https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

An open-source radar-based hail damage model for buildings and cars

Timo Schmid[1,2], Raphael Portmann[3], Leonie Villiger[1,2], Katharina Schröer[4], and David N. Bresch[1,2]

[1]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland
[2]Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland
[3]Agroscope Reckenholz, Zurich, Switzerland
[4]Faculty of Environment and Natural Resources, Albert-Ludwigs-University Freiburg, Freiburg, Germany

Correspondence: Timo Schmid (timo.schmid@usys.ethz.ch)

Abstract. Severe hailstorms cause substantial damages to buildings and vehicles, necessitating the quantification of associated risks. Here, we present a novel open-source hail damage model for buildings and cars based on single-polarization radar data and 250'000 geolocated hail damage reports in Switzerland from 2002 to 2021. To this end, we conduct a detailed evaluation of different radar-based hail intensity measures at 1 km resolution and find that the maximum expected severe hail size (MESHS) outperforms the other measures, despite a considerable false alarm ratio. Asset-specific hail damage impact functions for buildings and cars are calibrated based on MESHS and incorporated into the open-source risk modelling platform CLIMADA. The model successfully estimates the correct order of magnitude for the number of building damages in 91%, their total cost in 77%, the number of vehicle damages in 74%, and their total cost in 60% of over 100 considered large hail events. We found considerable uncertainties in hail damage estimates, which are largely attributable to limitations of radar-based hail detection. Therefore, we explore the usage of crowdsourced hail reports and find substantially improved spatial representation of severe hail for individual events. By highlighting the potential and limitations of radar-based hail size estimates, particularly MESHS, and the utilization of an open-source risk modelling platform, this study represents a significant step towards addressing the gap in risk quantification associated with severe hail events in Switzerland.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

1 Introduction

Severe hail storms constitute one of the leading damage drivers for buildings and cars in Switzerland, with a recent extreme event on 28 June 2021 causing building damages of 400 million Swiss francs (CHF) in a single canton (GVL, 2022). A potential increase in hail severity under climate change (Raupach et al., 2021) may further exacerbate such hail damages in the future. Structural damages to the roof and blinds are typically caused by hailstones with diameters larger than 2.5 cm (Stucki and Egli, 2007) and dents in the vehicle body from 2 cm (Hohl et al., 2002b). However, large damages such as destroyed roof tiles or broken car windows are only expected for hailstones greater than approx. 4–5 cm (Yeo et al., 1999; Heymsfield and Wright, 2014; Púčik et al., 2019).

Hail is formed when supercooled water is collected by a hail embryo (e.g. frozen drop) in the updraft of a convective storm (Allen et al., 2020). A key factor for the final hail size is the residence time of a hailstone within the hail growth zone and the supply of supercooled liquid in the thundercloud. Thus, hailstone sizes are primarily limited by the updraft speed in a thunderstorm cell which determines how large hailstones can grow before their fall speed exceeds the upwind velocity. However, this dependence is not a direct relationship because the spatial structure of local updrafts and horizontal winds in each individual thunderstorm influence the trajectory of hailstones (Kumjian and Lombardo, 2020).

The local hail intensity is primarily expressed by the maximum hailstone size or the hail kinetic energy (E_{kin}), although other measures such as the number concentration or the ice mass also exist (Grieser and Hill, 2019). Neither E_{kin} nor hailstone sizes can be measured directly, except for point measurements with hail pads (Punge and Kunz, 2016), hail sensors (Kopp et al., 2023), or verified crowdsourced reports, e.g. from the European Severe Weather Database (ESWD; Dotzek et al., 2009). Thus, for data with consistent spatial and temporal coverage, we rely on radar measurements. When using radar-based hail intensity estimates, it is important to know that hail stone sizes can approach or exceed radar wavelengths leading to resonance scattering effects (e.g. Bohren and Battan, 1982). Additionally, individual stones can have complex shapes with many side lobes, and may be coated with water (Atlas et al., 1960), which makes a direct hail size estimate from operational weather radars challenging.

Radar-based estimates of hail intensity in Switzerland include the probability of hail (POH) and maximum expected severe hail size (MESHS), which are based on an algorithm originally derived in Waldvogel et al. (1979) depending on the height between the 0 °C level and the highest elevation with a radar reflectivity threshold of 45 and 50 dBZ, respectively. Furthermore, hail kinetic energy has been estimated directly from radar reflectivity (Waldvogel et al., 1978; Hohl et al., 2002a). The radar network of MeteoSwiss (Fig. 1) consisted of three single-polarized C-band radars (Albis, Monte Lema, La Dôle) until it was expanded by two more (Plaine Morte, Weissfluh) in 2012. Also in 2011/12, all five radars were equipped with dual-polarization technology, which lead to improved clutter filtering (Germann et al., 2015) and the development of a hydrometeor classification algorithm that includes hail (Besic et al., 2016, 2018), but no hail intensity metric based on dual-polarization data is available to date.

In other European countries with C-band radar networks, it has been shown that hail can successfully be detected with the criterion of Waldvogel et al. (1979), on which also POH and MESHS are based. Puskeiler et al. (2016) found the highest skill in

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

detecting hail days over 7 years for an adjusted version of the Waldvogel et al. (1979) criterion, compared to a simple reflectivity threshold of 55 dBZ (Mason (1971) criterion). Analyzing one summer in the Netherlands, Holleman et al. (2000) found the Waldvogel et al. (1979) criterion to best estimate the probability of hail compared to Vertically Integrated Liquid (VIL; Amburn and Wolf, 1997), the severe hail index (SHI, see below), the Mason (1971) criterion, and a cloud-top temperature-based method by Auer (1994). Skripniková and Řezáčová (2014) developed a combination criterion of the Waldvogel et al. (1979) method, the SHI (see below), and a SHI-derived quantity called "probability of severe hail" to best detect large hail (>2 cm) in Germany and the Czech Republic from 2002-2011.

In the US, with an operational S-band radar network, the SHI (Witt et al., 1998) and the derived "maximum estimated size of hail" (MESH) are widely used. The SHI is calculated as the vertical integral of a reflectivity-based hail kinetic energy flux estimate (Waldvogel et al., 1978) weighted by the height of the reflectivity in relation to the 0 and -20 $^\circ\mathrm{C}$ height.

Accurate hail damage modelling is crucial for insurance companies, resulting in many proprietary models and few openly available studies on the topic. Hohl et al. (2002a, b) investigated relationships of radar-derived hail kinetic energy to car and building damage data per Swiss community level. The authors derived vulnerability curves using the relationship of radar reflectivity (Z) and E_{kin} developed by Waldvogel et al. (1978) with a calibration based on nine and twelve hail cells for buildings and cars, respectively. The obtained results also led to the development of the first probabilistic hail model for Europe, which was later purchased by "Risk Management Solutions" (Miller, 2007). Furthermore, Schmidberger (2018) developed a hail damage model for Germany with the primary purpose to serve as a risk assessment tool. It derives hail tracks from radar data and hail sizes from the ESWD. The damage model considers separate impact functions for different building types, where each function is split into damages at the building exterior and a value for the roof-penetrating damages. Expected damages with high return periods are modelled with a stochastic simulation of 10'000 years with hail cell tracks and hailstone sizes. Lastly, Yin et al. (2007) developed an event-based hail risk model for cars in the US, based exclusively on bias-corrected hail reports. Hail losses are estimated by simulating 10'000 years of hail events with a Monte Carlo approach, and using three separate impact function for dent repair labour cost, parts replacement, and depreciation costs.

While the existing models allow for a static, probabilistic hail risk assessment, they are not open-source and lack the capability to provide real-time hail impact estimates. This paper explores the potential of existing radar-based hail intensity measures to model damages to buildings and cars and is structured as follows: First, the radar, hail damage, and exposure data from Switzerland are introduced in Section 2. Leveraging 20 years of geolocated per-building damage data, we conduct a detailed verification of available radar products at 1 km resolution (Sect. 3) and find MESHS to be best performing. Within the CLIMADA framework (Aznar-Siguan and Bresch, 2019), we develop an open-source hail damage model (Sect. 4) with focus on real-time hail damage assessments, following the concept of risk as used by, e.g., the IPCC (Pörtner et al., 2022), namely the combination of hazard, exposure and vulnerability. A detailed model evaluation (Sect. 5) highlights the skill and the limitations of the MESHS-based hail damage model for buildings and cars, also in comparison with other radar variables.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

2 Data

To build a hail damage model, three datasets are required: hazard, informing about occurrence and intensity of hail (Sect. 2.1); exposure, informing about location and value of assets (Sect. 2.2; 2.3); and damage, informing about vulnerability of assets to hail (Sect. 2.2; 2.3). The datasets are described in the following. All hazard data in daily resolution comprise sub-daily information from 6 UTC to 6 UTC. This threshold marks the daily minimum of hail activity in the domain (Nisi et al., 2016; Schroeer et al., 2023) and so allows defining physically consistent hail events by minimizing the splitting of hail events over two calendar days.

2.1 Radar data

The five MeteoSwiss radars scan 20 elevation sweeps every five minutes (Fig. 1; Germann et al., 2015). Here, we use five different radar products: MESHS, reflectivity, VIL, and E_{kin} are used as hazard variables and POH is used for the pre-processing of damage data. Both POH and MESHS are based on the algorithms developed by Waldvogel et al. (1979) and extended by Witt et al. (1998). POH is calculated from the relation between the 45 dBZ contour height of radar reflectivity and the freezing level derived from the operational weather model COSMO (Baldauf et al., 2011). It thus represents an estimate for the zone where hail may grow by riming in deep convective storms (Betschart and Hering, 2012). MeteoSwiss is using the polynomial fit by Foote et al. (2005) to convert the height difference to a probability of hail (0-100%). Similarly, MESHS is calculated as the height difference between the 50 dBZ contour and the freezing level, following the approach of Treloar (1998) and its first operational implementation by Joe et al. (2004) during the Sydney 2000 Forecast Demonstration project. In contrast to POH, MESHS is an estimate of the maximum hailstone size for values from 2 cm and above. While the algorithm has no upper limit, uncertainties are high particularly above 6 cm because of verification data limitations. Exact formulas and visualizations for the MeteoSwiss implementation of the POH and MESHS algorithms are provided in Trefalt et al. (2023).

