neighbourhood was computed by the nearest match approach. In the panel b), the median of the neighbourhood was compared by the best match approach. Violin plots with kernel probability density and scatter plots are added to show smother distributions. The black dots represent the reportable sizes. The white dot represents a fictive value which was chosen for the larger than 5 Swiss frank coin size class in the best match approach.
.

The time lags of the MESHS matches among the MeteoSwiss crowd-sourced data show with both approaches a tendency of negative time lags (see figure 5.11). It has to be mentioned that the tendency of negative time lag (time of the matched radar time - reporting time) with the nearest match approach is forced by the method itself. In the nearest match approach, the matches have no time lag. Whereas the most matches computed by the best match approach are shifted by

a 5 minute period. The 5 minute negative time lag is the second most frequent time lag in the the nearest match approach.
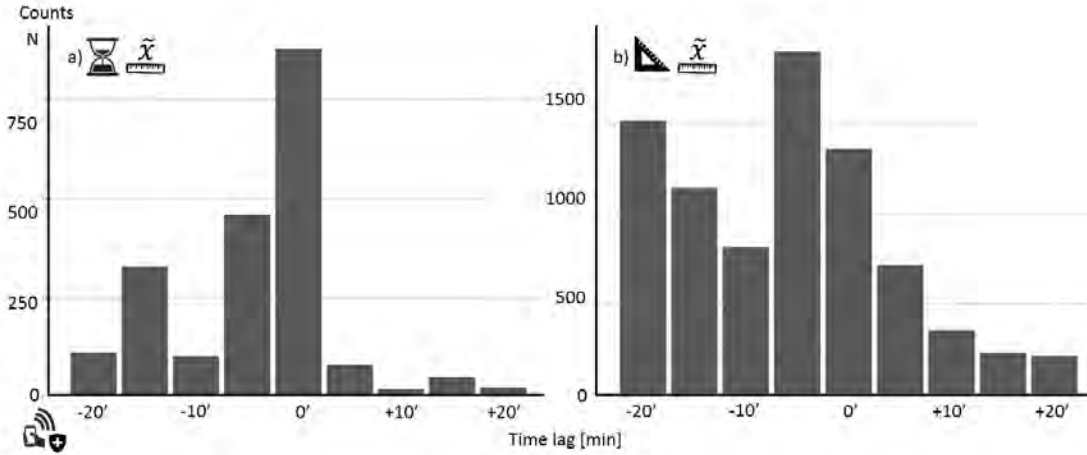
Figure 5.11: Distribution of the time lag (time lag = time of the matched radar time - reporting time) of the MeteoSwiss crowd-sourced data (only the matches). In the panel a), the time lag was derived by using the nearest match approach. In the panel b), the the time lag was computed by the Best match approach.

5.5   HS Verification with the MeteoSwiss App Data

An higher amount of matches are observed among the MeteoSwiss crowd-sourced data with the HS (see table 5.2) than with MESHS (see table 5.1). 27.1 % of the hail reports are classified as matches if HS is used as radar reference (compared to 15.4 % with MESHS). Without the no hail reports, 34.9 % are matched with radar values. The increased amount of matches is noticed throughout all size classes. The coffee bean size class reaches 40.6 % for the fraction of matched reports with the HS. The 1 frank coin class get the second highest fraction of matched reports with a score of 34.3 %. The fraction decreases with increasing reported sizes. The larger than 5 frank coin has the lowest match ratio among all size classes. The radar data was not available fully available over the 8 months. Some files were missing. There were 27'301 reports collected over the 8 months. But 26'733 reports were analysed for the MESHS and 26'534 reports for the HS. Fewer report were analysed when the HS was used, because the HS depends on two products (POH and MESHS).

The median of the neighbourhood medians respectively IDWs increases from the coffee bean to the 5 frank coin size class with both approaches (see figure 5.12). The medians computed with the best match approach increases from the no hail to the 5 frank coin class. The upper quartile increases for all the reported sizes with

Table 5.2: Matches and mismatches of the MeteoSwiss crowd-sourced data when compared to HS. The absolute and relative values are shown. The percentages are the ratios among the reported size classes.

	Matches		Mismatches		Total	
	#	%	#	%	#	%
⊘ *	979	11.2 %	7'767	88.8 %	8'746	100 %
bean	4'511	40.6 %	6'583	59.4 %	11'094	100 %
1	1'351	34.3 %	2'588	65.7 %	3'939	100 %
5	214	22.5 %	739	77.5 %	953	100 %
5+	125	3.6 %	1'677	93.1 %	1'802	100 %
Total	7'180	27.1 %	19'354	84.6 %	26'534	100 %

the exception of the larger than 5 frank coin class. The medians of the 1 frank coin, 5 frank coin and larger than 5 frank classes fall compared to the MESHS far below the set diameter of the reportable sizes. They show an increasing difference between the set diameter and the HS median with increasing size class. However, the median of the coffee bean size class is very close to the set sizes throughout both approaches and median respectively IDW computations. The Inter Quartile Range (IQR hereafter) of the coffee bean size class gets strongly concentrated around the median according to kernel probabilistic distribution. The lower quartile of the coffee bean size is reached with nearest match approach with a HS values of 4 mm. The lower quartile computed with the best match approach has a HS value of 5.9 mm.

The overall time lag distribution of the matches based on HS larger than 0 mm is quite similar to the one based on MESHS equal or larger than 20 mm (compare figure 5.11 and figure 5.13). Exceptional cases are the no hail and larger than 5 frank coin classes when they are computed by the best match approach. The other classes follow the overall distribution pattern. The highest abundance of matches have a time lag of -5 minute period in the best match approach. The second most frequent time lag is at 0 minute time lag. The number of matched reports increases from +20 minute time lag to the -5 minute time lag. The time lags from -10 minute to -20 minute are lowered compared to the -5, 0 and +5 minute time
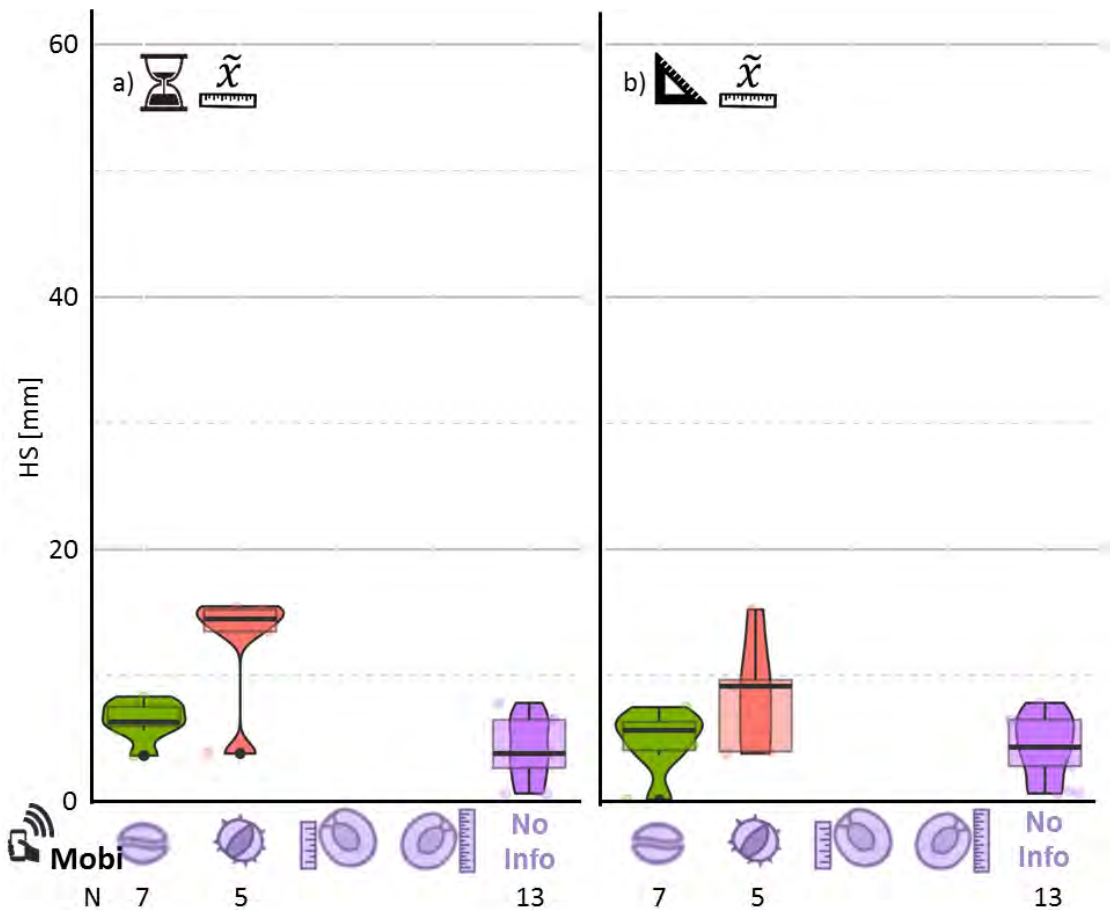
Figure 5.12: Boxplot of the HS median respectively IDW by the reported sizes of the MeteoSwiss crowd-sourced data (only the matches are shown). In the panel a), the median of the neighbourhood was computed by the nearest match approach. In the panel b), the median of the neighbourhood was derivated by the best match approach. The IDW of the neighbourhood was derived according to the nearest match approach in the c). In the panel d), the IDW of the neighbourhood was calculated by the best match match approach. Violin plots with kernel probability density and scatter plots are added to show smother distributions. The black dots represent the reportable sizes. The white dot represents the set value for the best match approach.

lag. The frequency of matches increases again from -10 to -20 minute time lag. The distribution pattern of the time lags derived from matches computed by the nearest match approach show more frequent negative time lags. The most matches have a time lag of 0 minute. The second and third most frequent time lags are the 5 and 15 time lags.
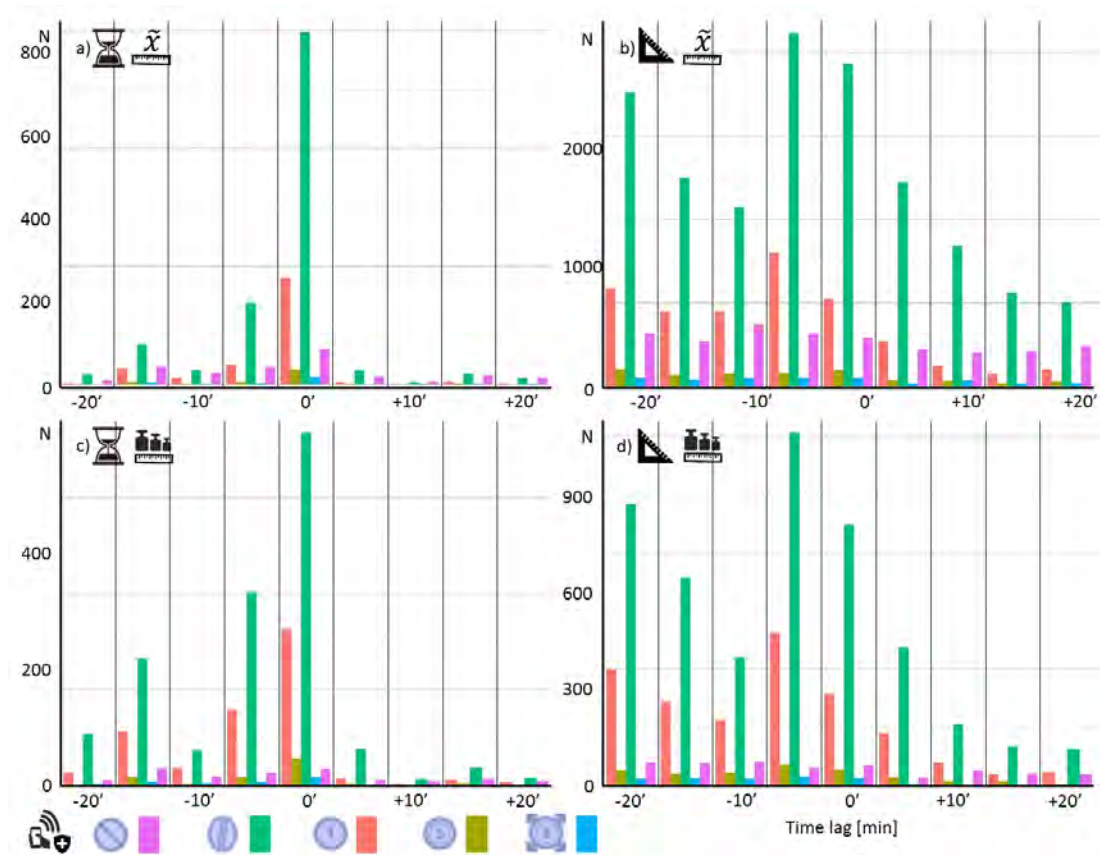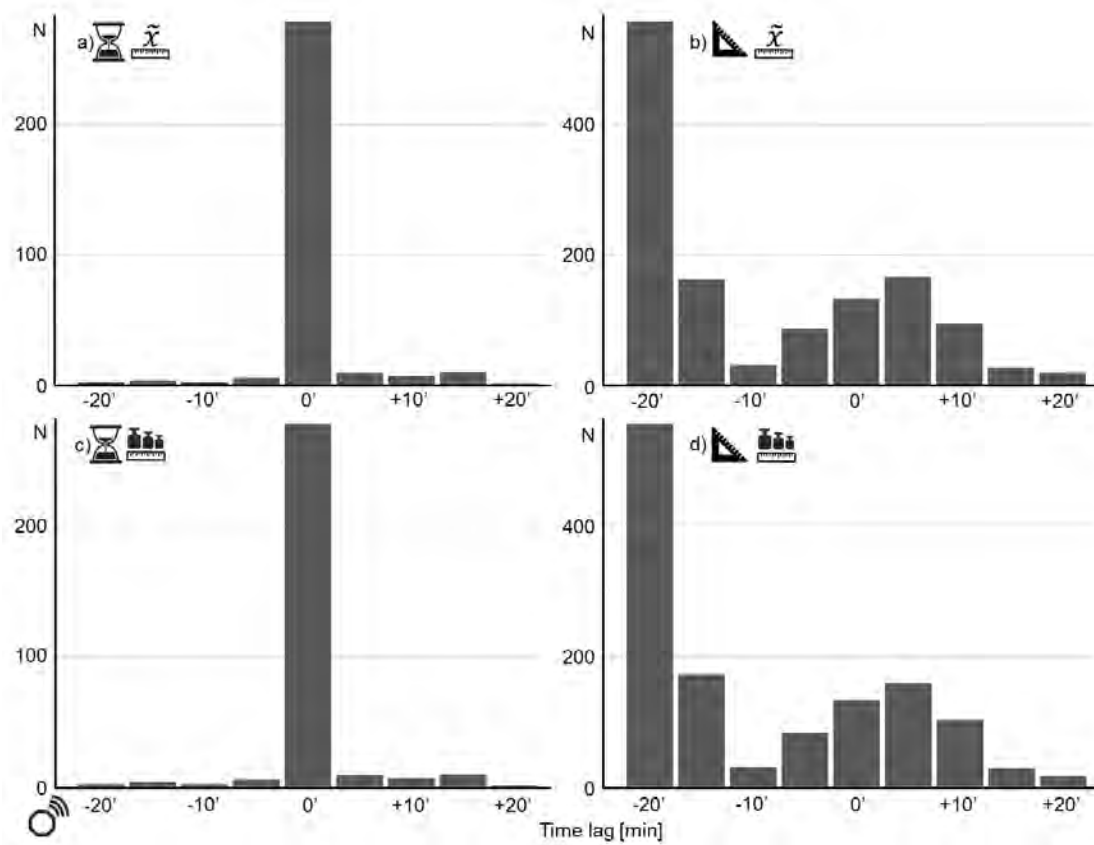

Figure 5.13: Distribution of the time lags of the MeteoSwiss crowd-sourced data (only the matches). In the panel a), the time lag was derived by the nearest match approach. In the panel b), the the time lag was computed by the Best match approach.

5.6   HS Verification with the Mobi App Data

3'393 reports were collected by the Mobiliar crowd-sourcing App (see table 5.3).
2 % of the all Mobiliar App reports could be matched to radar detection. 759
messages have a unusable information (No Info) about the size of the hailstones.
The unusable messages contain 13 matches. 7 matches have been found among
coffee bean size. Further 5 matches belong to the chestnut size class. The limited
number of matches do not guarantee stable boxplot statistics (see figure 5.14).
Therefore, no further statistics has been aimed to do.

Table 5.3: Matches and mismatches of the Mobiliar crowd-sourced data when
compared to HS. The absolute and relative values are shown. The percentages are
the ratios among the reported size classes.

	Matches #	Matches %	Mismatches #	Mismatches %	Total #	Total %
⊘ *	41	2.0 %	1'945	98.0 %	1'986	100 %
🫘	7	2.2 %	583	98.8 %	590	100 %
🌰	5	12.8 %	34	87.2 %	39	100 %
🥚	0	0.0 %	7	100 %	7	100 %
🥚	0	0.0 %	12	100 %	12	100 %
No Info	13	1.2 %	746	98.2 %	759	100 %
Total	66	27.1 %	3'327	84.6 %	3'393	100 %

Figure 5.14: Boxplot of the HS median respectively IDW by the reported sizes of the Mobiliar crowd-sourced data (only the matches are shown). In the panel a), the median of the neighbourhood was computed by the nearest match approach. In the panel b), the median of the neighbourhood was derivated by the best match approach. Violin plots with kernel probability density and scatter plots are added to show smother distributions.

5.7 Correlation of HS and Hail Sensor Data

The measured hailstone diameters by the hail sensor of inNET Monitoring AG range from 0 up to 30 mm (see figure 5.15). The range of the HS computed for the neighbourhoods goes from 0 up to 50 mm. There are measurement series of hailstones from individual hail sensor visible in the plots (the points which build lines). The time intervals of measurements within such measurement series vary from seconds up to minute scale. This results in horizontal lines in the panels a) and b), which represent the computation of the median by the nearest match and best match approach. Some of the horizontal lines become a positive or negative steepness when the IDW is computed. The comparison of the measured hailstones by the hail sensor and the HS in the nearest match approach shows a limited linearity. Whereas in the best match approach has a better correlation.

The distribution of the time lag according to the nearest match approach show that the most matches have a time lag of 0 minutes (see figure 5.16). The frequency of matches with other time lags are very low, but interestingly the positive time lags outnumber the negative time lags. The computation of the median and IDW by the best match approach lead to a total different time lag distribution. Most matches are now linked to HS values 20 minutes prior to the measurement time. A secondary cluster of matched reports are concentrated around the +5 minute time lag.
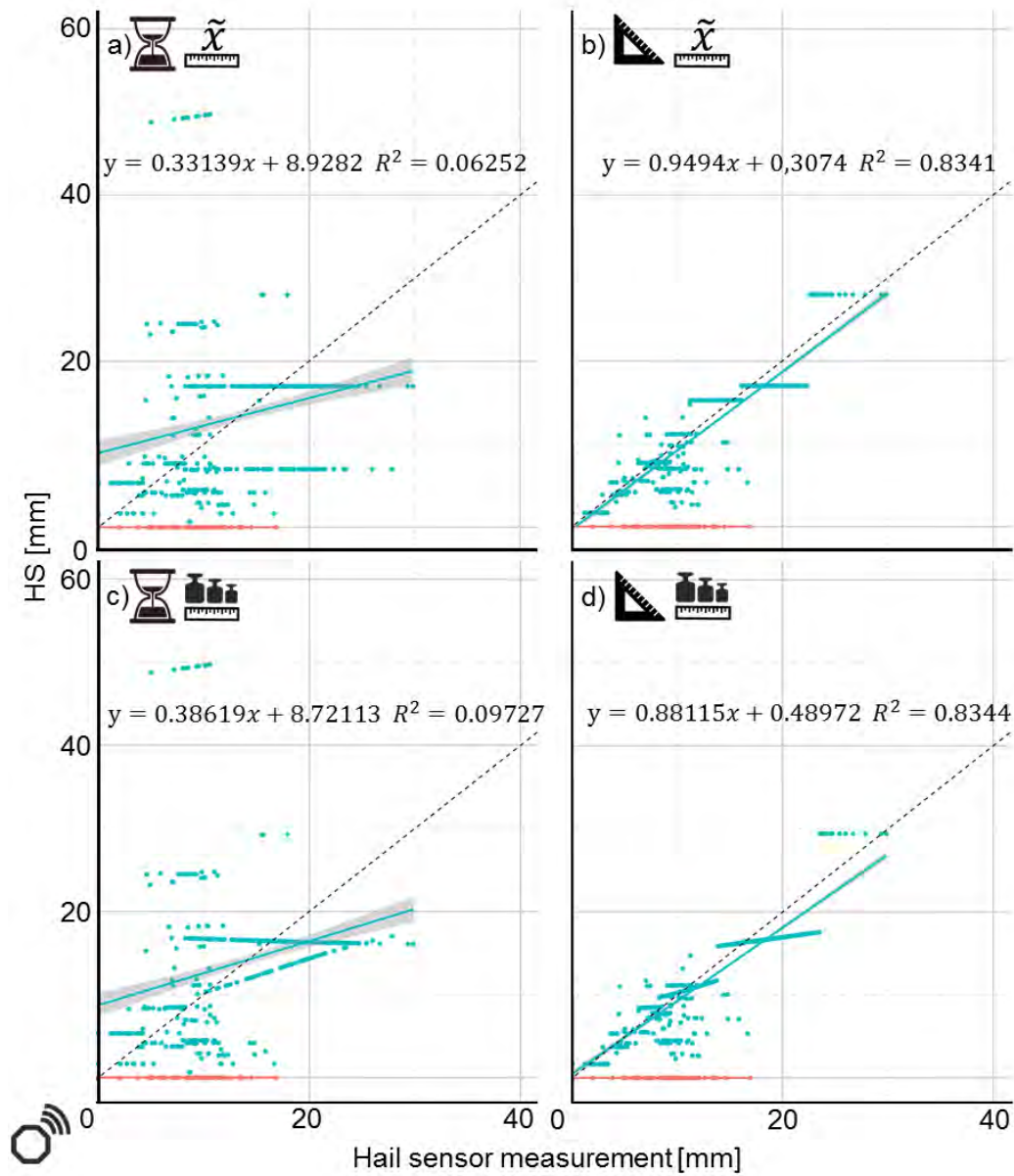
Figure 5.15: Regression of the median and IDW of HS in the neighbourhood on hailstone diameter measurements of all hail sensors. In the panel a), the median of the neighbourhood was computed by the nearest match approach. In the panel b), the median of the neighbourhood was computed by the best match approach. The IDW of the neighbourhood was derived according to the nearest match approach in the panel c). In the panel d), the IDW of the neighbourhood was calculated by the best match approach. The dashed line shows a theoretical perfect correlation.

5.8   Correlation of HS and POH

The HS has been compared to the POH values in figure 5.17. The distribution is similar to the one which shows the computation of HS out of MESHS and POH (compare to figure 4.3). The main difference is that there are values above the two regression lines (for the POH transformation). The mentioned gap between the transformation of the POH and the MESHS values (below 2 cm) is filled up when the median or IWD of the neighbourhoods are used instead of the pixels.

5.9   POH Categorical Verification

In figure 5.18 the overall POD is computed for all Swiss communities with living zone share of at least 30 %. The POD is low in many major cities. Very high POD are achieved in smaller communities in the agglomeration of the major cities or smaller urban communities on the countryside. The FAR is moderate or high in some major cities and as well in most other urban communities (see figure 5.19). Only a minority of towns are not affected by high FAR values. The CSI for the selected communities is shown in figure 5.20. The CSI is high in some lesser populated communities in the agglomeration. The CSI is low in the big cities and in smaller communities in the agglomeration and on the countryside. The number of hail events (hits, misses and false reports) is shown in figure 5.21. The FAR is high in communities with few hail events. The POD is low in many communities where the number of hail events is high. The overall-scores of the POD, FAR and CSI are shown below:

$$Overall - POD = 0.27 \tag{5.1}$$

$$Overall - FAR = 0.38 \tag{5.2}$$

$$Overall - CSI = 0.19 \tag{5.3}$$

Figure 5.16: Distribution of the time lags of the MeteoSwiss crowd-sourced data (only the matches). In the panel a), the time lag was derived by the nearest match approach. In the panel b), the the time lag was computed by the best match match approach.
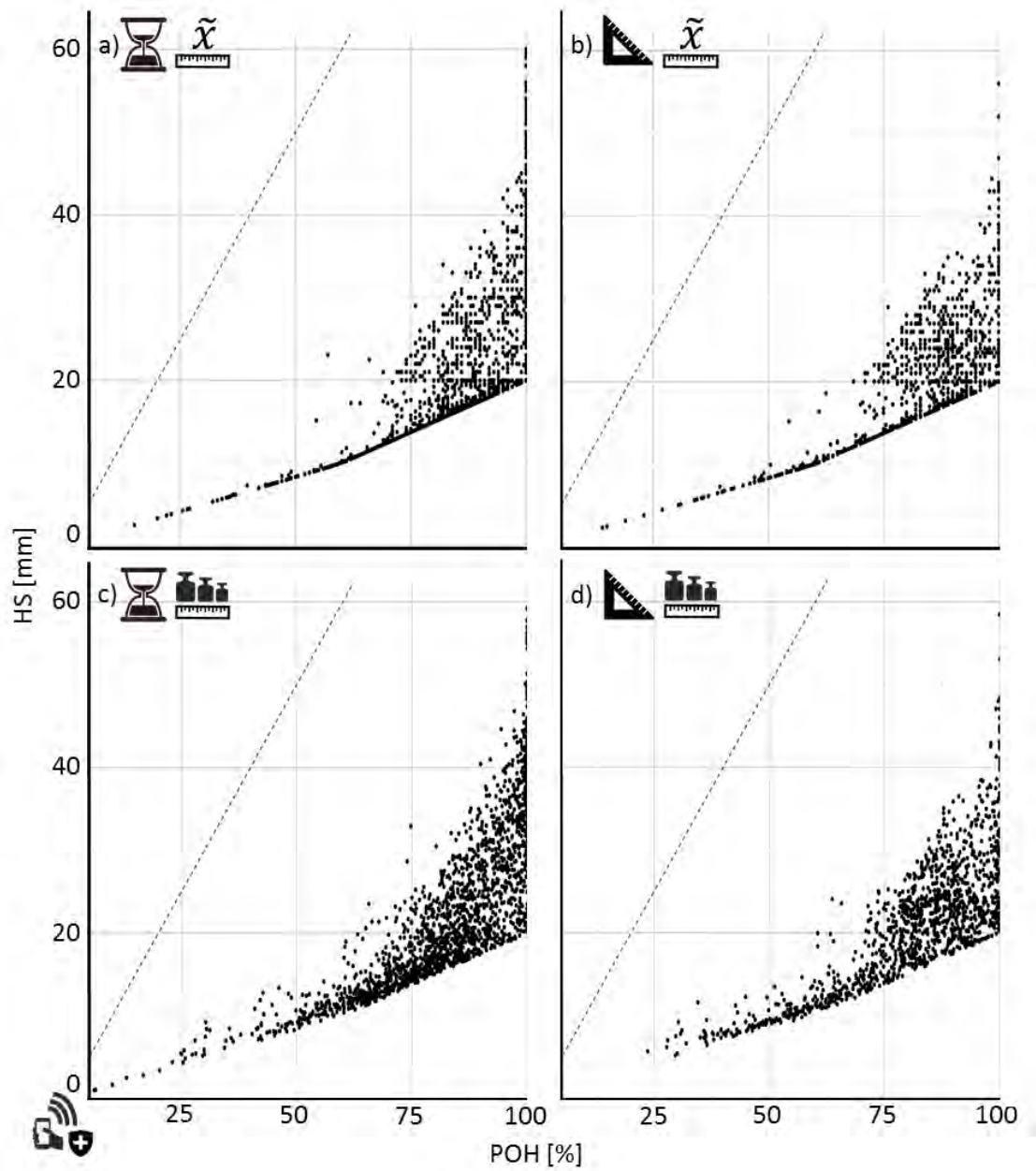
Figure 5.17: Relation of POH and HS in the neighbourhoods of MeteoSwiss crowd-sourced data. The median and IDW of POH and HS has been computed for the matches. The dashed line shows a theoretical perfect correlation.
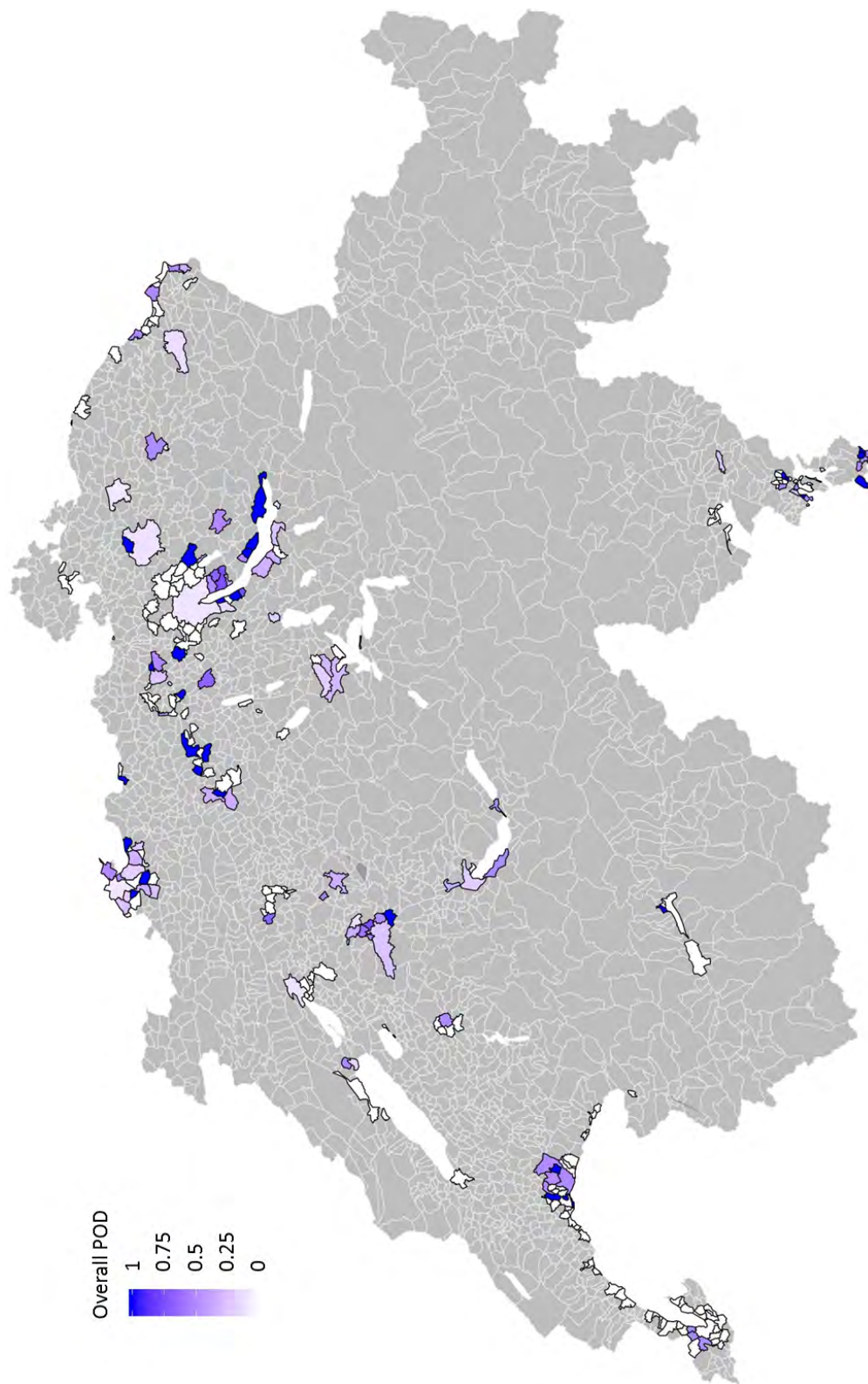
Figure 5.18: POD computed for each Swiss community with high share of living zones over 8 months.
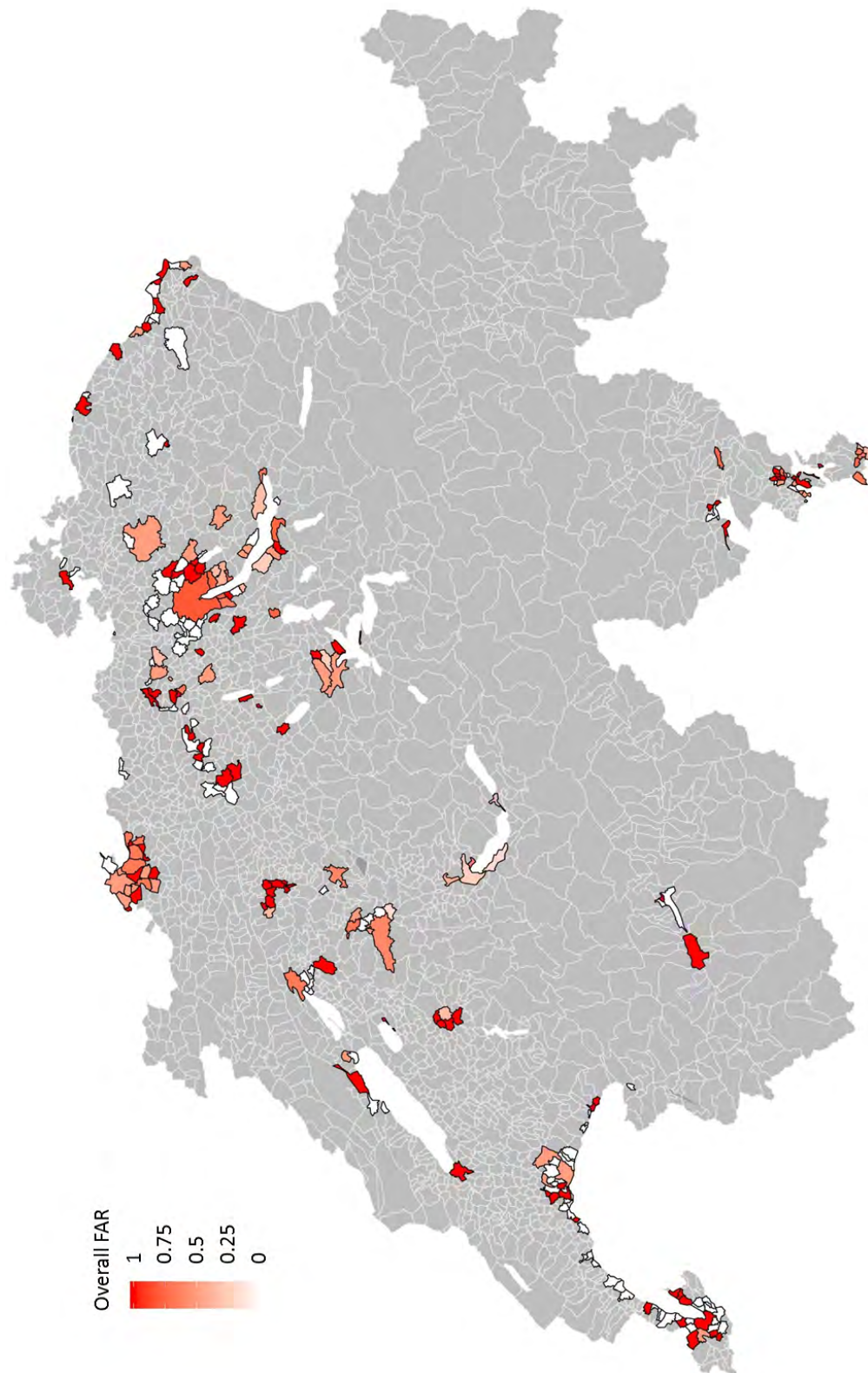
Figure 5.19: FAR computed for each Swiss community with high share of living zones over 8 months.
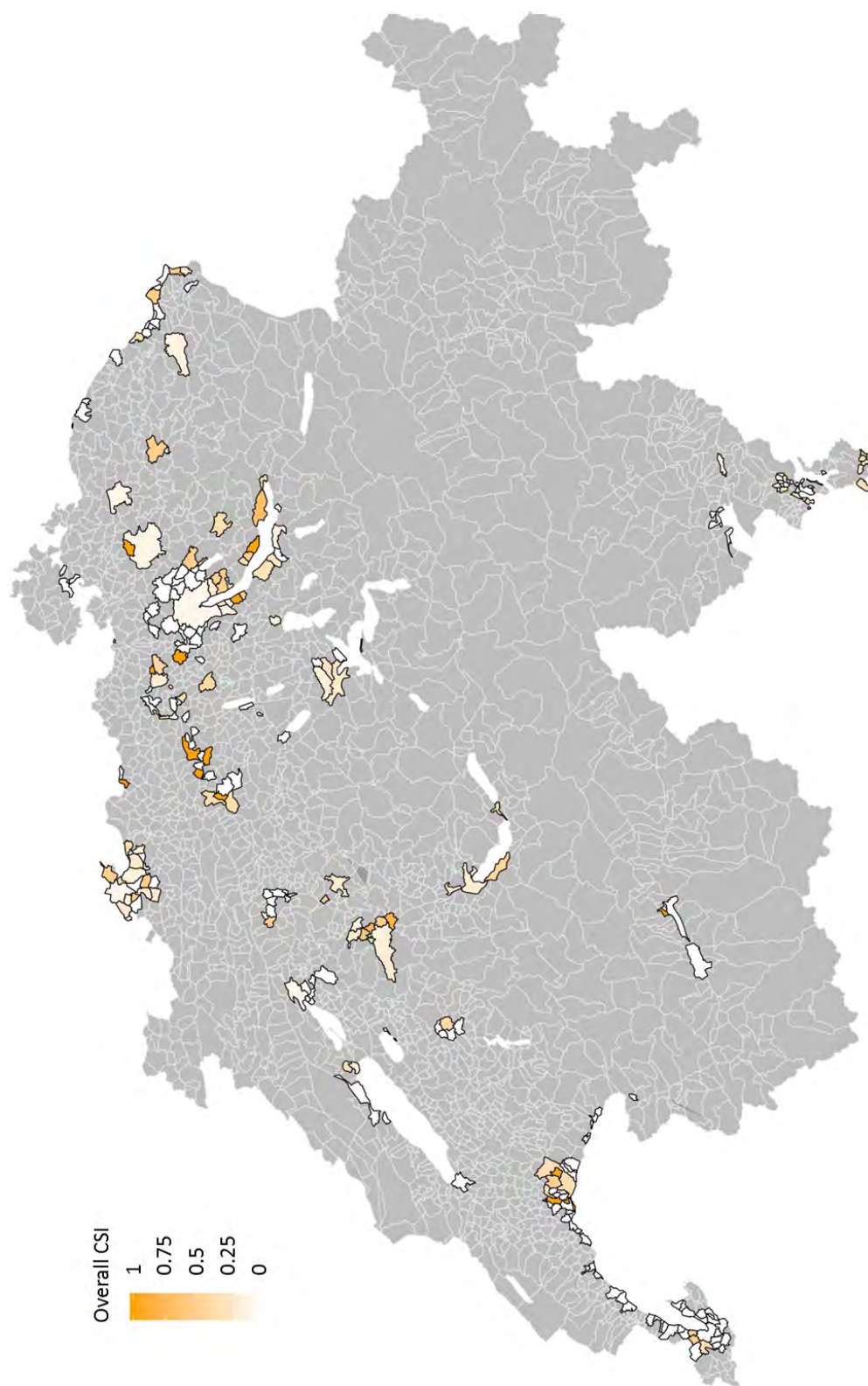
Figure 5.20: CSI computed for each Swiss community with high share of living zones over 8 months.
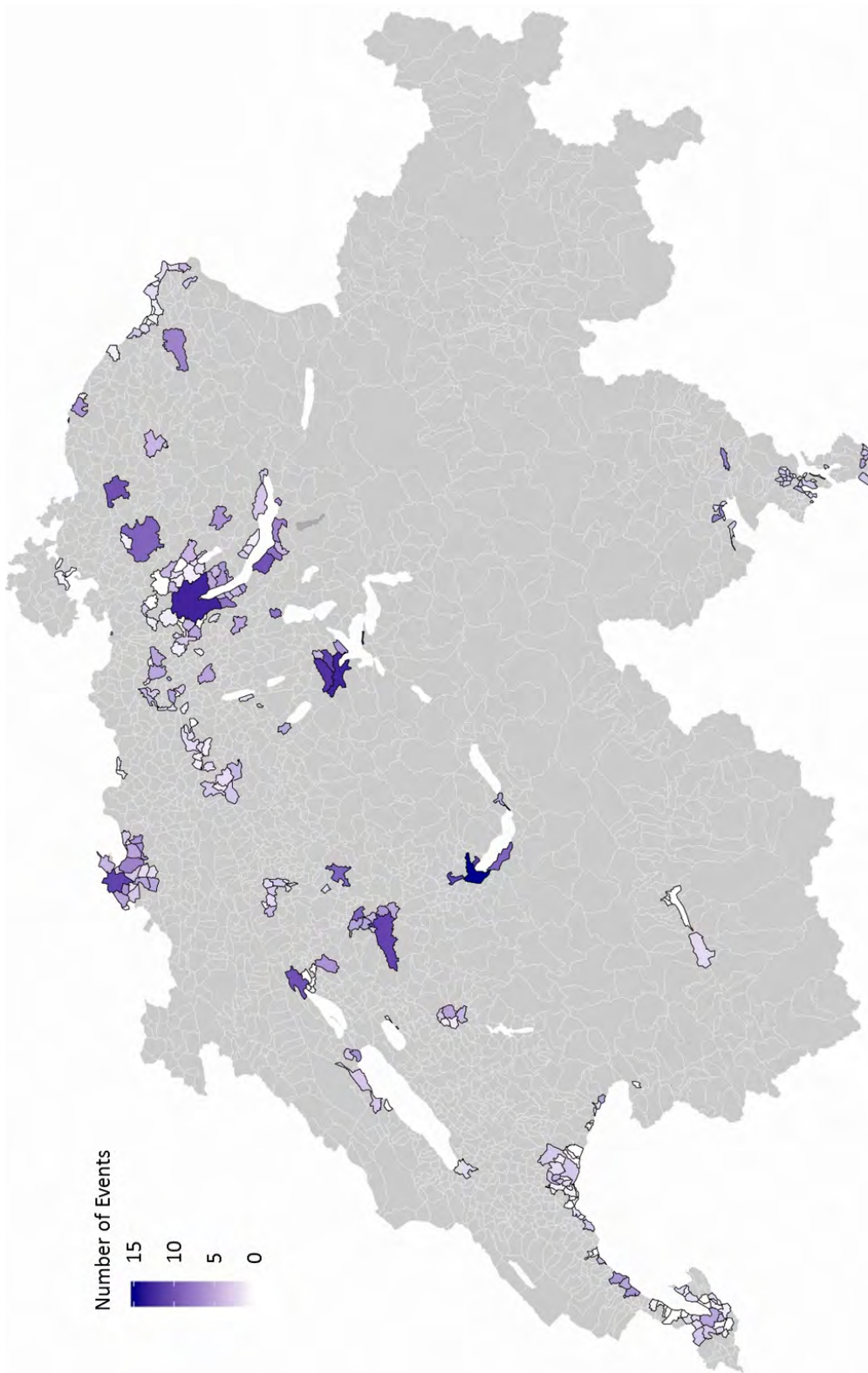
Figure 5.21: Number of hail events (hits+misses+false alarms) computed for each Swiss community with high share of living zones over 8 months.