The second hazard variable used in this study, E_{kin}, can also be estimated from radar reflectivity (e.g. Waldvogel et al., 1978) and has been used to model hail damages to cars and buildings by Hohl et al. (2002a). However, the method has large uncertainties and was originally calibrated with observations from an S-band radar. Cecchini et al. (2022) provide updated E_{kin} estimates for C- and S-band radars separately, based on the same dependence to radar reflectivity but with adjusted coefficients. We here use the C-band specific coefficients of Cecchini et al. (2022), but follow the approach of Hohl et al. (2002a) by integrating over time and using a low-level Constant Altitude Plan Position Indicator (CAPPI) in order to compare our model with previous hail damage modelling approaches. In a first step, kinetic energy flux ($\dot{E}$, in $\mathrm{J\,m^{-2}\,s^{-1}}$) is calculated as:

$$\log(\dot{E}) = -6.72 + 0.114Z \tag{1}$$

where Z is the radar reflectivity in dBZ at a $2\,\mathrm{km}$ CAPPI. In a second step, daily integrated values of E_{kin} are calculated as sum of $\dot{E}$ over all 5 min time steps of a day (6–6 UTC).

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

The vertical maximum reflectivity, which is directly measured by the radar, is used as a third hazard variable, as it has previously also been used as a proxy for hail intensity (e.g. Mason, 1971; Kunz and Puskeiler, 2010). Lastly, the VIL is quantified from vertically integrated radar reflectivity, which is converted into liquid water equivalent according to Greene and Clark (1972).

Only POH and MESHS are available in sufficient quality already from 2002-2012 due to extensive re-processing by Trefalt et al. (2023). Reflectivity, E_{kin}, and VIL are used from 2013, where dual-polarisation radars allow for an automatic clutter filtering of sufficient quality without re-processing (Germann et al., 2015).

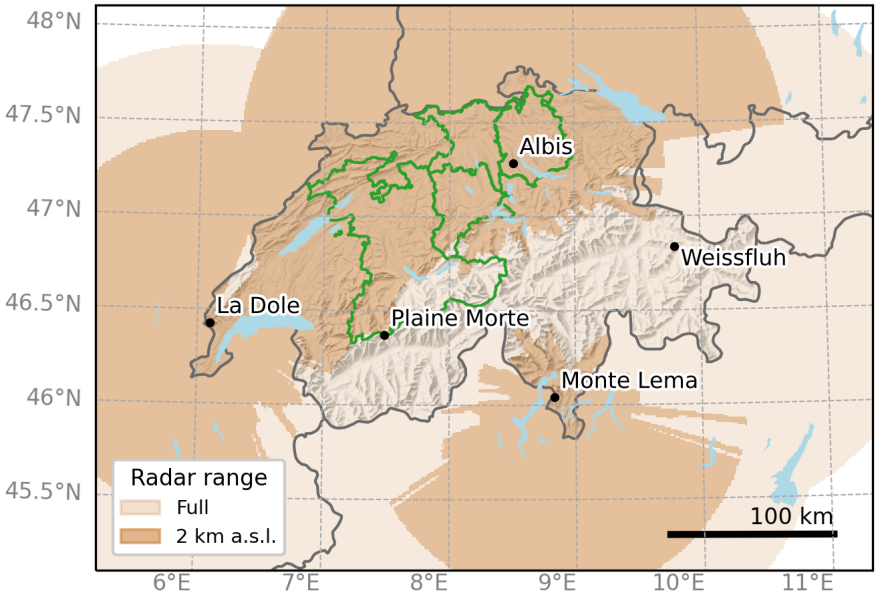

Figure 1. Radar locations in Switzerland with the full scanning range of 160 km (light brown) and the reduced scanning range at 2 km a.s.l. relevant for E_{kin} (dark brown). The four cantons (from west to east: Berne, Lucerne, Aargau, Zurich) with available building damage data are outlined in green.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

2.2 Building Data

Exposure and damage data from four cantonal building insurances in Switzerland are used: GVZ (Zurich), GVL (Lucerne), AGV (Aargau), GVB (Berne). These cantons represent 43% of the Swiss population. Damage reports are available from 2002-2021 and exposure data consist of the complete buildings stock (>99% insurance fraction) in 2021. Since spatial coordinates were not available from some of the insurances, for Lucerne and Aargau addresses were geocoded using the OpenStreetMap API Nominatim (Haklay and Weber, 2008). This introduces some additional uncertainty as, e.g. buildings with an unknown street number are assumed to be located in the center of their street. Furthermore, few addresses (<0.5%) were not geocoded within their canton and the corresponding buildings were removed from the exposure data. In total, the dataset contains the spatial coordinates, building value (in CHF), and year of construction of 989'000 buildings with the reference year 2021. The buildings are distributed over 89% of all 1 km gridcells within the four cantons, with 67% containing ten or more buildings (Fig. 2a). Average building values range from mostly below 1 million CHF in rural regions to over 10 million CHF in the cities of Zurich, Berne, and Lucerne, where land prices are high and buildings are larger on average (Fig. 2b).

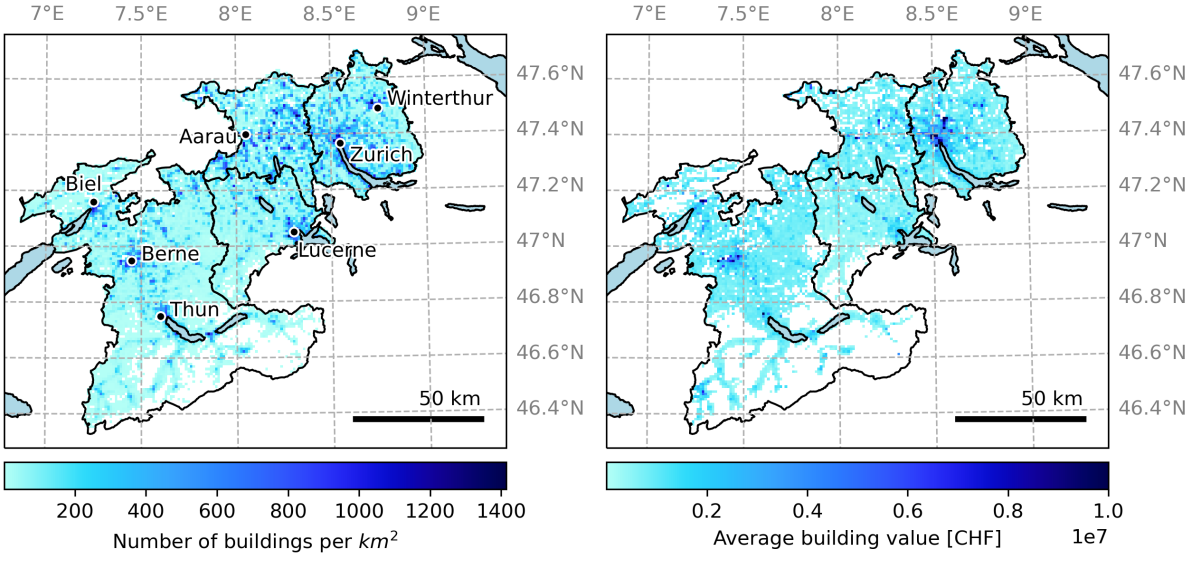

Figure 2. **(a)** Number of buildings per square kilometre in the four Swiss cantons with available data. **(b)** Average building value (for all 1 km^2 gridcells with >10 buildings).

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

2.2.1 Building damages

The building damage data include 250'000 reports with spatial coordinates, date, total damage in CHF (incl. deductible), and a building ID to match it with the exposure data. To ensure consistent damage and exposure data, our dataset does not contain damage reports of buildings that are not part of the 2021 portfolio (i.e. torn-down buildings). The total claim volume over the considered 20-year period is 1.39 billion CHF of which 90% (1.25 billion CHF) is caused by the strongest 30 hail days, with a single event on 28 June 2021 causing 35% (483 million CHF).

Some pre-processing of the damage data is required to use them for damage modelling. First, 38'787 damage claims report damages of zero and are removed, leaving 212'026 valid damage claims. Entries with zero represent claims that were not accepted as valid hail damages by the insurance. Second, all damage claims are indexed to the year 2021 using a canton-specific building cost index provided by the four building insurances. Finally, accurate temporal reporting of hail damages is not always guaranteed, in particular for damages that occur at night or, in fewer cases, when residents are not at home during a hail storm. Thus, damage data are pre-processed using a plausibility filter to correct for obvious misattributions.

Firstly, areas with plausible hail for each day are derived by requiring a POH of 10% with a 5 km buffer to account for the 2–4 km wind drift of hailstones (Hohl et al., 2002b; Barras et al., 2019). 10% POH represents a conservative threshold as damaging hail (>2cm) is expected for a POH of >80% (Saltikoff et al., 2010), but Nisi et al. (2016) note that soft hail or graupel in Switzerland can be detected in lower POH values of 20–50%. If within ±2 days of a hail damage report, no plausible hail is observed, the damage report is removed (3'460 of 212'026 cases). If only one day has plausible hail within ±2 days at the location of a hail damage report, the report is assigned to this day. If hail is plausible on more than one day, the date is only changed if another day has at least a 50% higher POH than at the originally reported date. The correction is applied to 57'432 of 212'026 damage reports, of which 92.5% were corrected to the previous day, indicating that damages were reported on the day after a hail storm. As hail activity often continues into the night and over midnight, this shifting of damage claims aligns the claims with consistent hail events (6–6UTC; Sect. 2). While the chosen POH-based pre-processing is not fully independent of other radar-based hazard data, it is required when working with insurance claim data and thresholds are chosen conservatively compared to other studies (e.g. assigned hail damage claims within ±30 days of the reported date in Warren et al., 2020).