6   Discussion

6.1   Reliability and Usability of the Crowd-Sourced Data

The crowd-sourced data strongly depend on the population density as well on the individual activities of the users. All kinds of people contributed information about hail events. Individual characters and perceptions of App users play a role in the reporting. Therefore, false reports have to be identified and filtered out, according to certain criteria. Moreover, most users do not know the difference between hail and graupel. Therefore, the crowd-sourced data likely contain reports about graupel.

A large fraction (84.6 % for the MESHS and respectively 72.9 % for HS) of the crowd-sourced data could not be matched to the radar-based hail detection algorithms. Many of them were sent when the sky is clear or the radar detected only low rain intensities. 32.9 % of the crowd-sourced data are No hail reports. The verifications methods focused on reports about hail and not about no hail reports. Therefore, the fractions of matched reports reach 21.5 % (78.5 % mismatched reports) for the MESHS respectively 34.9 % (65.1 % mismatched reports) for the HS, when the no hail reports are ignored. The reasons for such high amounts of mismatches are difficult to analyse. It can be speculated that many users just test the application or make unintentionally mistakes in the reporting procedure. For instance, forgetting to change the time and/or place. It cannot be excluded that a few users sent false reports on purpose. The narrow time window of 45 minutes excludes hail reports which fulfil the spatial tolerance, but the reporting time was outside the time tolerance. The narrow time window was chosen in order to guarantee the quality of the reports. It has not been tested how a wider or narrower time window would affect the analyses.

The fractions of mismatched reports is highest among the larger reported hail diameter classes such as the larger than 5 frank coin and 5 frank coin classes. The best fraction of matched reports is achieved by the coffee bean size class which is the smallest. The fraction of mismatched reports correlates with the reported size. The radar values are many times very low when large hailstone sizes are reported. It cannot be denied that there were really big hailstone, but it is unlikely. Furthermore, the MESHS gives the information about the larger hailstone in a distribution. Smaller hailstones can still fall together with large ones, since the hail growing processes are very complex.

Additionally, the radar resolution of 1 x 1 km^2 is coarser than the spatial structures of thunderstorms. The user reports the hailstones which lie in the instant and viewable surrounding. Many users are not aware of accurate reporting. It must be not the case, but it is expected that the selected size refers to the largest hailstone in the surrounding. Since the amount of reports are low among the larger reported size classes, the statistics for the large hailstone classes gets affected by such overestimations strongly and unconscious, bad sampling issues. *Punge et al.* (2014) analysed hail reports collected by European Severe Weather Database (ESWD) and found out that 23.8 % of all reports comment hailstone diameters are larger than 20 mm, but only 3.5 % larger than 40 mm.

6.2   Correlation of the MESHS respectively HS

Nevertheless, the MeteoSwiss crowd-sourced data correlate to some degree with the radar detections (MESHS and HS). The MESHS agrees better the larger hailstone size classes (see figure 5.10). The HS medians of the 5 frank coin size class boxplot are nearly identical to real size of the 5 frank coin. The HS and the larger hailstone classes do not agree well. A plausible reason is the spatially and temporally wider hail detection of the HS than the detection of the MESHS. This leads to more matched false reports and overestimations. The upper quartile has a HS value of approximately 40 mm. The 1 frank coin and coffee bean size classes show too high MESHS medians. The too high MESHS medians are due to the fact that the MESHS only detects hailstone diameters larger than 20 mm. The concentration of the reported coffee bean and 1 frank coin sizes close to the 20 mm speaks as well for the restrictiveness of MESHS representing the smaller size classes. Therefore, the extension of MESHS is crucial for representing small hailstones.

The HS suits the coffee bean size class since the medians is similar to real coffee bean size range (see figure 5.12). The 1 frank coin size class is represented only by the best match approach well. The MESHS median of the neighbourhoods by the nearest match approach overestimates the 1 frank coin class. The HS median of the neighbourhoods by the nearest match approach underestimates the 1 frank coin class. The HS medians of the larger hailstone classes are clearly smaller than the MESHS medians of the same classes. Here the question rise that the HS has a bimodal or a trimodal regime. The additional low value pixels next to higher value pixels pull down the HS median. Moreover, the HS contains larger areas,

which leads to higher amount over false reports and overestimations. The HS-POH correlation in the radar data shows a trimodal behaviour / regime (see figure 4.3). The HS-POH relation is more or less reflected in the matched reports of the MeteoSwiss crowd-sourced data (see figure 4.3). The MESHS-POH relation shows a limited linearity with a high heteroscedasticity. A second-degree or third-degree polynomial would probably describe better the MESHS(HS)-POH relation.

The hail sensor measurements do not show a linear relation to HS medians of the neighbourhood when the data is computed by the nearest match approach (see figure 5.15). The computation with the best match match approach leads to a much better linearity. The measurement of large hailstones is much influenced by rarity of larger hailstones as mentioned above. *Betschart and Hering* (2012) used maximum observed crowd-sourced report in a regression against the MESHS (see figure 6.1). Their coefficient of corrlation reached 0.45 for the median of MESHS and 0.4781 for the maximum MESHS. The correlation of max MESHS values and maximum observed hail diameter show similiar behaviour for small hailstones as the correlation of hail sensor measurements and HS medians compiled in this study (compare with figure 5.15). There are high medians of radar-based hail detections algorithm when low hailstone were observed/measured on ground. Neither the nearest match approach nor the best match approach may connect the reported hail or measured hail to the truly related radar detection time. The comparison of the measurements of the hail sensor station in Lucerne Sedel with the HS median and HS pixel values of the neighbourhood show that there are suitable values in the neighbourhood. But the approaches do only consider the median of the neighbourhood and ignore single pixel values. Similar studies such as *Schuster et al.* (2005) and *Betschart and Hering* (2012) used all pixel values inside the applied grid and did not use just suitable, single pixels.

6.3 Reflection of the Neighbourhood Verification

Just assigning the measured diameters with the next (temporally and spatially) best values does not lead to the real connection of detection and measurement. Hailstone can get shifted considerable horizontal distances or they can fall nearly vertically. The displacements of hails while they are falling is not well measurable. It takes some time while hailstones falling from the level of radar-based detection to the ground. *Matson and Huggins* (1980), *Mezeix and Admirat* (1978) and *Knight and Heymsfield* (1983) measured a terminal fall speed of 20 ms^{-1} for
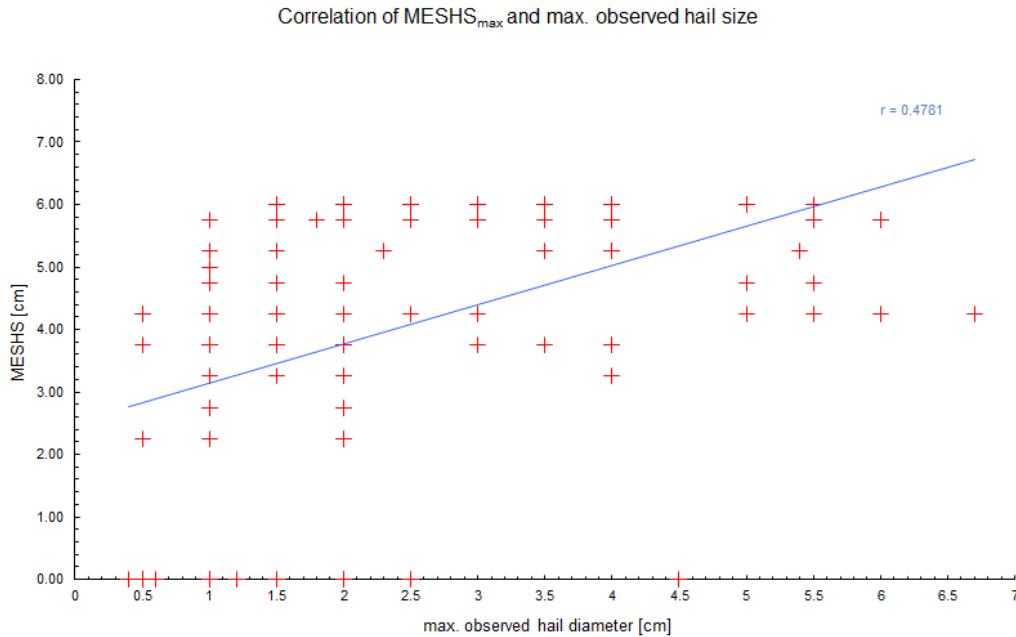
Figure 6.1: Correlation of max MESHS values and maximum observed hail diameter for the year 2011 (red crosses). The x-axis represents the maximum observed hail diameters. The y-axis shows the corresponding MESHS values. *Betschart and Hering* (2012)

hailstones with a diameter of 2 cm. If the freezing level lies at 3'900 m a.s.l. and the nomogram of *Treloar* (1998) is applied, the ET50 is then at a height of about 7'500 a.s.l. Falling directly and vertically from 7'500 m to ground takes 6 minutes and 15 seconds. Smaller hailstones need longer times and larger hailstones shorter times according to their fall speed. As a consequence, there must be a negative time lag between the impact time on ground and radar detection. The hail sensor registers the real impact time. Whereas the App user do rarely report at the impact time. An additional negative time lag has to be expected.

The matches of the MeteoSwiss crowd-sourced data with negative time lag (time lag = radar time - reporting time) outnumber the ones with positive time lag if the best match approach is computed (see figure5.11 and figure 5.13). Therefore, the best match approach is closer to the reality if the hailstones falls nearly vertically. The frequency decreases from the -5 minute time lag to the -10 minute time lag. The frequency increases from the -10 minute time lag to the -20 minute time lag. The reasons for the changing time lag frequencies might be due to transportation processes of hailstones. Hailstones can undergo several cycles inside the clouds and get transported many kilometres without exiting the cycle (*Markowski and*

Richardson, 2011). The best match approach applied on hail sensor measurements leads as well to an increase of the frequency from the time lag -15 to -20 minute (see figure 5.16). The secondary concentration of matches are around the +5 minute time lag with the best match match approach applied on the hail sensor data. Positive time lags of the hail sensor measurements are not realistic. Assessing matches according to the median and IDW of the neighbourhood is not the best approach. It has to be mentioned that the concept of taking the median and IDW as well as both match approaches are designed for the crowd-sourced data on a 5 minute time scale, and not for the hail sensor measurements. The positive time lag is supposed to the temporal tolerance for the customised crowd-sourced reports. The time lag computed by the nearest match approach is forced and may not represent the radar time of the detection.

The comparison of the radar data and the measurements of the hail sensor station in Lucerne Sedel on the 7th of June 2015 demonstrates that matches to -5 minute time lags are available. Furthermore, there are adequate radar pixel values in the neighbourhood available in time periods during, before and after the measured impacts. Picking temporally and spatially the nearest best suitable pixel value within the neighbourhood would lead to 0 minute time lags, which are not realistic in the most cases. Furthermore, it has been noticed that the impacts on the hail sensor station in Lucerne Sedel and its parallel station register different amount of impacts and impacts in different time periods. Measuring hailstone by the hail sensor is as well limited by chance of hitting the sensor area. Therefore, it can happen that the hail sensor does not measure a hail event or does measure only a small sample of hailstones.

Studies such as *Changnon Jr* (1971), *Fraile et al.* (1992), *Giaiotti et al.* (2001), *Fraile et al.* (2003) and *Sánchez et al.* (2009) used hailpads for measuring hailstones. Their results show that the frequency of hailstones decreases exponentially the larger the hailstone are. The location of the hail measurement or hail report cannot be representative for a whole area, which is important to determine the frequency of hail in an area (*Punge and Kunz*, 2016). According to *Smith and Waldvogel* (1989) the estimation of large hailstone diameters is statistically unstable because of the small surface of the hailpads. This is a small advantage of App users that they are not limited by a super small area. Anyway, both datasets are effected by the rareness of larger hail stones. As already mentioned above, *Punge et al.* (2014) found similar results regarding the rarity of hail reports of the ESWD. Melting affects only graupel and small hailstones due to the small ratio of

volume to surface (*Mahoney et al.*, 2012).

Hailstones and especially the larger hailstones are more likely reported by crowd-sourced data than measured by hail sensors. Spatial and temporal representativeness of crowd-sourced data is high in densely populated regions. On the other hand, user can overestimate or underestimate hailstone diameters unconsciously or on purpose. In this sense, the hail sensor are more trustful than the hail report. But the hail sensor may measure hailstone wrongly. The hailstones have to fall in the right angle on the hail sensor. According to *Lozowski and Strong* (1978) wind can have an strong effect on hailpad-based measurements of hail. Additionally, changes in the shape of the hailstone may lead to different impact and momentum energies.

The hail sensor station in Zell registered 16 impacts ranging from 4.9 to 10.5 mm on the 7th of June 2015 (from 18:15 UTC to 21:22 UTC). None of the 16 impacts could be assigned to any matches in the 4 km neighbourhood. There were no matches or pixel even on daily scale. Wind can lift particles from the ground that can lead to impacts on the sensor. But the amount of matches on the hail sensor, Max Echo values above 25 dBZ, the nearby crowd-sourced data and the lightning activity imply impacts of hailstones on the sensor. Hailstone growth can still be possible, but not likely with Max Echo of 25 dBZ and higher. Nevertheless, it seems to be more likely that the hailstone were drifted from the thunderstorms which passed the hail sensor station in the south and the north. The Max Echo pixels nearby the station reaches values up to 50 dBZ. It cannot be excluded that the radar algorithms may not have detected the hail south of the station.

The many crowd-sourced hail reports over Bern on the 7th of June 2015 have to be mismatched reports. The Max Echo values are relatively low for hail growth and the lightning activity is low in the region of Bern. Many of the reports claimed very large hailstone sizes, which are unrealistic with the low Max Echo pixel values. It is high likely that many users wanted to report hailstones from the severe hailstorm on the 6th of June 2015. They may got reminded by the news, the rainy day or some other reasons. Consequently, the spatial and temporal concentration of reports cannot be used as a quality measure for future filtering of the crowd-sourced data.