The pre-processed building damage data reveal almost log-normally distributed values averaging at 6'650 CHF and ranging from 395 to 21'900 CHF (5-95% quantile), with 1187 claims above 100'000 CHF (Fig. 3). The peaks at CHF 500 and CHF 5000 stem from cases where the insurance paid a fixed amount rather than the actual repair costs. As expected, claim values increase with the value of the affected building, but not linearly (parallel to the 1:1 line in Fig. 3). On average, the damage is proportional to the building value to a power of 0.27, indicating a decrease in relative damage with increasing building values. This decrease reflects the fact that while the value of a building is approximately proportional to its volume, hail damages are proportional to the area (and value) of exposed vulnerable building parts (blinds, roof, and facade) which scale roughly with the building surface area. Within this study, we use both, the absolute building value and a scaled version which approximates the value of exposed vulnerable buildings parts as being proportional to the building value to a power of 0.27.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

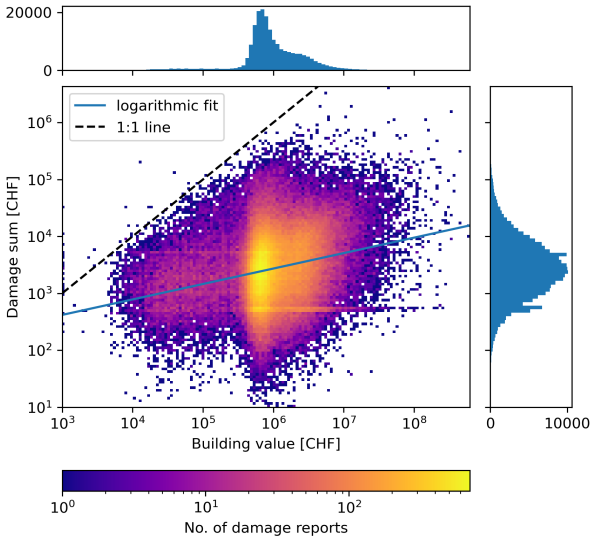

Figure 3. Density plot of building values vs. reported damage for the 212'026 valid damage claims in the considered cantons from 2002–2021. A linear fit (in the logarithmic space) is shown in blue, and the 1:1 line in black.

2.3 Car data

Car damage and exposure data for the model calibration are provided by a private insurance company for the years 2017-2021. Each year, over 500'000 cars are insured over all Swiss cantons, with a total value of over 10 billion CHF. Reported claims vary strongly from year to year, with a total of over 50'000 claims over the five-year period 2017–2021. Due to conditions of the data provider, per-event reported car damages are only explicitly shown as normalized values.

In contrast to buildings, the exact location of a car (and whether it is directly exposed to hail or covered for protection) at any given time is unknown. Thus, the location of each vehicle is inferred from the municipality of the registered address of its most frequent driver (mostly the owner). Since we are interested in modelling the portfolio losses, rather than an individual car, the location of each car is assumed to be a random point within the municipality. The average area per Swiss municipality is 12.9 km^2 with larger areas in mountainous regions and smaller values in cities where the density of cars is highest.

Given the uncertainty in the spatial coordinates of a vehicle, the POH-based plausibility filtering must be adjusted, compared to the building damages. Rather than defining if a claim is in an area with plausible hail, for each day we create a likelihood function that a car registered at a given coordinate could have driven into a hail storm. Based on the average Swiss municipality area of 12.9 km^2 and the average car ride distance in Switzerland (14,8 km with 63% of all rides <10 km; Biedermann, 2023) we select a 50 km radius which covers most cars, as visually determined by analyzing reported car damage claims during large events with known hail extent. We assume the likelihood to be proportional to the number of gridcells with a POH>10% (see Sect. 2.2.1) within the 50 km radius. As damages are particularly often reported the day after a hail event (due to nocturnal hailstorms), the date is changed if the likelihood at the location of a report is over twice as large as on the reported date. For 2

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

days before or 1–2 days after, the date is only changed if the likelihood is >10 times higher than on the reported date (usually cases where no POH at all is detected on the date of the report). This approach removes 1.7% of all claims where no POH within 50 km is reported, and moves 9.5% to the previous day and 4.9% to ± 2 days. Of course, individual cars drive further than 50 km and a few correct damage reports will be shifted to a different date, which is unavoidable when working with spatio-temporally imprecise damage reports.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

3 Evaluation of radar-based hail intensity measures

Before calibrating a hail damage model, we evaluate the available radar-based hail intensity metrics using the geolocated building damage and exposure data. The focus of this evaluation lies on the occurrence of damaging hail and its spatial match with radar data rather than its intensity, which will only be used in the model calibration in Sect. 4. Since exact spatial coordinates of reported hail damages are required to verify occurrences of damaging hail, only data for buildings and not cars are used in this section. Both damage and hazard data are classified to binary values. Each $1\,\mathrm{km}^2$ gridcell with 10 or more buildings is classified as hail damage (yes/no) on each date. Hail intensity metrics are divided into gridcells with damaging hail expected (yes/no) by a threshold which is varied over the respective intensity range (e.g., from 20 to 80 mm for MESHS). For each variable and intensity threshold, we so obtain hits (H), false alarms (FA), misses (M), and correct negatives (CN) to create a contingency table (Table 1). From these fractions, the skill metrics probability of detection (POD), false alarm ratio (FAR), and Heidke-skill-score (HSS; Wilks, 2019) are calculated as follows:

$$POD = \frac{H}{H+M} \tag{2}$$

ranging from 0 (no skill) to 1 (perfect model).

$$FAR = \frac{FA}{FA+H} \tag{3}$$

ranging from 0 (perfect model) to 1 (no skill).

$$HSS = \frac{2\cdot(H\cdot CN - FA\cdot M)}{(H+M)(M+CN)+(H+FA)(FA+CN)} \tag{4}$$

ranging from -1 to 1 (perfect model), with positive values denoting better and negative values lower skill than a random guess.

Table 1. Standard 2x2 contingency table for dichotomous events, used for validation of damaging hail occurrence.

	damage observed	no damage observed
hail predicted	Hits (H)	False alarm (FA)
no hail predicted	Misses (M)	Correct Negatives (CN)

Figure 4 shows the skill scores for MESHS, maximum reflectivity, and E_{kin} over their whole intensity range. As expected, both POD and FAR decrease with an increasing intensity threshold for each variable. Thus, there is a trade-off between detecting most hail events at a low threshold and obtaining few false alarms at high thresholds. The HSS considers both of these effects and peaks at intermediate thresholds. Note that an evaluation on a 1 km gridcell level yields more conservative skill estimates relative to studies with less accurate damage data (e.g. on community level; Skripniková and Řezáčová, 2014; Kunz and Kugel, 2015; Puskeiler et al., 2016; Nisi et al., 2016). Thus, both POD and FAR are also shown with a 4 km buffer which accounts for horizontal drift of hail (Barras et al., 2019). For the POD_{4km}, a gridcell with a damage report is considered a hit

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences Discussions

if the hazard threshold is reached anywhere within 4 km. For the FAR_{4km}, a gridcell which exceeds the hazard threshold (i.e. hail damage expected) is only considered a false alarm if there is no damage report within 4 km.

At the minimum MESHS value of 20 mm, the POD reaches 60% for an exact spatial match (and 80% with a 4 km buffer), indicating that 60% of all gridcells with a hail damage report lie within a MESHS footprint. When analyzing individual claims (not shown), detection probabilities of 88% (97% with 4 km buffer) are reached, indicating that missed events are often cells with few claims, while cells with many claims are more likely to be within a MESHS footprint. The gridcell-wise FAR decreases almost linearly with increasing MESHS from over 80% at 20 mm to 50% at 80 mm. Thus, even at extreme MESHS values only one in two gridcells with >10 buildings contains a damage claim and in 20% of cases there is not a single damage claim within 4 km. Overall, the skill of MESHS is highest between 20 and 50 mm with a HSS of >0.2 and a peak of 0.3 at MESHS equal to 35 mm.

In contrast to MESHS, maximum reflectivity is defined on a continuous scale without minimum value and the shown range starts at 40 dBZ, where both POD and FAR reach almost 100%. As POD and FAR decrease, maximum reflectivity starts showing some predictive skill with the highest HSS values of over 0.2 in a window between 55 and 65 dBZ. The enhanced HSS above 50 dBZ, which is used as threshold in MESHS, indicates that there is additional information regarding the occurrence of damaging hail in high reflectivity values up to 65 dBZ.

E_{kin} values are distributed in a long-tailed distribution with few high values and the majority below 200 $\mathrm{Jm^{-2}}$. Thus, the POD decreases rapidly while the FAR decreases linearly with increasing E_{kin} reaching 60% at 1000 $\mathrm{Jm^{-2}}$ and 20% with a 4 km buffer, comparable to the FAR of MESHS at 80 mm. However, few gridpoints show such high E_{kin} values leading to low POD and, accordingly, HSS. The highest predictive skill is achieved between 200 and 400 $\mathrm{Jm^{-2}}$ with a HSS of 0.2. Lastly, VIL (not shown) has lower HSS and higher FAR than the three variables shown in Fig. 4 indicating a weak distinction of hail and non-hail events even at high values.

A maximum HSS of 0.3 over all variables shows that there is no radar variable threshold that can reliably identify all areas with damaging hail on a gridcell level. Spatial footprints in Fig. 4 and A1 show that within an intense hail streak nearly all gridcells consistently contain a damage claim, but no radar variable consistently distinguishes these hail streaks. Of the three analyzed radar metrics, MESHS stands out as most promising with an extended range of relatively high HSS which offers good trade-off between POD and FAR. However, in particular the high FAR confirms that MESHS cannot be interpreted as *actual* local hail size, but (as the name suggests) as *maximum expected* hail size, given the storm intensity. For a deeper discussion of MESHS and local hail size see Schroeer et al. (2023). Since no radar-based proxy corresponds accurately to hail size or energy, the calibration in the following section is based on empirical hazard-vs-damage relationships, rather than physical considerations of the vulnerability of buildings materials to hail sizes (e.g. Stucki and Egli, 2007; Macdonald and Stack, 2021).

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

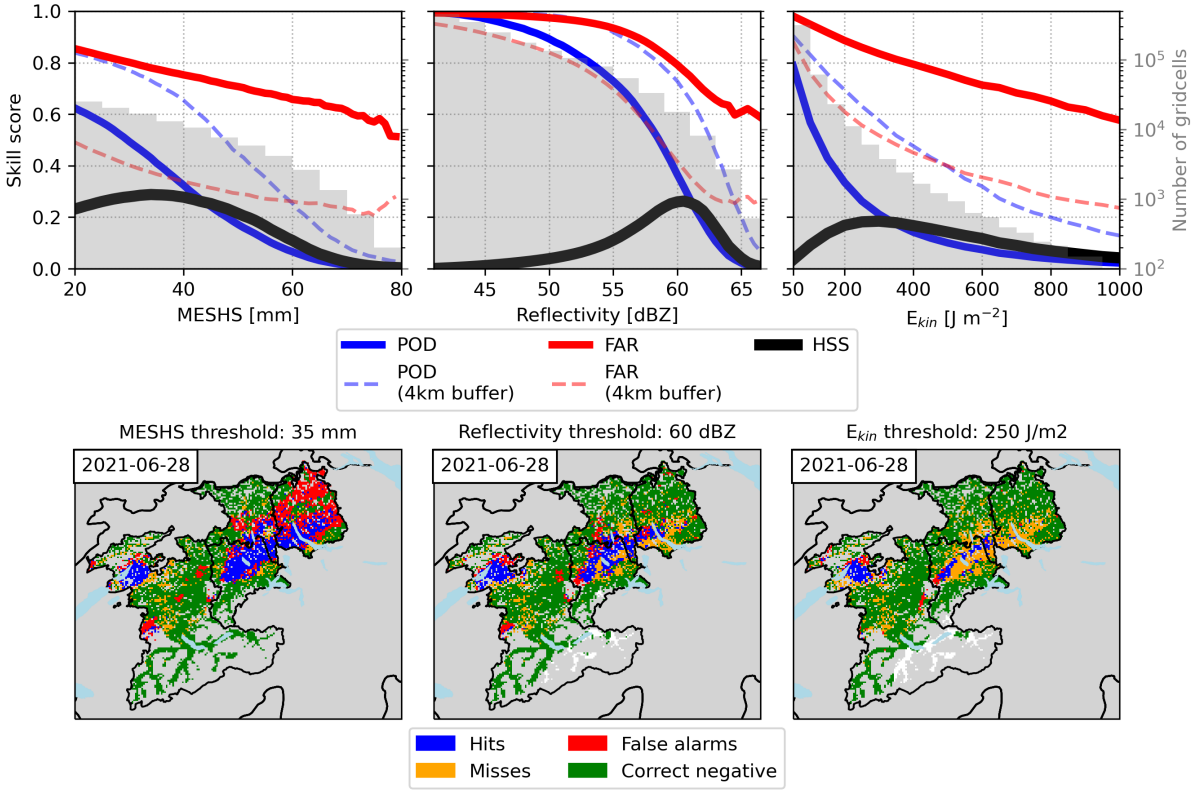

Figure 4. (Upper panel) MESHS, maximum reflectivity, and E_{kin} verification statistics probability of detection (POD), false alarm ratio (FAR), and Heidke-skill-score (HSS) calculated per gridcell for all 1km cells with >10 buildings. Grey bars show the number of gridcells with given hazard intensity over the whole time period with a logarithmic scale. **(Lower panel)** Visualization of contingency table variables per gridcell for a single event (28 June 2021) and selected hazard thresholds corresponding to the maximum HSS. Grey areas do not contain sufficient exposure data (<10 buildings) and white areas mark cells with missing hazard data.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences Discussions

4 Model calibration

We build a hail damage model for each of the introduced radar variables MESHS, maximum reflectivity, and E_{kin}. Based on the analysis in the previous section, we here focus on the MESHS-based model. Following the concept of risk as a combination of hazard, exposure, and vulnerability (IPCC; Pörtner et al., 2022), the model estimates hail damages by translating a hazard intensity (e.g. MESHS) to a relative damage with an impact function. This relative damage is translated to absolute values (e.g. number of damaged buildings or damage sum in CHF) by multiplication with the exposure data in a spatially explicit fashion (Aznar-Siguan and Bresch, 2019).

Impact functions for different natural hazards are typically calibrated by assuming a parameterized impact function (e.g. sigmoid function) and minimize a loss function between modelled and observed per-event damages (e.g. Schwierz et al., 2010; Eberenz et al., 2021; Lüthi et al., 2021). In this study, the point-based building damage data allow for a spatially explicit calibration which can additionally inform about the shape of the impact function and provide detailed uncertainty estimates. The approach consists of the following steps:

1. Select all days with non-zero hazard intensity and/or a damage report from the considered time period of April–September (i.e. convective season) 2002-2021 (806 days for MESHS and the building damage reports).

2. Assign the local hazard intensity to each damage report and exposure point in the dataset (for events before 2021 the buildings that were not yet built are removed).

3. Aggregate exposure and impact variables (number and value of exposed buildings/cars, number of claims, damage claim sum) to hazard intensity levels (MESHS values from 20-100 mm in 1 mm intervals). Note that for high hazard intensities data points become increasingly sparse.

4. For each hazard intensity level, calculate the average percent of assets affected (PAA) and mean damage ratio (MDR):

$$PAA = \frac{N_{dmg}}{N_{exp}}, \quad MDR = \frac{V_{dmg}}{V_{exp}} \tag{5}$$

where N_{dmg} is the number of damaged assets, N_{exp} is the number of exposed assets, V_{dmg} the damage costs, and V_{exp} the monetary value of exposed assets. For high hazard intensities purely empirical PAA and MDR become more uncertain and eventually unstable due to under-sampling of the underlying distribution (see Sect. 4.1).

5. Calculate two types of impact functions from the raw impact data per hazard intensity. First, calculate a monotonically increasing smooth-empirical fit by using a running mean (10 mm running mean for MESHS). Given the aforementioned under-sampling, above a certain hazard threshold, the smooth-empirical impact function is given as constant by the average PAA/MDR for all values above this threshold (60 mm for MESHS). Secondly, fit a sigmoid-type function as proposed by Emanuel (2011) by minimizing the root-mean-squared error (RMSE) between the function and the purely empirical PAA or MDR:

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

$$f = \frac{v_n^3}{1+v_n^3} \cdot S, \quad v_n = \frac{MAX[(V - V_{thresh}), 0]}{V_{half} - V_{thresh}} \tag{6}$$

where f is the relative impact (PAA or MDR), S the scale parameter (i.e. maximum relative loss), V is the hazard intensity (here MESHS), V_{thresh} the hazard intensity below which no damage occurs, and V_{half} the hazard intensity at which the relative loss is half of its maximum (i.e. half of the scale parameter S). The RMSE is minimized using the Nelder-Mead algorithm (Gao and Han, 2012) with parameter bounds informed from the empirical PAA or MDR. For MESHS, V_{thresh} ranges from 0-20 mm, V_{half} from 20 mm to the maximum considered MESHS value (100 mm), and S is constrained within two orders of magnitude of the maximum PAA or MDR of the empirical function (for MESHS-based building damages: PAA: 0.6%–60%, MDR: 0.005%–0.5%).

6. To quantify effects of sampling uncertainty on the impact functions, a bootstrapping is used: Steps 3–5 are repeated for 1000 subsamples, each including 806 days. The subsamples are obtained by randomly selecting 806 days with replacement from all 806 days identified in (1).

4.1 Calibrated impact functions for buildings

Before showing the calibrated impact function (step 5 of the calibration procedure), we shortly consider the observed gridcell-wise variability in PAA and MDR (step 3). For all MESHS intervals shown in Fig. 5, many gridcells with buildings exposed to high MESHS values still have zero reported damages (false alarms) and, thus, zero PAA and MDR. However, the fraction of these false alarms reduces with increasing MESHS (also shown in Fig. 4) while the number of gridcells with high damages of >10% PAA and >0.1% MDR increases (Fig. 5). Note that this increase is more pronounced for the PAA (Fig. 5c) because the MDR is additionally dependent on the building value and, thus, increasingly variable when considering fewer gridcells. The transitioning distribution of PAA and MDR as MESHS increases results in ascending average values of PAA and MDR (step 4) which ultimately determine the shape of the vulnerability curve (step 5). This, in turn, means that the calibrated impact function does reproduce the observed gridcell-wise variability of PAA and MDR. Instead, it assigns the same relative damage (PAA or MDR) to each gridcell with the same MESHS value, corresponding to the average of the distribution in Fig. 5a,b.