The crowd-sourced data can be used as well to assess model output such as WRF, which produced in both schemes hail in the region where the most intensity

was detected by the radar and observed by users. The WRF model can even be used to assess the hail detection algorithms. The WRF shows the same hotspot region as the hail detection algorithms on the 6th of June 2015.

6.4   POH Verification

The overall FAR in the categorical verification is relatively high in the most Swiss communities with high share of settlement areas. Many communities with high FAR have low number of events (hits, false reports and misses together). Surprisingly, the FAR is moderate or high in the major big cities, where the amount of potential reports are high. The CSI and POD are high where the FAR is moderate or low and where at least one hail event occurred. The POD is low in cities with many events. Even though the number of crowd-sourced reports are higher in the major city communities, they cannot confirm all the hail events during the day. Consequently, the number of events affects the POD and FAR negatively when too many events respectively far too few events were detected and/or observed. The hailstone may get drifted out of the community and increase in this way the FAR. Additionally, the settlement area do not cover all the land of the communities. Hilly or mountainous parts of the communities are favourable locations for the formation of thunderstorms and are often not populated. The POD can get decreased if the hailstone drift into the community area.

The overall-POD score over all communities is very low with 0.27. The overall-FAR has a score of 0.38. The overall-CSI reach a score of 0.19. The low POD score might be due to the many false report among the crowd-sourced data (due to a too high sensitivity). 27.1 % of the crowd-sourced data were counted as matches (in the HS anaylsis). The moderate FAR score might be low due to the fact that many reports were sent, including false reports. The hail reports were already filtered by a threshold of the 45 dBZ for the median of the neighbourhood. *Nisi et al.* (2016) used insurance loss claims on cars collected over 10 years for the Verification of POH and MESHS. That verification leads to better results when 80 % POH-threshold is considered. The overall-POD achieves a score of 0.89. The overall-FAR is 0.54 and the CSI reach a score of 0.45 %. On the one hand, the overall-POD score of this study is so low because the analysis are restricted to just 8 months in comparison to the 10 years in the study of *Nisi et al.* (2016). On the other hand, insurance claims are based on a higher reliability than the crowd-sourced reports. The users feel free due to the perception of anonymity in

the reporting process. The assessment of FAR is known to be more problematic than the POD (*Delobbe et al.*, 2003; *Saltikoff et al.*, 2010; *Nisi et al.*, 2016) and sometimes even not possible to compute (*Delobbe and Holleman*, 2006). *Betschart and Hering* (2012) used a different approach to calculate the POD, FAR and CSI for the verification of POH in Switzerland (FAR and POD area). Therefore, it cannot be compared with this study.

Schuster et al. (2006) claimed that hailstone drift can be a strong error source, when strong winds occur. The POH categorical verification gets affected by the drift since the verification is restricted to the borders of community. Moreover, the hail senor measurements and the crowd-sourced data has to deal with the hailstone drift processes. *Nisi et al.* (2016) and *Kunz and Kugel* (2015) discussed about the high FAR in their verification of the radar-based detection algorithms with insurance loss claims of vehicles. Uncertainties are generated by the spatial and temporal allocation of the claims and the available potential of damageable cars. The availability of damage claims depends on the population density. Similar error sources in the allocation occur in the crowd-sourced data. Furthermore, the rarity of large hailstones, the limited spatial representativeness of the hail sensor station and the availability of potential hail reporters restrict the verifications of radar-based hail detection algorithms.

7   Summary and Conclusions

This study intended to analyse the correlation of ground observations such as crowds-sourced App reports and hail sensors and radar-based hails detection algorithms. Two radar-based hail detection algorithms, the Maximum Expected Severe Hail Size (MESHS) and Hail Size (HS), are measures for hailstone diameters. Wind drift can cause displacements up to 2.8 km. Temporal and spatial errors in the reporting process can lead to additional displacements. All kinds of people can contribute in the crowd-sourcing. This affects the reporting positively and negatively. So the crowd-source data can contain false reports and overestimations.

Therefore, neighbourhood methods have been introduced in order to assign ground observation to radar-based hail detection algorithms. The neighbourhood methods consist of temporal and spatial windows which only consider radar values within 45 minute period around the reporting time and within 2 km (4km) radius. The temporal and spatial restriction is a quality measure to get accurate reports. The median and the Inverse Distance Weighing (IDW) were used to get a single value representing pixels within the neighbourhoods. Two approaches were developed to determine a single match for each report out of the spatial and temporal windows: The nearest match approach takes the temporally nearest match with priority to the 2 km neighbourhoods. The best match approach chooses the closest radar value to the reported hailstone diameter with priority to the 2 km neighbourhoods.

Usability and Reliability of Crowd-Sourced Data

All kinds of users can report hail in the crowd-sourcing App of MeteoSwiss and Mobiliar. Individuals estimate hailstone diameters differently. Some users may over- or underestimate the reported hailstones. Furthermore, the reporting procedure is unknown and can vary widely among the users. Mistakes and testing purposes can happen while reporting hail with the interface of the crowd-sourcing App. The time and location of the cell phone can be wrongly allocated. Consequently, the crowd-sourced data is negatively affected by false reports. Narrow time and spatial windows as filtering means can reduce the amount of false reports in the proceeded dataset. Datasets with fully reliable and accurate reports cannot be created out of crowd-sourcing. The high spatial and temporal coverage of the crowd-sourced hail reports is an advantage over conventional hail reports. Many crowd-sourced reports may outbalance the false reports. But the results have

shown that a contraction of many reports does not mean that they are not false reports. This study could match 21.5 % for the MESHS respectively 34.9 % for the HS algorithms if the no hail reports were removed. The developed approaches in this study did not make use of the no hail reports.

Correlation of MESHS respevtively HS

The size classes of the crowd-sourced data do correlate with the MESHS to a certain degree. The median (including false reports and overestimations). The upper quartile of the reported size classes correlates well with the MESHS. The median leads to a correlation for all hailstone size classes apart from the largest. The larger reported hailstone classes suffer from the limited amount of reports. Other studies observed that the frequency of hailstone diameters decrease the larger hailstone diameters get (measured by hailpads). False reports and overestimation affect the statistics of the larger hailstone size classes strongly. The MESHS can only detect hail equal or larger than 20 mm.

A prototype of a radar-based hail detection algorithm (Hail Size, HS) has been introduced in order to detect the whole range of hailstones. Where MESHS is 0 mm, converted POH values has been used to fill the gap from 0 to 20 mm. HS represents smaller hailstone sizes better than MESHS. A bit problematic is transition of 20 mm. MESHS represents the larger hailstones better than HS. Nevertheless, there is an increase from the median of the smaller hailstone size classes to the lager ones. The median of the largest hailstone class does not follow the increasing series. Low HS pixels values around high pixel values lower the median as well. The larger hailstone size classes suffer from the observation and rarity of larger hailstones. False reports and overestimations corrupt the statistics. The figure 7.1 shows a combination of the MESHS and HS verification, which is an appropriate way to represent better the hailstone size classes than the single verifications do.

POH Verification

A categorical verification of the presence of POH has been conducted on daily basis. Therefore, Swiss communities with at least 30 % settlement area has been selected in order to guarantee a high potential for hail reporting. The maximum of POH values from 07:00 am UTC (09:00 local time) to 19:00 pm UTC (21:00 local time) has been assimilated. The ground observations were not restricted to the daytime. Only radar values and ground observations within the borders of the selected communities were considered. A filter with a threshold of 45 dBZ has been applied on the crowd-sourced data for removing false reports. The categorical ver-
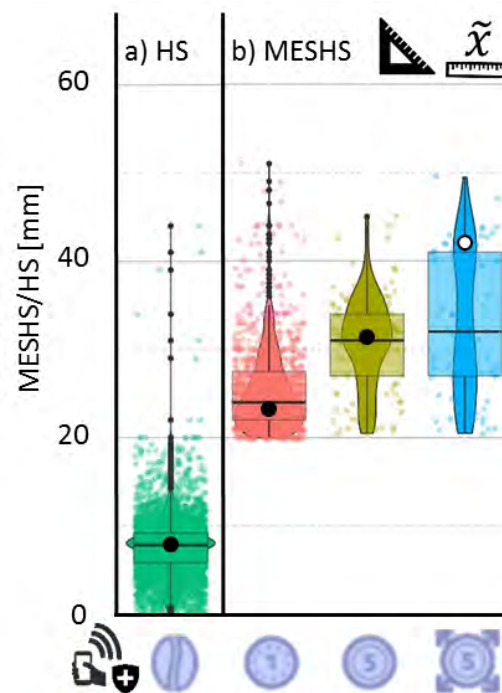
Figure 7.1: Boxplot of the HS respectively MESHS median by the reported sizes of the MeteoSwiss crowd-sourced data computed by the best match approach (only the matches are shown). In part a), the HS median of the neighbourhood for the coffee bean size class was computed. In part b), the MESHS median of the neighbourhood was calculated for the 1 Swiss frank coin, 5 Swiss frank coin and larger than 5 Swiss frank size classes. Violin plots with kernel probability density and scatter plots are added to show smother distributions. The black dots represent the reportable sizes. The white dot represents the set value for the best match approach.

ification of POH achieved low score for Probability Of Detection (POD) and and a moderate scores for Critical Succession Index (CSI) and the False Alarm Rate (FAR). The verification suffers from the short, analysed time period of 8 month data. Communities with high and low number of events lead to low POD scores respectively high FAR scores. The applied filter did not remove all false reports. The crowd-sourced data contains many false reports as result of testing, temporal and spatial allocation (too late reporting and moving users) and cell phone reception. Additionally, hailstone drift can cause large displacements across the border of communities. Moreover, the amount of potential users may not guaranteed throughout the day and area of the community. Apart from the overlapping issue, the mentioned error sources affect other analyses of this study.

High-Resolution Verifications with Crowd-Sourced Data

The case studies of the 6th and 7th of June 2016 showed that the crowd-sourced data capture hailstorm in high spatial and temporal resolution. The crowd-sourced data does agree with the MESHS, POH and HS well in highly populated areas. Regions with high HS/MESHS values were detect by the crowd-sourced data. The hail reports reflect well the life cycle of the hailstorm and its structure on the 5 minute time resolution of the radar data. There are some hail reports in areas which has not been detected neither by HS nor by MESHS. Many reports are counted as false reports since they were sent during dry conditions or in low Max Echo regions. Cases on the case study of the 7th of June 2015 show that missed hail detections can be possible, but far distance drifts could not be excluded. Overall, the HS overcomes the restrictions of MESHS and definite advance in hailstone occurrence and diameter detection.

Only 34.9 % (all no hail reports removed) of the collected crowd-sourced data of MeteoSchweiz could be assigned to any non-zero radar values. Nevertheless, the crowd-source data has much higher spatial and temporal coverage than the hail sensor measurements. It is assumed to be impossible to assign the reported or measured hailstone on ground to the related hail detection. The analyses of the time lags (time lag = radar time - reporting time) imply that none of the applied approaches can fully represent the complex processes of hailstone drift. A negative time lag of few minutes has to be expected due to drift. The analyses of the time lags shows high frequencies around the -20 minute time lag aside the most frequent -5 minute time lag. The large negative time lags might be caused by large distance drift through several hailstone growth cycle. Such case is supported by measurements of the hail sensor station in Zell on the 7th of June 2015.

Investigation of the neighbourhood and measurements of the hail sensor station in Lucerne Sedel on the 7th of June 2015 show that the measured hailstone come from subregion in the neighbourhood 5 minute prior the measurements. The nearest match and best match approach are mainly designed to address the chaotic nature of the allocation of crowd-source data. Positive time lags can be realistic if the users customise the time and/or place after the hail event.

8   Outlook

This study used data over relatively short time period. Longer time series of the crowd-source data and hail sensor measurements are crucial for doing stable statistics. Especially, the evaluation of the larger reported hail size classes and the assessment of measurement of large hailstone by the hail sensor station would benefit the most. The frequency of large hailstone sizes observed or measured on ground decreases exponentially the larger the hailstone are. The HS still needs further adjustments for hailstone diameters lower than 20 mm and for smoothing the transition from small hailstone to larger ones. An extension of the hail sensor network in general would be beneficial for data allocation and the spatial coverage. MeteoSwiss and Mobiliar should agree on common hailstone size classes in the future. Separate scales for the hailstone sizes make a comparison nearly impossible. Additional size classes (smaller and larger ones) would enrich the information about hailstone diameters. Possible hailstone size classes would be pinhead, ping-pong ball, tennis ball, baseball and so on. The largest size class may attract overestimation and false reports and would serve mainly as dump size. Large datasets of crowd-source data over time scales would allow to introduce filter mechanisms based on the reliability of users. Black and white list of users can be created out of analyses. These lists can be used to filter the crowd-sourced data efficiently. This filtered data can then be used for thunderstorm nowcasting purposes, for weather monitoring or for the assessment of insurance damage claims.