The resulting smooth-empirical impact functions for buildings (Fig. 6) start at almost zero at 20 mm MESHS, reaching 5% for PAA and 0.05% for MDR at 60 mm with a bootstrap 90% confidence interval (CI) between 3-7% for PAA and 0.02-0.09% for MDR. The larger relative CI for MDR indicates higher sensitivity... in hail damage sums towards individual events compared to the number of affected buildings. The bootstrap CIs accurately quantify the sampling uncertainty but do not reflect the variability of gridcell-wise PAA and MDR given a certain MESHS value (discussed earlier; Fig. 5a,b). Above 60 mm MESHS, fewer damage reports (red bars in Fig. 6) and fewer affected gridcells (6304 from 40-60mm vs. 1435 from 60-80mm) remain. Hence, and given the long-tailed distribution of PAA and MDR values (Fig. 5a,b), we refrain from continuing the smooth-empirical impact function for these MESHS values.

To provide impact function estimates for higher values, we use a sigmoidal fit (Emanuel, 2011), which is fitted to all data (including values above 60 mm), but weighted by the number of samples (see Sect. 4). This approach avoids overfitting to high

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

MESHS values with few data points and provides vulnerability information for the complete MESHS range, with increasing uncertainty for high values. The sigmoidal fit (Fig. 6, in black) agrees well with the empirical PAA and extends it to reach 5-10% at 80 mm MESHS. For the MDR, the empirical fit exhibits a steeper increase between 50 and 60 mm, which cannot be fully captured by the chosen sigmoidal fit. The sharp increase occurs due to extreme (roof-penetrating) damages that can occur at high MESHS values. However, regarding the final model skill the difference between the empirical fit and the sigmoidal fit is small (Appendix C) as hazard data uncertainties dominate (Sect. 5). For the subsequent damage modelling the sigmoidal fit is used. Note that impact functions were also derived for the hazard variables E_{kin}, maximum reflectivity, and VIL (see Table D1 for the fitted parameters and Sect. 5 for their discussion).

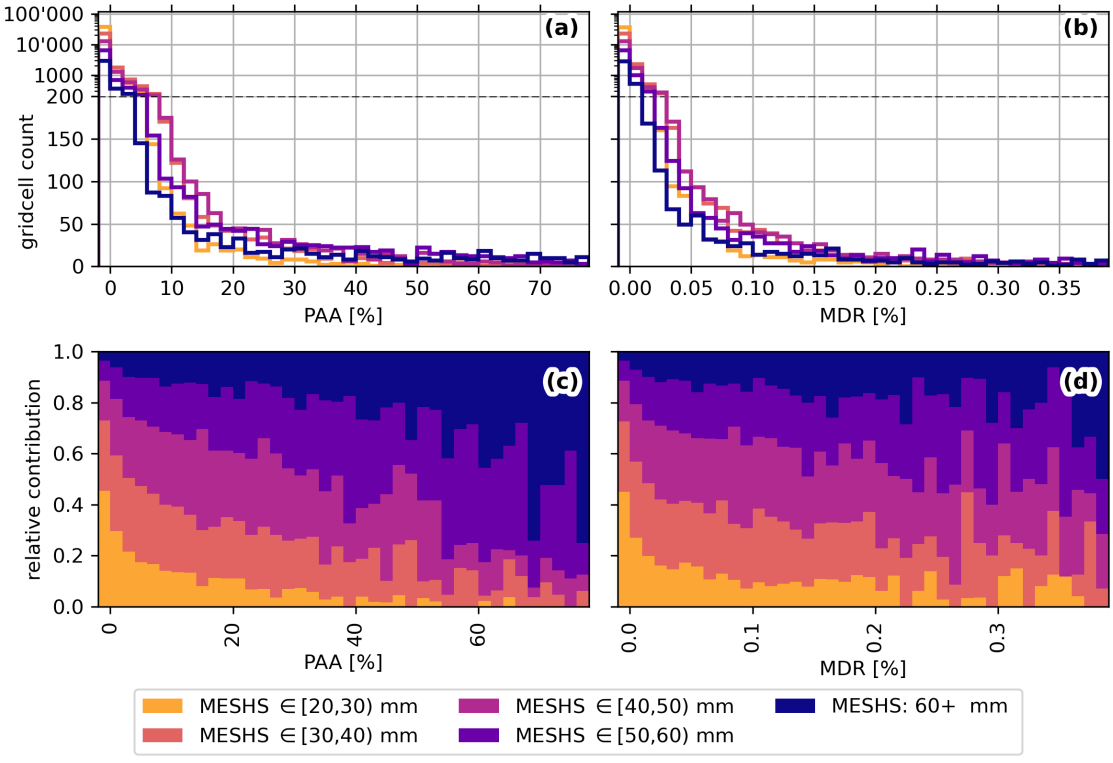

Figure 5. (a,b) Distribution of PAA and MDR values for all 1 km^2 gridcells with 10 or more buildings, split by observed MESHS values. **(c,d)** Corresponding relative contribution of each MESHS class. Note that the first bin comprises gridcells with a PAA and MDR value of zero and contains over 50% of all gridcells.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

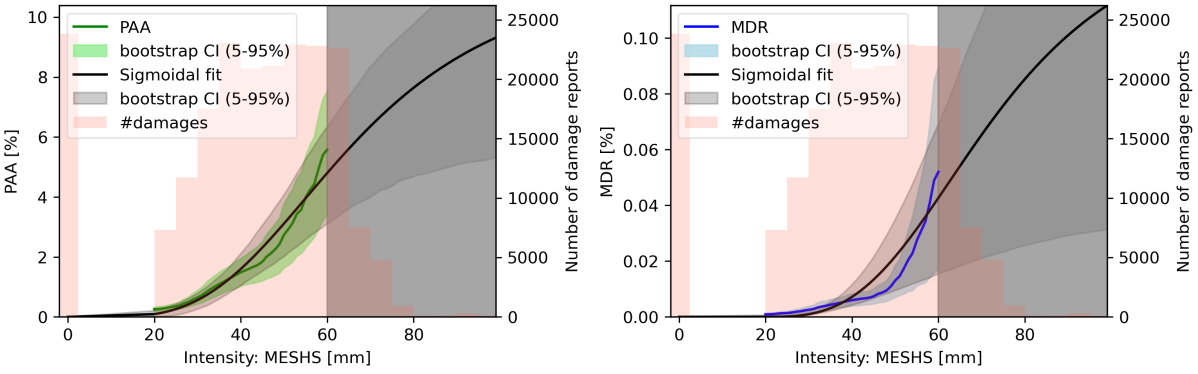

Figure 6. (a) Impact function for the number of damaged buildings, calibrated with data from 2002-2021. The solid green line indicates the 10 mm moving average for the empirical percent of assets affected (PAA) and the green shading its 5-95% bootstrap confidence interval. In black, a sigmoidal fit (Emanuel, 2011) is shown, and pink bars indicate the number of damage reports per MESHS intensity. The grey shaded area indicates MESHS values without sufficient data points for a meaningful empirical fit. **(b)** same as (a), but for damage sums in CHF, with the empirical mean damage ratio (MDR) in blue.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

4.2 Calibrated impact functions for cars

Regarding car damage modelling, the inaccurate spatial coordinates of each individual car constitute an additional challenge for the impact function calibration. As the exposure data contains a large portfolio and the probability distribution of the actual car location decreases with increasing distance to the assumed location in its municipality (Sect. 2.3), an impact function can still be calibrated empirically. As expected, many car damages (42% vs. 12% for buildings) seemingly occur at a MESHS value of zero (Fig. 7), many of which likely were in an area with MESHS when the hail damage occurred. Crucially, the PAA and MDR of the car impact functions include the probability of cars being covered (e.g. in a garage) or located elsewhere. Thus, the impact functions can be directly applied to a car portfolio with the location given by the address of each vehicles' most frequent driver. Empirical impact functions (Fig. 7) show a robust increase with MESHS from a PAA of 0.2% to 2.75% at 60 mm MESHS and a MDR of 0.02% to 0.25%. For both PAA and MDR, the sigmoidal fit follows the empirical function closely, with slightly lower vulnerability at MESHS below 30 mm. The lower PAA compared to building damages is consistent with many cars being covered in a garage, preventing hail damage. In contrast, the high MDR compared to buildings is due to lower exposure values for cars, relative to the inflicted damages by hail. As for the buildings, the simgoidal fit is used for subsequent damage modelling.

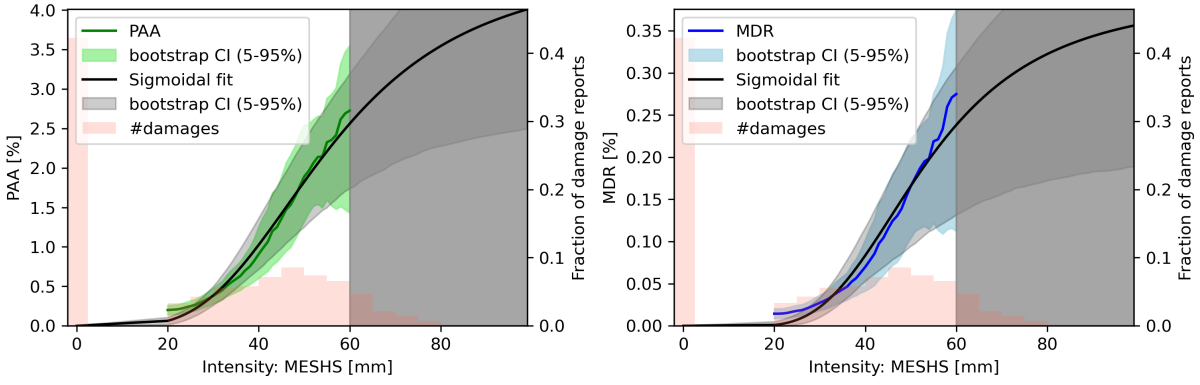

Figure 7. (a) Impact function for the number of damaged cars, calibrated with data from 2017-2021. The solid green line indicates the 10 mm moving average for the empirical percent of assets affected (PAA) and the blue shading its 5-95% bootstrap confidence interval. In black, a sigmoidal fit (Emanuel, 2011) is shown, and pink bars indicate the fraction of damage reports per MESHS intensity. The grey shaded area indicates MESHS values without sufficient data points for a meaningful empirical fit. **(b)** same as (a), but for damage sums in CHF, with the empirical mean damage ratio (MDR) in blue.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

5 Model evaluation

The calibrated impact functions from the previous section allow estimating hail impacts to buildings and cars for each day with available hazard and exposure data. Here, we evaluate the model performance of deterministic hail impact estimates based on daily MESHS footprints from 2002-2021 and the available building exposure data. Both the number of damaged buildings and the total damage are compared to reported hail damage claims of each day. Results are evaluated separately for the periods 2002-2012 (MESHS only) and 2013-2021 (all hazard variables).