Appendix
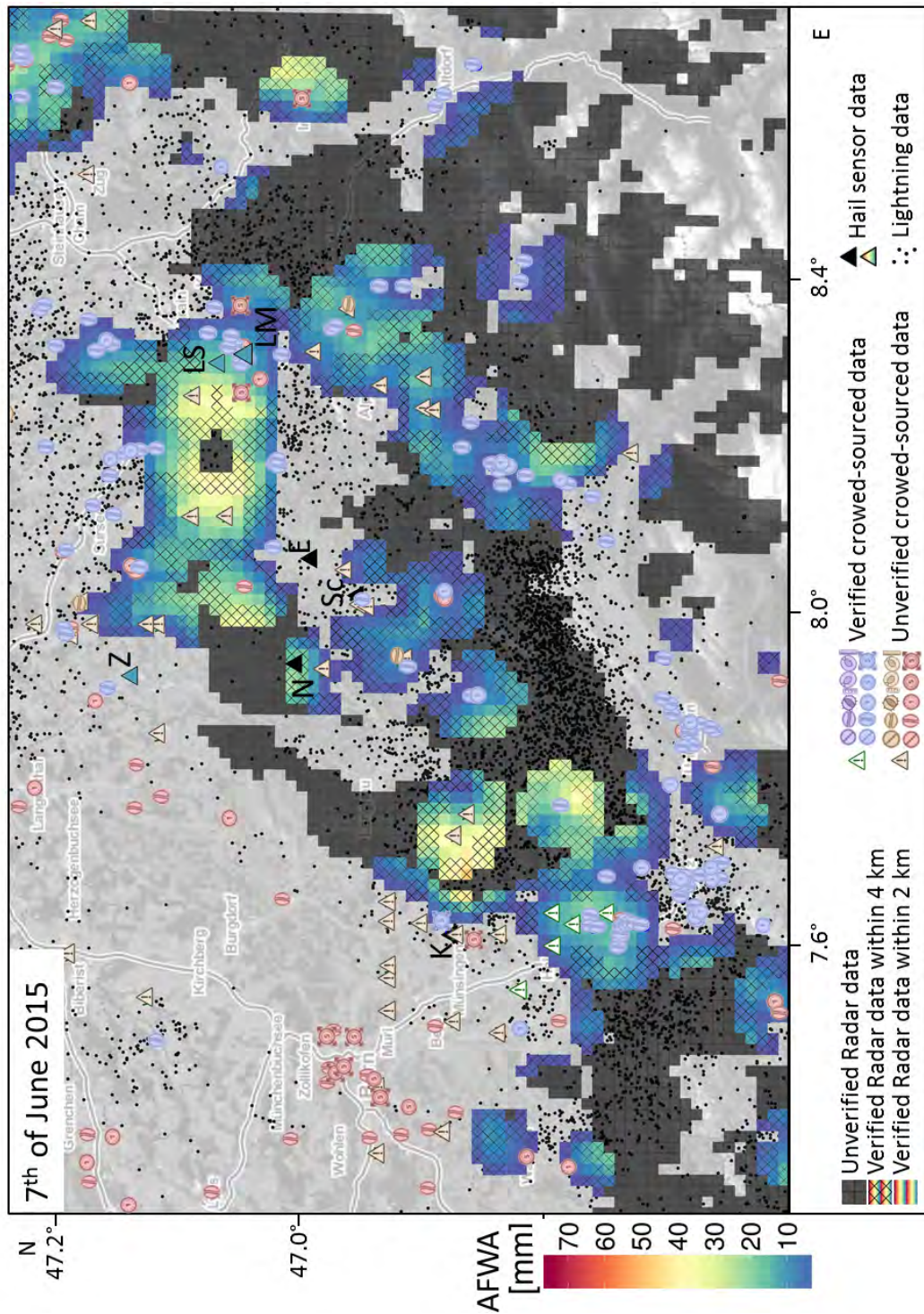
Figure 8.1: Map showing the daily hail from the Morrison double-moment scheme, hail sensor measurements and the crowed-sourced data over the Napf-region for the 7th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim, E Entlebuch, Z Zell, LM Lucerne Moosstrasse and LS Lucerne Sedel.
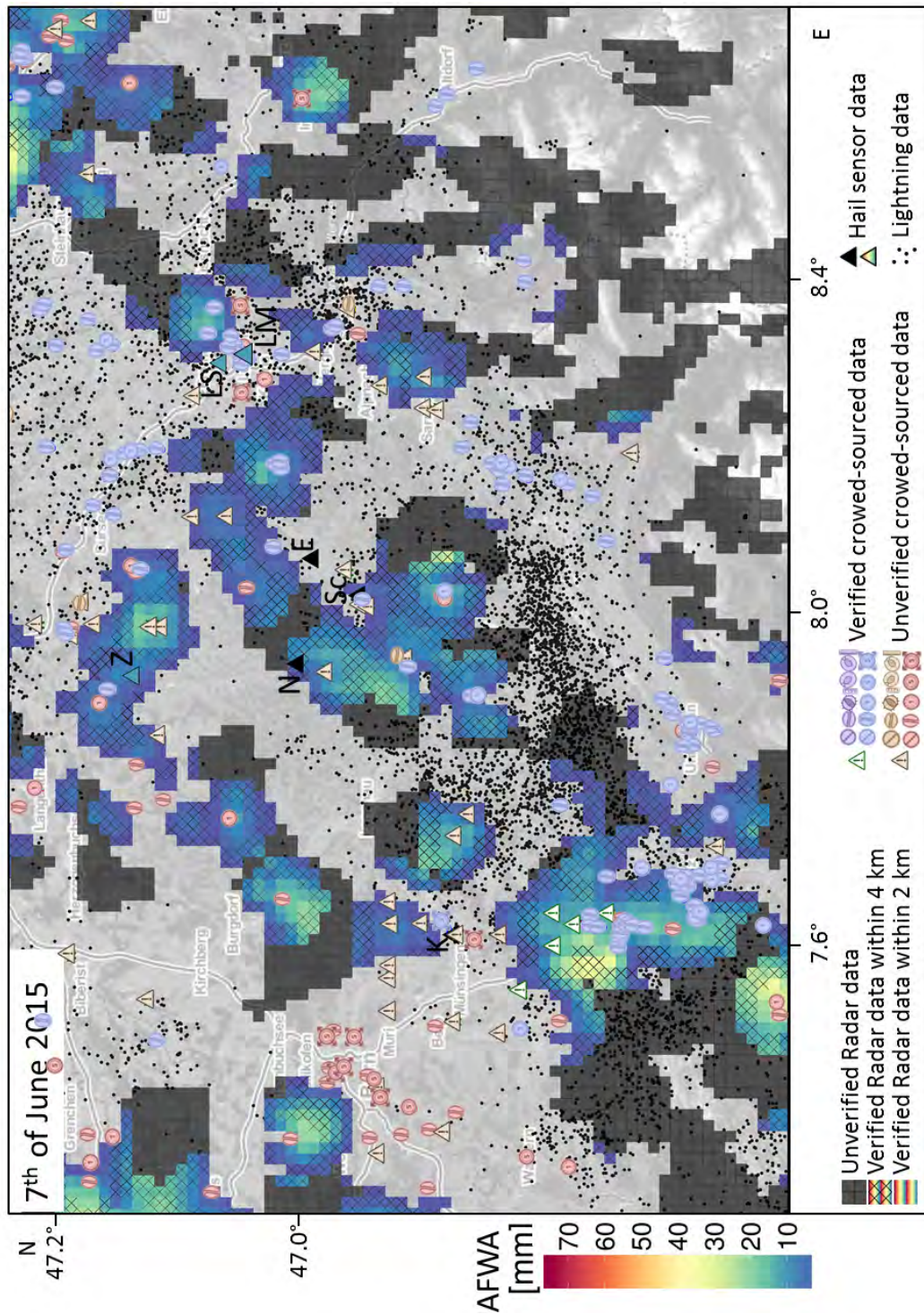
Figure 8.2: Map showing the daily hail from the Thompson scheme, hail sensor measurements and the crowed-sourced data over the Napf-region for the 7th of June 2015. The hail sensor stations are represented by: K Konolfingen, N Napf, Sc Schüpfheim, E Entlebuch, Z Zell, LM Lucerne Moosstrasse and LS Lucerne Sedel.

References

Angsheng, W. (1993), Characteristics of large natural hailstones and studies of artificial hailstones, *Frontiers in atmospheric sciences, Allerton Press*, p. 375.

Atger, F. (2001), Verification of intense precipitation forecasts from single models and ensemble prediction systems, *Nonlinear Processes in Geophysics*, *8*(6), 401–417.

Austin, P. M. (1987), Relation between measured radar reflectivity and surface rainfall, *Monthly Weather Review*, *115*(5), 1053–1070.

Awan, N. K., H. Truhetz, and A. Gobiet (2011), Parameterization-induced error characteristics of mm5 and wrf operated in climate mode over the alpine region: an ensemble-based analysis, *Journal of Climate*, *24*(12), 3107–3123.

Aydin, K., T. Seliga, and V. Balaji (1986), Remote sensing of hail with a dual linear polarization radar, *Journal of Climate and Applied Meteorology*, *25*(10), 1475–1484.

Basara, J. B., D. R. Cheresnick, D. Mitchell, and B. G. Illston (2007), An analysis of severe hail swaths in the southern plains of the united states, *Transactions in GIS*, *11*(4), 531–554.

Battan, L. J. (1973), Radar observation of the atmosphere, *The University of Chicago Press*, *99*, 324.

Berthet, C., E. Wesolek, J. Dessens, and J. Sánchez (2013), Extreme hail day climatology in southwestern france, *Atmospheric Research*, *123*, 139–150.

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 nr. 238, http://www.meteoschweiz.admin.ch/content/dam/meteoswiss/en/Ungebundene-Seiten/Publikationen/Fachberichte/doc/ab238.pdf.

Bonelli, P., and P. Marcacci (2008), Thunderstorm nowcasting by means of lightning and radar data: algorithms and applications in northern italy, *Natural Hazards and Earth System Science*, *8*(5), 1187–1198.

Boudreau, K. J., and K. R. Lakhani (2013), Using the crowd as an innovation partner, *Harvard business review*, *91*(4), 60–69.

Brimelow, J. C., G. W. Reuter, R. Goodson, and T. W. Krauss (2006), Spatial forecasts of maximum hail size using prognostic model soundings and hailcast, *Weather and forecasting*, *21*(2), 206–219.

Brooks, H. E. (2009), Proximity soundings for severe convection for europe and the united states from reanalysis data, *Atmospheric Research*, *93*(1), 546–553.

Brooks, H. E., J. W. Lee, and J. P. Craven (2003), The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data, *Atmospheric Research*, *67*, 73–94.

Browning, K. (1977), *A Review of Hail Science and Hail Suppression*, vol. 16, chap. The structure and mechanisms of hailstorms, pp. 1–43, Meteor. Monogr, American Meteorological Society.

Browning, K., and G. Foote (1976), *Airflow and hail growth in supercell storms and some implications for hail suppression*, vol. 102, 499–533 pp., Wiley Online Library.

Casati, B., L. Wilson, D. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. Ebert, B. Brown, and S. Mason (2008), Forecast verification: current status and future directions, *Meteorological applications*, *15*(1), 3–18.

Changnon, S. A. (1999), Data and approaches for determining hail risk in the contiguous united states, *Journal of Applied Meteorology*, *38*(12), 1730–1739.

Changnon Jr, S. A. (1971), Note on hailstone size distributions, *Journal of Applied Meteorology*, *10*(1), 168–170.

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, *Weather and Forecasting*, *27*(5), 1235–1248.

Cossu, F., and K. Hocke (2014), Influence of microphysical schemes on atmospheric water in the weather research and forecasting model, *Geoscientific Model Development*, *7*(1), 147–160.

Creighton, G., E. Kuchera, R. Adams-Selin, J. McCormick, S. Rentschler, and B. Wickard (2014), Afwa diagnostics in wrf.

Cullen, M., and A. Brown (2009), Large eddy simulation of the atmosphere on various scales, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *367*(1899), 2947–2956.

Damrath, U. (2004), Verification against precipitation observations of a high density network—what did we learn?, in *Int. Verification Methods Workshop; Montreal, Canada.*

Davies, J. M. (2004), Estimations of cin and lfc associated with tornadic and nontornadic supercells, *Weather and forecasting*, *19*(4), 714–726.

Delobbe, L., and I. Holleman (2006), Uncertainties in radar echo top heights used for hail detection, *Meteorological Applications*, *13*(4), 361–374.

Delobbe, L., D. Dehenauw, K. Hamid, and J. Neméghaire (2003), Hail detection using radar observations: case studies in the summer 2002, *Scientific report*, (29).

Delobbe, L., I. Holleman, D. Dehenauw, and J. Neméghaire (2005), Verification of radar-based hail detection product, in *Preprints WWRP Symposium on Nowcasting and Very Short Range Forecasting (WSN05).*

Diendorfer, G. (2002), Euclid—technical structure and performance of the european wide lightning location system, in *Proc. Int. Conf. Grounding Earthing 3rd Brazilian Workshop Atmos. Electr*, pp. 4–7.

Dixon, M., and G. Wiener (1993), Titan: Thunderstorm identification, tracking, analysis, and nowcasting-a radar-based methodology, *Journal of Atmospheric and Oceanic Technology*, *10*(6), 785–797.

Doms, G., U. Schättler, and M. Baldauf (2015), Description of the nonhydrostatic regional cosmo model, part i: Dynamics and numerics, *Tech. rep.*, www.cosmo-model.org/content/model/documentation/core/cosmoDyncsNumcs.pdf.

Donaldson, J. R. (1961), Radar reflectivity profiles in thunderstorms, *Journal of Meteorology*, *18*(3), 292–305.

Doswell, C. A., and J. S. Evans (2003), Proximity sounding analysis for derechos and supercells: An assessment of similarities and differences, *Atmospheric Research*, *67*, 117–133.

Droegemeier, K. K., S. M. Lazarus, and R. Davies-Jones (1993), The influence of helicity on numerically simulated convective storms, *Monthly weather review*, *121*(7), 2005–2029.

Ebert, E. E. (2008), Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework, *Meteorological applications*, *15*(1), 51–64.

Ebert, E. E. (2009), Neighborhood verification: A strategy for rewarding close forecasts, *Weather and Forecasting*, *24*(6), 1498–1510.

Elmore, K. L., Z. Flamig, V. Lakshmanan, B. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz (2014), mping: Crowd-sourcing weather reports for research, *Bulletin of the American Meteorological Society*, *95*(9), 1335–1342.

Foote, G. B., and C. A. Knight (1979), Results of a randomized hail suppression experiment in northeast colorado. part i: Design and conduct of the experiment, *Journal of Applied Meteorology*, *18*(12), 1526–1537.

Foote, G. B., T. W. Krauss, and V. Makitov (2005), Hail metrics using conventional radar, in *Proc., 16th Conference on Planned and Inadvertent Weather Modification.*

Foresti, L., M. Kanevski, and A. Pozdnoukhov (2011), Data-driven exploration of orographic enhancement of precipitation, *Advances in Science and Research*, *6*(1), 129–135.

Fraile, R., A. Castro, and J. Sánchez (1992), Analysis of hailstone size distributions from a hailpad network, *Atmospheric research*, *28*(3-4), 311–326.

Fraile, R., C. Berthet, J. Dessens, and J. L. Sánchez (2003), Return periods of severe hailfalls computed from hailpad data, *Atmospheric Research*, *67*, 189–202.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger (2006), Radar precipitation measurement in a mountainous region, *Quarterly Journal of the Royal Meteorological Society*, *132*(618), 1669–1692.

Germann, U., M. Boscacci, M. Gabella, and M. Satori (2015), Radar design for prediction in the swiss alps, *Meteorological Technology International*, *4*, 42–45.

Germann, U., M. Boscacci, M. Gabella, and M. Schneebeli (2016), *From weather observations to atmospheric and climate sciences in Switzerland: Celebrating 100 years of the Swiss Society for Meteorology*, chap. Weather radar in Switzerland, vdf Hochschulverlag AG.

Giaiotti, D., E. Gianesini, and F. Stel (2001), Heuristic considerations pertaining to hailstone size distributions in the plain of friuli-venezia giulia, *Atmospheric research*, *57*(4), 269–288.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert (2009), Intercomparison of spatial forecast verification methods, *Weather and Forecasting*, *24*(5), 1416–1430.