The model is primarily evaluated based on its ability to correctly estimate the total number and cost of hail damages over the four considered cantons. As the vulnerability of buildings differs and hail is a small-scale difficult-to-observe hazard (Martius et al., 2018), co-occurring with strong winds and precipitation which additionally influence damages, a large spread in hail damage costs is expected. To compare model performance between different time periods, hazard variables, exposure data, impact metric, and number of events, we split per-event hail damages estimates into three categories. Hits are all events with modelled damages within one order of magnitude (OOM) of the reported values (green in Fig. 8), false alarms are events where the modelled damages are more than one OOM higher than reported damages (red in Fig. 8), and misses are days with modelled damages of at least one OOM lower than reported damages (orange in Fig. 8). As by default all events are weighted equally, but we are interested in events with considerable hail damage, we calculate metrics considering a minimum damage threshold of 100 affected buildings or 100'000 CHF (modelled or observed) to evaluate model skill. The chosen thresholds ensure that all medium- to high-impact events are included, with the non-considered events (grey shaded in Fig. 8) constituting <1% of all damages and <5% of all claims. From the fraction of the hits, false alarms, and misses we derive the three model evaluation metrics: model-probability of detection (POD_M) as one minus the fraction of misses, given observed damages above the threshold, the model-false alarm ratio (FAR_M) as the fraction of false alarms given modelled damages above the threshold, and the fraction of hits (FH_M) as the number of hits divided by all considered events (red, green, and orange in Fig. 8). As model skill is only weakly dependent on the chosen impact function type (Table C1), the results here focus only on the sigmoidal impact function, and not the smooth-empirical fit.

Table 2 shows skill scores for hail impact model versions with different hazard and exposure data. Compared to other hazard variables, MESHS performs best for both the number of damaged buildings and the total damage with a FH_M of 91 and 77%, respectively. When also considering data before 2013, the POD_M remains almost constant, but significantly more false alarms reduce the model skill. Many of these false alarms are days with multiple few-km-wide patches of positive MESHS values but no recorded damages, which occur less often with the newest generation radar network due to the improved processing (Sect. 2.1). As evident from Fig. 9, both the POD_M and FH_M increase with increasing thresholds for minimum considered damage, while the FAR_M remains stable. For example, if one is only interested in high-impact events (CHF 1 million and larger), model skill increases from POD_M 80 to 94% and from FH_M 77 to 83% for MESHS-based building damages. Focusing on observed extreme hail days from 2013–2021 (2002–2021), the largest 17 (33) events which make up 92% (91%) of all hail damages are modelled with the correct OOM. In comparison, the threshold for the minimum number of claims has a lower impact on the skill scores, as visible in Fig. 9a.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

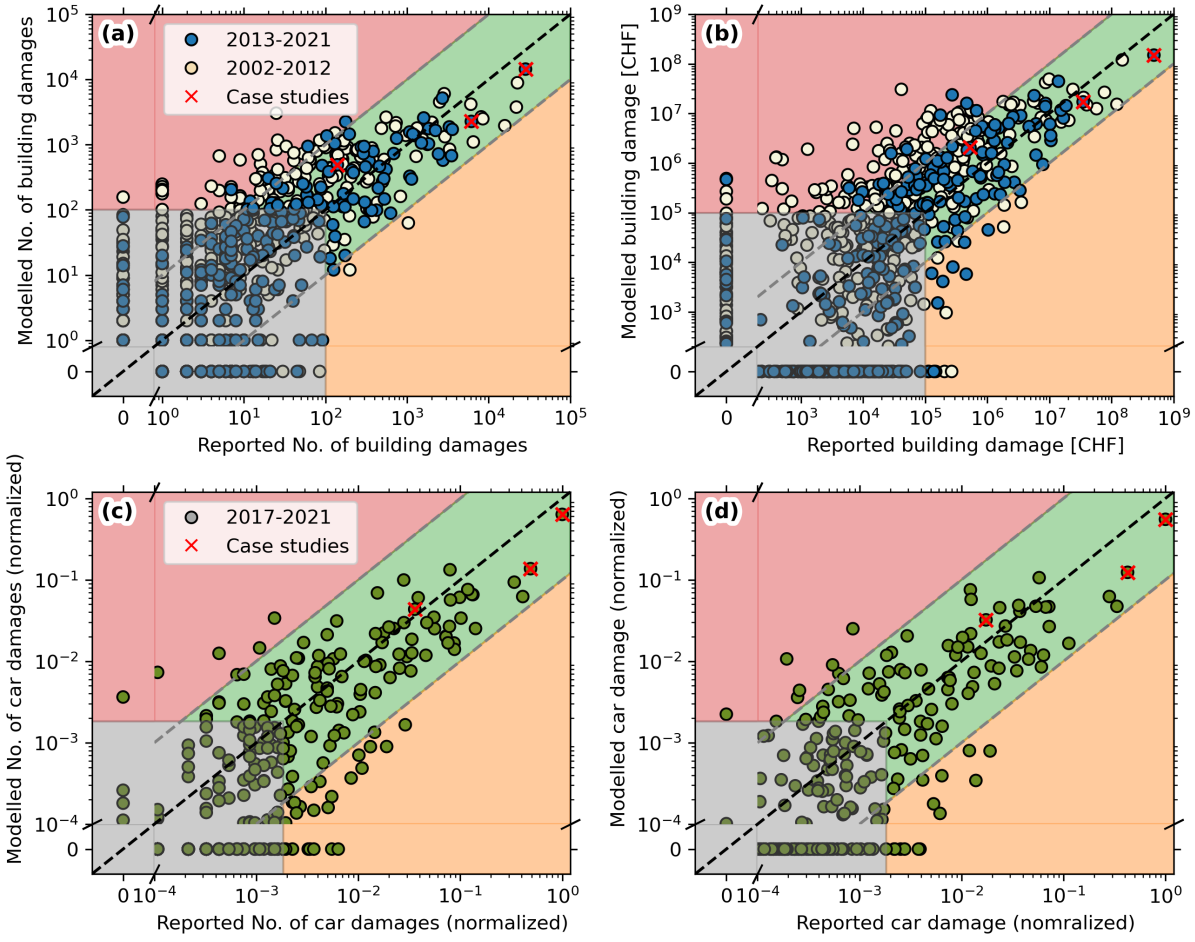

Figure 8. **(a)** Reported vs. modelled number of damaged buildings per hail event (1 day) for the years 2002-2012 (beige) and 2013-2021 (blue). Grey shaded are events with <100 affected buildings which are excluded for the calculation of the skill metrics. Background colours indicate hits (green), false alarms (red), and misses (orange). Case studies from Fig. 10 are marked with a red x. **(b)** Same as (a), but for total damages per event, with events <100'000 CHF damages excluded from the skill metrics calculation. **(c)** Reported vs. modelled number of car damages per event for the years 2017-2021 (green), normalized relative to the maximum impact event. Grey shaded are excluded events, with the threshold chosen to include the same percentage of the total claims as for buildings (99%). **(d)** Same as (c), for total car damages per event, with the threshold chosen to include the same percentage of total damages as for buildings (95%).

The other hazard variables E_{kin}, reflectivity, and VIL have a similar POD_M compared to MESHS, but a higher FAR_M. Namely, there are more events where high damages would be expected from the hazard variable but are not observed, which also leads to an increased number of events which exceed the chosen thresholds of 100 damage claims or 100'000 CHF. Due to the significantly lower skill of other hazard variables, model skills for different exposure layers are only shown for models with MESHS as hazard. Considering scaled building exposure which approximates the value of the exposed building exterior

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

(Sect. 2.2) does not change the model performance for aggregated damages substantially (Table 2). It does reduce overestimation of damages on individual expensive buildings, but as our model aims to estimate aggregated rather than building-level damages, there is little added value in using the scaled exposure.

To compare the model skill to car damages, appropriate thresholds for the minimum considered number of claims and damages must be selected. As the absolute exposure value differs, we choose thresholds that correspond to the same percentage of included damages. The chosen 100 buildings correspond to 95% of all damage claims included, and 100'000 CHF correspond to 99% of the total damage sum included. We choose the equivalent percentage for cars, but only normalized values are shown in Fig. 8c,d due to the conditions of the data provider.

Figure 8 shows that the car damage model consistently predicts the correct order of magnitude for the 25 most impactful events, which contain over 80% of the total damage. However, skill metrics are lower than for buildings (Table 2) with a POD_M and FH_M of 74% for the number of damaged cars and 60% for total damages. The lower skill is explained by a high number of missed events with zero modelled damages but observed damages above the chosen threshold, which relates to two main factors. Firstly, the assumed random location increases uncertainties particularly in the damage data pre-processing, which can lead to missed events. Secondly, the shorter considered time period contains fewer extreme events and, thus, a certain percentage of included total damages contains more small-scale events. Furthermore, specifically for car damages the time of day is relevant due to increased traffic during rush hour, which is not considered in this study but may further improve damage estimates. As with buildings, there are events where missed damages in one location are compensated by over-predicted damages in another location where the hazard variable wrongly predicts intense hail (see next section). This compensation mechanism is more effective for larger spatial scales, which benefits skill scores for car damages that cover all Swiss cantons.

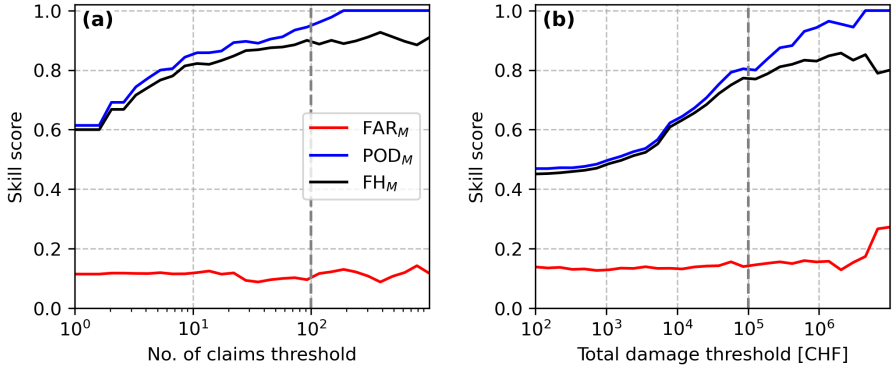

Figure 9. Dependence of the skill metrics POD_M, FAR_M, FH_M to the chosen threshold for the minimum **(a)** number of claims for buildings and **(b)** minimum building damage per event. The threshold used for the calculations in Table 2 is marked with a grey dotted line.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

Table 2. Skill scores for calculating hail damages to buildings (cars) with different variable combinations for the period 2013-2021 (2017-2021). For rows where data is available, skill scores for the time period 2002-2021 are shown in parentheses. The first half (a) refers to predictions for the number of damage claims where POD_M, FAR_M, FH_M refer to events with >100 building damage claims (for car threshold refer to text) and the second half (b) to predictions of the damage sums in CHF where POD_M, FAR_M, FH_M refer to events with >100'000 CHF building damages (for car threshold refer to text). The column "No. of events" refers to the number of events that exceed the damage threshold and are considered for the skill metrics.

Exposure	**Hazard**	**FAR_M**	**POD_M**	**FH_M**	**No. of events**
(a) No. of damage claims					
buildings	E_{kin}	33%	87%	64%	112
buildings	VIL	43%	94%	58%	125
buildings	dBZ	36%	95%	63%	120
buildings	MESHS	9% (27%)	94% (94%)	91% (73%)	78 (212)
buildings (scaled)	MESHS	11% (30%)	96% (95%)	90% (71%)	81 (232)
cars	MESHS	7%	74%	74%	121
(b) Damages					
buildings	E_{kin}	65%	78%	34%	284
buildings	VIL	57%	78%	41%	229
buildings	dBZ	65%	71%	34%	241
buildings	MESHS	14% (37%)	80% (74%)	77% (60%)	116 (320)
buildings (scaled)	MESHS	18% (39%)	81% (75%)	75% (58%)	122 (331)
cars	MESHS	8%	60%	60%	233

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

5.1 Limitations of the MESHS-based model and potential of crowdsourced data

As this section addresses limitations regarding the spatial structure of MESHS-based hail damage estimates, we focus only on building damage data where exact spatial coordinates are available. Due to the high POD of MESHS, almost all observed hail damages are captured by the model. The model represents the overall spatial structure of hail damages well for both multiple isolated hail cells and large hail streaks from one thunderstorm (Fig. 10). However, spatial patterns of actual hail damages tend to be narrower and locally more intense than modelled damages and there are areas with MESHS signal but no reported damages despite exposed buildings (e.g. southwest Berne on 26 June 2020, or northern Zurich on 28 June 2021; Fig. 10). This underlines the limitations from currently available radar-based hail observations, which is also reflected in the high FAR in Fig. 4.

For particularly intense events, such as 28 June 2021 (see also Kopp et al., 2022), local extreme damages from roof-penetrating hail are not well represented by the MESHS-based model. In fact, they are partly compensated by a larger spatial extent, leading to a total damage estimate with the correct order of magnitude. However, this compensation is less effective on smaller scale (e.g. Canton) leading to a reduced model skill. In the case of 28 June 2021, underestimated damages in Lucerne are compensated by overestimated damages in Zurich due to high relative exposure values and overestimated MESHS. For damages on even smaller scales down to individual buildings, the here developed model is not fit-for-purpose. Nevertheless, it may be used as input for statistical post-processing to model damages on a building level (Miralles et al., 2023).

Our analysis based on long-term and extensive damage data suggests that the spatial extent of damaging hail streaks on the ground are actually narrower than the radar-derived MESHS footprints (Fig. A1). Consequently, long-term averages of MESHS-based damage estimates are spatially smoother and depend more on the distribution of exposure values (Fig. 2) than long-term averages of actual damages (Fig. 11). While some densely populated areas (e.g. the cities of Biel, Thun, Lucerne and the greater Zurich area) are visible as spots with high total hail damages in both modelled and observed long-term damage patterns, others (e.g. Berne and Winterthur) are only clearly visible in modelled damages. This indicates that multiple hailstorms may have either passed nearby with MESHS extending over the city, while the actual hail streaks missed densely populated areas, or thunderstorms with high MESHS but no damaging hail were observed over the city. The intense hail streak on 28 June 2021 in the canton of Lucerne is evident in the observed 20-year damage record (Fig. 11a). It was characterized by extreme total hail damages in a rural area where only relatively few buildings are located (Fig. 2a). Since MESHS-based damage estimates suffer from potentially over-estimating the spatial extent of a hail streak and consequently have lower per-area damage estimates, individual extreme events such 28 June 2021, are not directly visible in the long-term modelled hail damages (Fig. 11b).

While from a nowcasting-perspective, a spatial over-estimation of the potential hail hazard may be intended, this is not practical for damage modelling. Having learned about the shortcomings of using radar-based hail intensity estimates, such as MESHS, as a hazard variable for damage modelling, it is worth exploring alternatives. One possibility to improve the hail hazard variable is the usage of crowdsourced reports (Appendix B; Barras et al., 2019). An exploratory hail impact estimate based on crowdsourced data of the extreme hail event on 28 June 2021 in Fig. B2 shows a more spatially accurate hail damage foot-

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

print with more realistic extreme local damages, compared to radar-based estimates. This improvement is caused by narrower footprints and fewer occurrences of high hazard intensities, allowing the crowdsource-based impact function for PAA to reach higher values of up to 40% (compared to 10% for MESHS). Thus, crowdsource-based hail damage estimates have a reduced low-intensity large-area bias. However, drawbacks include inconsistent reporting behaviour (particularly "joke" reports of very large hail), lower accuracy at night, and short data records (2017-2021). Current pre-processing still cannot filter out all false reports of very large hail (Barras et al., 2019), which is crucial for accurate damage modelling. Due to these false alarms and numerous missed events (e.g. nocturnal hailstorms), the model skill of crowdsource-based hail impact estimates over all events (not shown) remains much lower than for radar-based impact estimates. Nevertheless, there is large potential for hail damage modelling based on crowdsourced data, especially because areas with high exposure (buildings or cars) mostly also have high population density and many crowdsourced reports. To fully exploit this potential, dedicated pre-processing of hail damage reports needs to be developed, which is a promising direction for future research.

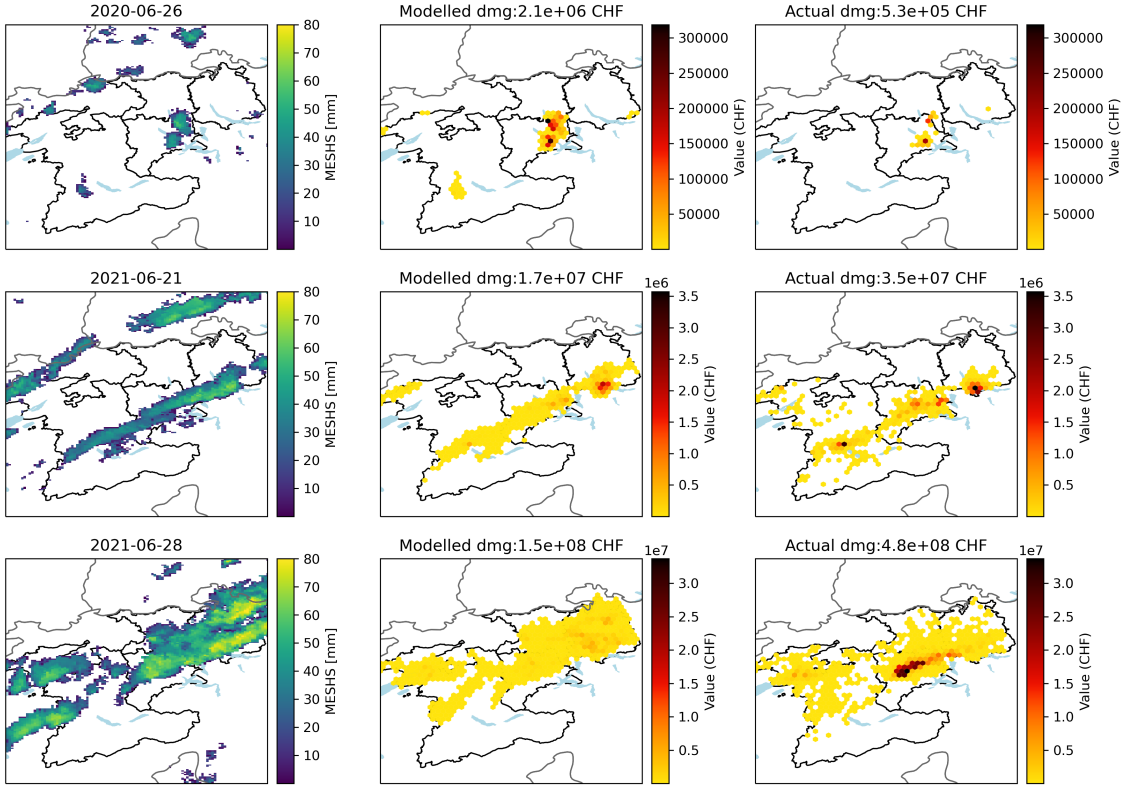

Figure 10. Example of model performance for three hail events on 26 June 2020, 21 June 2021, and 28 June 2021. Shown are (left) MESHS, (middle) modelled, and (right) reported building damages. Note the different damage value scales for each event.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