Grasso, V., I. Zaza, F. Zabini, G. Pantaleo, P. Nesi, and A. Crisci (2016), Weather events identification in social media streams: tools to detect their evidence in twitter, *PeerJ Preprints*, *4*, e2241v1.

Groenemeijer, P., and A. Van Delden (2007), Sounding-derived parameters associated with large hail and tornadoes in the netherlands, *Atmospheric research*, *83*(2), 473–487.

Grzych, M. L., B. D. Lee, and C. A. Finley (2007), Thermodynamic analysis of supercell rear-flank downdrafts from project answers, *Monthly weather review*, *135*(1), 240–246.

Halder, M., A. Hazra, P. Mukhopadhyay, and D. Siingh (2015), Effect of the better representation of the cloud ice-nucleation in wrf microphysics schemes: A case study of a severe storm in india, *Atmospheric Research*, *154*, 155–174.

Hanson, G. S. (2016), Impact of assimilating surface pressure observations from smartphones on a regional, high resolution ensemble forecast: Observing system simulation experiments, Ph.D. thesis, The Pennsylvania State University.

Havlik, D., M. Egly, H. Huber, P. Kutschera, M. Falgenhauer, and M. Cizek (2013), Robust and trusted crowd-sourcing and crowd-tasking in the future internet, in *International Symposium on Environmental Software Systems*, pp. 164–176, Springer.

Held, G. (1978), The probability of hail in relation to radar echo heights on the south african highveld, *Journal of Applied Meteorology*, *17*(6), 755–762.

Hering, A., C. Morel, G. Galli, S. Sénési, P. Ambrosetti, and M. Boscacci (2004), Nowcasting thunderstorms in the alpine region using a radar based adaptive thresholding scheme, in *Proceedings of ERAD*, vol. 1.

Hering, A., S. Sénési, P. Ambrosetti, and I. Bernard-Bouissières (2005), Nowcasting thunderstorms in complex cases using radar data, in *WMO Symposium on Nowcasting and Very Short Range Forecasting*.

Hering, A., L. Nisi, G. Della Bruna, M. Gaia, D. Nerini, P. Ambrosetti, U. Hamann, S. Trefalt, and U. Germann (2015), Fully automated thunderstorm warnings and

operational nowcasting at meteoswiss, in *European Conerence on Severe Storms 2015*.

Hering, A. M., U. Germann, M. Boscacci, and S. Sénési (2008), Operational nowcasting of thunderstorms in the alps during map d-phase, in *Proceedings of the 5th European Conference on Radar Meteorology (ERAD 2008), June*.

Heymsfield, A. J., A. R. Jameson, and H. W. Frank (1980), Hail growth mechanisms in a colorado storm: Part ii: Hail formation processes, *Journal of the Atmospheric Sciences, 37*(8), 1779–1807.

Hohl, R., H.-H. Schiesser, and D. Aller (2002a), Hailfall: the relationship between radar-derived hail kinetic energy and hail damage to buildings, *Atmospheric Research, 63*(3), 177–207.

Hohl, R., H.-H. Schiesser, and I. Knepper (2002b), The use of weather radars to estimate hail damage to automobiles: an exploratory study in switzerland, *Atmospheric research, 61*(3), 215–238.

Holleman, I. (2001), *Hail detection using single-polarization radar*, Scientific report WR-2001-01, Tech. rep., Royal Netherlands Meteorological Institute (KNMI) (2001) (74 pp.).

Holleman, I., H. Beekhuis, S. Noteboom, L. Evers, H. Haak, H. Falcke, and L. Bähren (2006), Validation of an operational lightning detection system, in *19th International Lightning Detection Conference, Tucson, USA, 12pp*.

Houze, R., W. Schmid, R. Fovell, and H. Schiesser (1993), Hailstorms in switzerland: Left movers, right movers, and false hooks, *Monthly weather review, 121*(12), 3345–3370.

Howe, J. (2006), The rise of crowdsourcing, *Wired magazine, 14*(6), 1–4.

Hoy, J., and D. Klein (2016), Mobile crowdsourcing for high-accuracy thunderhead prediction, *Technical Disclosure Commons, (June 03, 2016) http://www.tdcommons.org/dpubs_series/207*.

Huntrieser, H., H. Schiesser, W. Schmid, and A. Waldvogel (1997), Comparison of traditional and newly developed thunderstorm indices for switzerland, *Weather and Forecasting, 12*(1), 108–125.

Hyvärinen, O., and E. Saltikoff (2010), Social media as a source of meteorological observations, *Monthly Weather Review, 138*(8), 3175–3184.

Imran, M., S. Elbassuoni, C. Castillo, F. Diaz, and P. Meier (2013), Extracting information nuggets from disaster-related messages in social media, in *Proceedings of the 10th International ISCRAM Conference – Baden-Baden, Germany*.

inNET (2016), Hagelsensor mit automatischer datenerfassung. http://www.innetag.ch/hagelsensor.109.0.html., *Tech. rep.*

Jankov, I., W. A. Gallus Jr, M. Segal, B. Shaw, and S. E. Koch (2005), The impact of different wrf model physical parameterizations and their interactions on warm season mcs rainfall, *Weather and forecasting, 20*(6), 1048–1060.

Jankov, I., J.-W. Bao, P. J. Neiman, P. J. Schultz, H. Yuan, and A. B. White (2009), Evaluation and comparison of microphysical algorithms in arw-wrf model simulations of atmospheric river events affecting the california coast, *Journal of Hydrometeorology, 10*(4), 847–870.

Joe, P., D. Burgess, R. Potts, T. Keenan, G. Stumpf, and A. Treloar (2004), The s2k severe weather detection algorithms and their performance, *Weather and Forecasting, 19*(1), 43–63.

Johnson, J., P. L. MacKeen, A. Witt, E. D. W. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas (1998), The storm cell identification and tracking algorithm: An enhanced wsr-88d algorithm, *Weather and Forecasting, 13*(2), 263–276.

Jolliffe, I. T., and D. B. Stephenson (2012), *Forecast verification: a practitioner's guide in atmospheric science*, John Wiley & Sons.

Joss, J., B. Schaedler, G. Galli, R. Gavalli, M. Boscacci, E. Held, G. D. Bruna, G. Kappenberger, V. Nespor, and R. Spiess (1998), *Operational Use of Radar for Precipitation Measurements in Switzerland*, Hochschulverlag AG ETH Zurich, Switzerland.

Kaltenboeck, R., and M. Steinheimer (2015), Radar-based severe storm climatology for austrian complex orography related to vertical wind shear and atmospheric instability, *Atmospheric Research, 158*, 216–230.

Kessinger, C., M. Hjemfelt, and J. Wilson (1983), Low-level microburst wind structure using doppler radar and pam data, *Reprints, 21st Conference on Radar Meteorology*, pp. 609–615.

Kessinger, C. J., D. B. Parsons, and J. W. Wilson (1988), Observations of a storm containing misocyclones, downbursts, and horizontal vortex circulations, *Monthly weather review, 116*(10), 1959–1982.

Kessinger, C. J., E. A. Brandes, and J. W. Smith (1995), A comparison of the nexrad and nssl hail detection algorithms, in *In Reprints of the 27th Conference on Radar Meteorology, Vail, Colorado*, pp. 603–605.

Kim, N.-Y., Y.-H. Kim, Y. Yoon, H.-H. Im, R. K. Choi, and Y. H. Lee (2015), Correcting air-pressure data collected by mems sensors in smartphones, *Journal of Sensors*, *2015*.

King, J. R. (2016), Master's thesis: Environmental conditioning of cool season, low instability thunderstorm environments in the tennessee and ohio valleys and southeastern us, *North Carolina State University*.

Knapp, K. R., J. L. Matthews, J. P. Kossin, and C. C. Hennon (2016), Identification of tropical cyclone "storm types" using crowd-sourcing, *Monthly Weather Review*, (2016).

Knight, N. C., and A. J. Heymsfield (1983), Measurement and interpretation of hailstone density and terminal velocity, *Journal of the Atmospheric Sciences*, *40*(6), 1510–1516.

Knupp, K. R. (1987), Downdrafts within high plains cumulonimbi. part i: General kinematic structure, *Journal of the atmospheric sciences*, *44*(6), 987–1008.

Kober, K., and A. Tafferner (2009), Tracking and nowcasting of convective cells using remote sensing data from radar and satellite, *Meteorologische Zeitschrift*, *18*(1), 75–84.

Koole, M., and P. Siegmund (2016), Evaluating the quality and usability of crowd-sourced weather data, in *EGU General Assembly Conference Abstracts*, vol. 18, p. 3959.

Kunz, M., and P. I. Kugel (2015), Detection of hail signatures from single-polarization c-band radar reflectivity, *Atmospheric Research*, *153*, 565–577.

Kunz, M., and M. Puskeiler (2010), High-resolution assessment of the hail hazard over complex terrain from radar and insurance data, *Meteorologische Zeitschrift*, *19*(5), 427–439.

Kunz, M., J. Sander, and C. Kottmeier (2009), Recent trends of thunderstorm and hailstorm frequency and their relation to atmospheric characteristics in southwest germany, *International Journal of Climatology*, *29*(15), 2283–2297.

Leuenberger, D., and A. Rossa (2007), Revisiting the latent heat nudging scheme for the rainfall assimilation of a simulated convective storm, *Meteorology and atmospheric physics*, *98*(3-4), 195–215.

Levi, L., N. E. Castellano, O. B. Nasello, and F. Prodi (1999), Requirements for low density riming and two stage growth on atmospheric particles, *Atmospheric research*, *50*(1), 21–35.

Löffler-Mang, M., D. Schön, and M. Landry (2011), Characteristics of a new automatic hail recorder, *Atmospheric Research*, *100*(4), 439–446.

Li, P., and E. S. Lai (2004), Short-range quantitative precipitation forecasting in hong kong, *Journal of Hydrology*, *288*(1), 189–209.

Liljequist, H., Gösta, and C. Konrad (1994), *Allgemeine Meteorologie*, Friedr. Vieweg & Sohn Verlagsgesellschaft.

List, R. (2014), New hailstone physics. part i: Heat and mass transfer (hmt) and growth, *Journal of the Atmospheric Sciences*, *71*(4), 1508–1520.

Longmore, S., S. Miller, D. Bikos, D. Lindsey, E. Szoke, D. Molenar, D. Hillger, R. Brummer, and J. Knaff (2015), An automated mobile phone photo relay and display concept applicable to operational severe weather monitoring, *Journal of Atmospheric and Oceanic Technology*, *32*(7), 1356–1363.

Lorenz, E. N. (1969), Atmospheric predictability as revealed by naturally occurring analogues, *Journal of the Atmospheric Sciences*, *26*(4), 636–646.

Lozowski, E., and G. Strong (1978), On the calibration of hailpads, *Journal of Applied Meteorology*, *17*(4), 521–528.

Mahoney, K., M. A. Alexander, G. Thompson, J. J. Barsugli, and J. D. Scott (2012), Changes in hail and flood risk in high-resolution simulations over colorado's mountains, *Nature Climate Change*, *2*(2), 125–131.

Mandapaka, P. V., U. Germann, L. Panziera, and A. Hering (2012), Can lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex alpine orography?, *Weather and Forecasting*, *27*(1), 28–49.

Markowitz, A. H. (1976), Raindrop size distribution expressions, *Journal of Applied Meteorology*, *15*(9), 1029–1031.

Markowski, P., and Y. Richardson (2011), *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell.

Marshall, J., and K. Gunn (1952), Measurement of snow parameters by radar, *Journal of Meteorology*, *9*(5), 322–327.

Marshall, J., R. Langille, and W. M. K. Palmer (1947), Measurement of rainfall by radar, *Journal of Meteorology*, *4*(6), 186–192.

Marshall, J., T. East, and K. Gunn (1952), *The microwave properties of precipitation particles*, McGill University. (Sci. Rep. MW-7, Stormy Weather Group), 40 pp.

Marshall, J. S., and W. M. K. Palmer (1948), The distribution of raindrops with size, *Journal of meteorology*, *5*(4), 165–166.

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella (2005), The cosmoleps mesoscale ensemble system: validation of the methodology and verification, *Nonlinear Processes in Geophysics*, *12*(4), 527–536.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle (2002), Does increasing horizontal resolution produce more skillful forecasts?, *Bulletin of the American Meteorological Society*, *83*(3), 407–430.

Matson, R. J., and A. W. Huggins (1980), The direct measurement of the sizes, shapes and kinematics of falling hailstones, *Journal of the Atmospheric Sciences*, *37*(5), 1107–1125.

McMaster, H. (1999), The potential impact of global warming on hail losses to winter cereal crops in new south wales, *Climatic Change*, *43*(2), 455–476.

Mecklenburg, S., J. Joss, and W. Schmid (2000), Improving the nowcasting of precipitation in an alpine region with an enhanced radar echo tracking algorithm, *Journal of Hydrology*, *239*(1), 46–68.

Meischner, P. (2013), *Weather radar: principles and advanced applications*, Springer Science & Business Media.

Mercader, J., B. Codina, A. Sairouni, and J. Cunillera (2010), Results of the meteorological model wrf-arw over catalonia, using different parameterizations of convection and cloud microphysics, *Journal of Weather and Climate of the Western Mediterranean*, *7*, 75–86.

Mezeix, J.-F., and P. Admirat (1978), The measurement of hait at ground level, *Atmosphere-Ocean*, *16*(1), 61–68.

Michael Illingworth, S., C. Louise Muller, R. Graves, and L. Chapman (2014), Uk citizen rainfall network: a pilot study, *Weather*, *69*(8), 203–207.

Mittermaier, M., and N. Roberts (2010), Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score, *Weather and Forecasting*, *25*(1), 343–354.

Mittermaier, M., N. Roberts, and S. A. Thompson (2013), A long-term assessment of precipitation forecast skill using the fractions skill score, *Meteorological Applications*, *20*(2), 176–186.

Mittermaier, M. P. (2014), A strategy for verifying near-convection-resolving model forecasts at observing sites, *Weather and Forecasting*, *29*(2), 185–204.

Mohr, S., and M. Kunz (2013), Recent trends and variabilities of convective parameters relevant for hail events in germany and europe, *Atmospheric Research*, *123*, 211–228.

Morrison, H., G. Thompson, and V. Tatarskii (2009), Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes, *Monthly Weather Review*, *137*(3), 991–1007.

Muller, C., L. Chapman, S. Johnston, C. Kidd, S. Illingworth, G. Foody, A. Overeem, and R. Leigh (2015), Crowdsourcing for climate and atmospheric sciences: Current status and future potential, *International Journal of Climatology*, *35*(11), 3185–3203.

Nelson, S. P. (1983), The influence of storm flow structure on hail growth, *Journal of the Atmospheric Sciences*, *40*(8), 1965–1983.

Nelson, S. P. (1987), The hybrid multicellular-supercellular storm-an efficient hail producer. part ii. general characteristics and implications for hail growth, *Journal of the atmospheric sciences*, *44*(15), 2060–2073.

Nisi, L., P. Ambrosetti, and L. Clementi (2014), Nowcasting severe convection in the alpine region: the coalition approach., *Quarterly Journal of the Royal Meteorological Society*, *140*(1), 1684–1699.

Nisi, L., O. Martius, A. Hering, M. Kunz, and U. Germann (2016), Spatial and temporal distribution of hailstorms in the alpine region: a long-term, high resolution, radar-based analysis, *Quarterly Journal of the Royal Meteorological Society*, *142*(697), 1590–1604.

NWS (2015), U.s. national weather service: Social media, http://www.weather.gov/socialmedia.

Ortega, K. L., T. Smith, and K. Scharfenberg (2006), An analysis of thunderstorm hail fall patterns in the severe hail verification experiment, in *Preprints, 23rd Conference on Severe Local Storms*.

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. C. Kolodziej, and J. J. Gourley (2009), The severe hazards analysis and verification experiment, *Bulletin of the American Meteorological Society*, *90*(10), 1519.

Otkin, J. A., and T. J. Greenwald (2008), Comparison of wrf model-simulated and modis-derived cloud data, *Monthly Weather Review*, *136*(6), 1957–1970.

Pflaum, J. C. (1980), Hail formation via microphysical recycling, *Journal of the Atmospheric Sciences*, *37*(1), 160–173.

Prodi, F., G. Santachiara, and A. Franzini (1986), Properties of ice accreted in two-stage growth, *Quarterly Journal of the Royal Meteorological Society*, *112*(474), 1057–1080.

Pruppacher, H., and J. Klett (1997), Microphysics of clouds and precipitation: With an introduction to cloud chemistry and cloud electricity, *Kluwer Acad., Norwell, Mass*, p. 954 pp.

Pruppacher, H. R., and J. D. Klett (2010), *Microphysics of clouds and precipitation*, Springer, Dordrecht; New York, doi:10.1007/978-0-306-48100-0.

Punge, H., and M. Kunz (2016), Hail observations and hailstorm characteristics in europe: A review, *Atmospheric Research*, *176*, 159–184.

Punge, H., K. Bedka, M. Kunz, and A. Werner (2014), A new physically based stochastic event catalog for hail in europe, *Natural Hazards*, *73*(3), 1625–1645.

Puskeiler, M. (2013), *Radarbasierte analyse der hagelgefährdung in deutschland*, vol. 59, KIT Scientific Publishing.

Que, L.-J., W.-L. Que, and J.-M. Feng (2016), Intercomparison of different physics schemes in the wrf model over the asian summer monsoon region, *Atmospheric and Oceanic Science Letters*, *9*(3), 169–177.

Rajeevan, M., A. Kesarkar, S. Thampi, T. Rao, B. Radhakrishna, and M. Rajasekhar (2010), Sensitivity of wrf cloud microphysics to simulations of a severe

thunderstorm event over southeast india, in *Annales geophysicae: atmospheres, hydrospheres and space sciences*, vol. 28, p. 603.

Raupach, T. H., and A. Berne (2016), Spatial interpolation of experimental raindrop size distribution spectra, *Quarterly Journal of the Royal Meteorological Society*, *142*, 125–137, doi:10.1002/qj.2801.

Rezacova, D., Z. Sokol, and P. Pesice (2007), A radar-based verification of precipitation forecast for local convective storms, *Atmospheric Research*, *83*(2), 211–224.

Roberts, N. M., and H. W. Lean (2008), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Monthly Weather Review*, *136*(1), 78–97.

Rotach, M. W., P. Ambrosetti, C. Appenzeller, M. Arpagaus, L. Fontannaz, F. Fundel, U. Germann, A. Hering, M. A. Liniger, M. Stoll, et al. (2009), Map d-phase: Real-time demonstration of weather forecast quality in the alpine region, *Bull. Amer. Meteorol. Soc.*, *90*(9), 1321–1336.

Salek, M., J.-L. Cheze, J. Handwerker, L. Delobbe, and R. Uijlenhoet (2004), Radar techniques for identifying precipitation type and estimating quantity of precipitation. http://www.smhi.se/.

Saltikoff, E., J.-P. Tuovinen, J. Kotro, T. Kuitunen, and H. Hohti (2010), A climatological comparison of radar and ground observations of hail in finland, *Journal of Applied Meteorology and Climatology*, *49*(1), 101–114.

Sánchez, J., B. Gil-Robles, J. Dessens, E. Martin, L. López, J. Marcos, C. Berthet, J. Fernández, and E. García-Ortega (2009), Characterization of hailstone size spectra in hailpad networks in france, spain, and argentina, *Atmospheric Research*, *93*(1), 641–654.

Sassen, K. (1987), Ice cloud content from radar reflectivity, *Journal of climate and applied meteorology*, *26*(8), 1050–1053.

Schemm, S., L. Nisi, A. Martinov, D. Leuenberger, and O. Martius (2016), On the link between cold fronts and hail in switzerland, *Atmospheric Science Letters*, *17*(5), 315–325.

Schiesser, H. (1990), Hailfall: the relationship between radar measurements and crop damage, *Atmospheric Research*, *25*(6), 559–582.

Schiesser, H., R. Hohl, and W. Schmid (1999), Über die beziehung hagelfall-gebäudeschäden: Fallstudie luzern-hagelsturm vom 21 juli 1998, in *Tech. rep.*, Atmospheric Science ETH Zürich, Switzerland.

Schmid, W., H. Schiesser, and A. Waldvogel (1992), The kinetic energy of hailfalls. part iv: Patterns of hailpad and radar data, *Journal of Applied Meteorology*, *31*(10), 1165–1178.

Schulz, W., D. Poelman, S. Pedeboy, C. Vergeiner, H. Pichler, G. Diendorfer, and S. Pack (2014a), Performance validation of the european lightning location system euclid, in *International Colloquium on Lightning and Power Systems*, *CIGRE Lyon*.

Schulz, W., S. Pedeboy, C. Vergeiner, E. Defer, and W. Rison (2014b), Validation of the euclid lls during hymex sop1, in *Proceedings, International Lightning Detection Conference ILDC*.

Schuster, S. S., R. J. Blong, and M. S. Speer (2005), A hail climatology of the greater sydney area and new south wales, australia, *International Journal of Climatology*, *25*(12), 1633–1650.

Schuster, S. S., R. J. Blong, and K. J. McAneney (2006), Relationship between radar-derived hail kinetic energy and damage to insured buildings for severe hailstorms in eastern australia, *Atmospheric research*, *81*(3), 215–235.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio (2009), Next-day convection-allowing wrf model guidance: A second look at 2-km versus 4-km grid spacing, *Monthly Weather Review*, *137*(10), 3351–3372.

Segawa, T., and Y. Honda (2007), The verification of high-resolution precipitation forecasts of the operational jma mesoscale model, in *JWGV Workshop*, 16, p. 31.

Seinfeld, J. H., and S. N. Pandis (2016), *Atmospheric chemistry and physics: from air pollution to climate change*, John Wiley & Sons.

Sekhon, R., and R. Srivastava (1970), Snow size spectra and radar reflectivity, *Journal of the Atmospheric Sciences*, *27*(2), 299–307.

Shepard, D. (1968), A two-dimensional interpolation function for irregularly-spaced data, in *Proceedings of the 1968 23rd ACM national conference*, pp. 517–524, ACM.

Skripniková, K., and D. Řezáčová (2014), Radar-based hail detection, *Atmospheric Research*, *144*, 175–185.

Smith, D. M., R. Eade, and H. Pohlmann (2013), A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction, *Climate dynamics*, *41*(11-12), 3325–3338.

Smith, P. L. (1984), Equivalent radar reflectivity factors for snow and ice particles, *Journal of Climate and Applied Meteorology*, *23*(8), 1258–1260.

Smith, P. L., and A. Waldvogel (1989), On determinations of maximum hailstone sizes from hallpad observations, *Journal of Applied Meteorology*, *28*(1), 71–76.

Spilhaus, A. F. (1948), Drop size, intensity, and radar echo of rain, *Journal of Meteorology*, *5*(4), 161–164.

Sumner, T. (2015), The future of forecasting: Technology promises faster weather predictions on a smaller scale, *Science News*, *187*(9), 20–23.

Theis, S., A. Hense, and U. Damrath (2005), Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach, *Meteorological Applications*, *12*(3), 257–268.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall (2008), Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part ii: Implementation of a new snow parameterization, *Monthly Weather Review*, *136*(12), 5095–5115.

Torriani-Braga, Y. (2009), Ein verheerender hagelschlag, *Die Zeitschrift der Schweizerischen Hagel-Versicherungsgesellschaft*, *3*, 6–7.

Treloar, A. (1998), Vertically integrated radar reflectivity as an indicator of hail size in greater sidney region of australia, *American Meteorological Society, 19th Conference on Severe Local Storms*, pp. 48–51.

Turner, B., I. Zawadzki, and U. Germann (2004), Predictability of precipitation from continental radar images. part iii: Operational nowcasting implementation (maple), *Journal of Applied Meteorology*, *43*(2), 231–248.

Tustison, B., D. Harris, and E. Foufoula-Georgiou (2001), Scale issues in verification of precipitation forecasts, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *106*(D11), 11,775–11,784.

Uijlenhoet, R., and J. Pomeroy (2001), Raindrop size distributions and radar reflectivity? rain rate relationships for radar hydrology, *Hydrology and Earth System Sciences Discussions*, *5*(4), 615–628.

Vetterli, M. (2016), Hagel: Eine webreportage zu einem der letzten wettergeheimnissen. beobachter webreportagen. http://webreportagen.beobachter.ch/hagel/.

Villarini, G., and W. F. Krajewski (2010), Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall, *Surveys in Geophysics*, *31*(1), 107–129.

Waldvogel, A., B. Federer, and P. Grimm (1979), Criteria for the detection of hail cells, *Journal of Applied Meteorology*, *18*(12), 1521–1525.

Webb, J., D. Elsom, and D. Reynolds (2001a), Climatology of severe hailstorms in great britain, *Atmospheric Research*, *56*(1), 291–308.

Webb, R. M., A. Treloar, J. Colquhoun, R. Potts, J. Bally, T. Keenan, and P. May (2001b), Overview of sydney weather during the forecast demonstration project. in preprints of 30th international conference on radar meteorology, 19–24 july 2001. munich, germany: 477–479. american meteorological society: Boston.

Weisman, M. L., and J. B. Klemp (1982), The dependence of numerically simulated convective storms on vertical wind shear and buoyancy, *Monthly Weather Review*, *110*(6), 504–520.

Wexler, R. (1948), Rain intensities by radar, *Journal of Meteorology*, *5*(4), 171–173.

Weygandt, S. S., A. F. Loughe, S. G. Benjamin, and J. L. Mahoney (2004), Scale sensitivities in model precipitation skill scores during ihop, in *Preprints, 22nd Conf. on Severe Local Storms, Hyannis, MA, Amer. Meteor. Soc. A*, vol. 16.

Wieringa, J., and I. Holleman (2006), If cannons cannot fight hail, what else?, *Meteorologische Zeitschrift*, *15*(6), 659–669.

Wiggins, A., and K. Crowston (2011), From conservation to crowdsourcing: A typology of citizen science, in *System Sciences (HICSS), 2011 44th Hawaii international conference on*, pp. 1–10, IEEE.

Wilks, D. S. (2006), Forecast verification, *Statistical methods in the atmospheric sciences*, pp. 260–268.

Willemse, S. (1995), A statistical analysis and climatological interpretation of hailstorms in switzerland, Ph.D. thesis, Diss. Naturwiss. ETH Zürich, Nr. 11137, 1995. Ref.: A. Waldvogel; Korref.: C.-D. Schönwiese; Korref.: D. Heimann.

Wilson, C. J., K. Ortega, and V. Lakshmanan (2009), Evaluating multi-radar, multi-sensor hail diagnosis with high resolution hail reports, in *Preprints, 25th Conf. on Interactive Information Processing Systems, Phoenix, AZ, Amer. Meteor. Soc. P*, vol. 2.

Wilson, J. W., R. D. Roberts, C. Kessinger, and J. McCarthy (1984), Microburst wind structure and evaluation of doppler radar for airport wind shear detection, *Journal of Climate and Applied Meteorology*, *23*(6), 898–915.

Wilson, J. W., E. E. Ebert, T. R. Saxen, R. D. Roberts, C. K. Mueller, M. Sleigh, C. E. Pierce, and A. Seed (2004), Sydney 2000 forecast demonstration project: convective storm nowcasting, *Weather and forecasting*, *19*(1), 131–150.

Witt, A., M. D. Eilts, G. J. Stumpf, J. Johnson, E. D. W. Mitchell, and K. W. Thomas (1998), An enhanced hail detection algorithm for the wsr-88d, *Weather and Forecasting*, *13*(2), 286–303.

Zepeda-Arce, J., E. Foufoula-Georgiou, and K. K. Droegemeier (2000), Space-time rainfall organization and its role in validating quantitative precipitation forecasts, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *105*(D8), 10,129–10,146.

List of Figures

List of Tables

Declaration

under Art. 28 Para. 2 RSL 05

Last, first name: Noti, Pascal-Andreas

Matriculation number: 10-608-537

Programme: Master of Science in Climate Sciences
with special qualification in Atmospheric Science

Bachelor ☐ Master X Dissertation ☐

Thesis title: Hailstorms over Switzerland:
Verification of Radar-Based Hail Detection Algorithms
with Crowd-Sourced Data and Hail Sensor Data

Thesis supervisor: Prof. Dr. Olivia Romppainen-Martius

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, except where due acknowledgement has been made in the text. In accordance with academic rules and ethical conduct, I have fully cited and referenced all material and results that are not original to this work. I am well aware of the fact that, on the basis of Article 36 Paragraph 1 Letter o of the University Law of 5 September 1996, the Senate is entitled to deny the title awarded on the basis of this work if proven otherwise. I grant insprection of my thesis.

Brig, 9.11.2016

Pascal-Andreas Noti