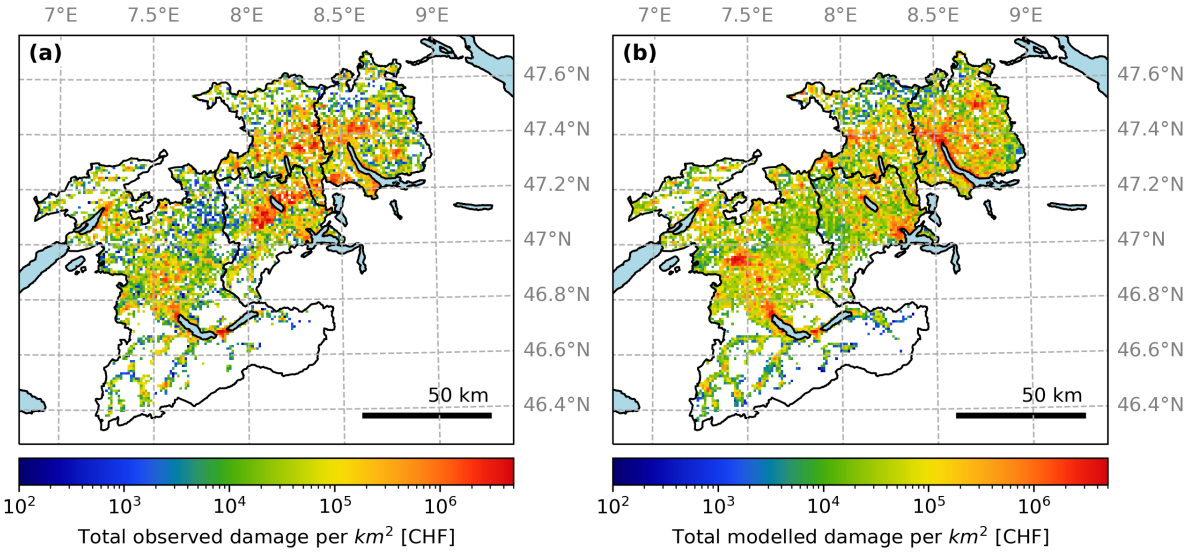

Figure 11. Long-term observed and modelled hail damages to buildings: **(a)** Total damage over 20 years per $1\,\mathrm{km}^2$ gridcell. **(b)** Total modelled hail damage over 20 years, based on MESHS data.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

6 Discussion

Radar verification with insurance claims. The available geolocated hail damage data reveal coherent hail streaks with damaging hail where a maximum hail size of >2cm can be expected (Stucki and Egli, 2007). In combination with complete exposure data, the insurance claims constitute a particularly useful verification metric in contrast to individual reports (e.g. ESWD; Púčik et al., 2017), which do not allow a quantification of false alarms (Delobbe and Holleman, 2006). While many studies have used hail damage data to verify radar-based hail detection (Hohl et al., 2002b; Holleman et al., 2000; Kunz and Puskeiler, 2010; Skripniková and Řezáčová, 2014; Nisi et al., 2016), only few have used building-scale data (Schuster et al., 2006; Kunz and Kugel, 2015; Brown et al., 2015; Warren et al., 2020) and to our knowledge, with 250'000 claims, we base our work on the largest number of per-building hail damage reports. A limitation of insurance claim data is their inaccurate temporal reporting (e.g. Warren et al., 2020), which requires a plausibility check (Sect. 2.2.1). Thus, the data is not fully independent of radar-based hazard variables, but the rare occurrence and small spatial extent of hail streaks allow for a pre-processing approach with high confidence in the corrected building damage claim dates.

Reflection on the model calibration approach. Compared to earlier radar-based hail damage modelling in Switzerland (Hohl et al., 2002a, b), we calibrate our model with spatio-temporally complete data rather than individual storms. Using all data leads to more robust results and direct usability with operational radar data as input. However, we acknowledge that attenuation can lead to a bias in individual hail cells, but we expect this bias to be negligible compared to the inherent uncertainty of (large) hail detection from radar data (e.g. Blair et al., 2011).

The chosen spatially explicit calibration approach reveals the underlying long-tailed distribution of actual hail damages (Fig. 5), given a certain hazard intensity. Sampling from this distribution instead of directly using the mean PAA or MDR to model hail damages was considered, but finally not implemented. As visible in the shown examples (Fig. 4, 10, and A1), the extreme PAA and MDR values are typically concentrated in one hail streak which is a subset of the MESHS footprint. Its location within the MESHS footprint often does not correspond to the highest hazard intensity derived from radar data. Thus, while a random sampling may improve per-gridcell statistics, the modelled hail damage footprints would not be improved towards the observed hail streaks, and aggregated damages would remain unaffected.

Horizontal hail drift. Some studies explicitly consider horizontal drift and shift radar footprints to obtain an optimal overlap to insurance claims, which improves correlations (Hohl et al., 2002a; Schuster et al., 2006). Here, we consider this drift in the insurance claims pre-processing, and by providing additional verification metrics with a 4 km buffer. For the model calibration and evaluation, however, no spatial shift is considered for two main reasons. First, verification results of the primary hazard variable MESHS show that most insurance claims (88%) are within a MESHS footprint without considering spatial shift. Our findings suggest that MESHS generally spatially overpredicts hail streaks on the ground, rather than suffering from wind-related directed shifts. Secondly, an optimal shift (as in Hohl et al., 2002a; Schuster et al., 2006) could not be applied in a real-time application, before damage claim data is available. For E_{kin} the improvement with a spatial shift may be larger due to the smaller scale of high E_{kin} values, but is not expected to rival MESHS in terms of model skill due to much lower skill metrics (Table 2).

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

Model evaluation and applicability. The developed model focuses on representing aggregated hail damages to buildings and cars, based on the exposed number of assets and their value. While different types of used building materials, building age, and spacing between buildings strongly influence the vulnerability of an individual building (Hohl et al., 2002a; Stucki and Egli, 2007; Schmidberger, 2018), the developed impact function represents a mean vulnerability which yields realistic estimates on an aggregated spatial scale. Exploratory results show that e.g. newer buildings tend to be more vulnerable with an average PAA of 9% at MESHS of 60 mm for buildings built after 2002 vs. 5% for all buildings (Fig. 6). However, the per-event model skill remains largely unaffected when using separate impact functions for different classes of construction year. Similarly, the consideration of a scaled building value that approximates the value of exposed building parts does not improve model skill (Table 2).

This lack of improvement when considering more details of the exposure data highlights that currently the overall large uncertainty in the hail damage model mainly stems from the available single-polarization radar metrics. Since none of the considered radar variables accurately and consistently distinguishes large hail from smaller (not damaging) hail, damage estimates have high uncertainty. While uncertainties of over one order of magnitude are common in natural hazard risk modelling due to the complexity of modelled processes (e.g. Röösli et al., 2021; Eberenz et al., 2021; Lüthi et al., 2021), it has important implications for the model applicability.

The presented MESHS-based hail damage model is suitable to provide approximate estimates for the number of claims and total cost of building and car damages for both an exposure layer including all asset values (e.g. derived through LitPop; Eberenz et al., 2020) or a user-provided portfolio. Note that uncertainties increase with a smaller spatial extent of the exposure portfolio (Sect. 5.1). Assuming the ratio between building/car values and repair costs is similar between countries, the impact functions derived in this study can be used for other regions. A scaling of the impact function to reflect different building materials could be explored, but our results suggest that the uncertainty in MESHS exceeds the differences in exposure properties, such as building volume and year of construction, when analyzing aggregated hail damage estimates. The areas with expected hail damages are well covered, but extreme local impacts are not well represented, since modelled hail damages present with larger spatial extent and lower local impact. Thus, for applications where an accurate representation of intense hail streaks is crucial, statistical post-processing (Miralles et al., 2023) or a more accurate hazard layer is indispensable. A data product for the latter is not yet available in Switzerland, but promising work is currently underway to derive more accurate hail measures, e.g., from a combination of radar and crowdsourced data or from improved hail detection algorithms based on dual-polarization radar signatures.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

7 Conclusions

Severe hail causes large building damages in the selected four cantons of Switzerland with 1.39 billion CHF in the last 20 years, averaging 1'400 CHF per building and including 1187 claims above 100'000 CHF. 90% (1.25 billion CHF) of the total damage to buildings occurred during the strongest 30 hail events, with a single event on 28 June 2021 causing 35% (483 million CHF). Hail damages are mostly concentrated in 3–10 km wide hail streaks with consistently reported damage claims in all gridcells with exposed buildings.

The maximum expected severe hail size (MESHS) is the best performing radar-based hail intensity measure, with a high probability of detection of 60% on a gridcell-level and 88% for individual damage claims. However, a significant false alarm ratio of over 50% even for extreme MESHS values indicates that high MESHS often occur without damaging hail, confirming that MESHS should not directly be interpreted as local hailstone size, but as "maximum expected" size, given the storm intensity.

Using a novel empirical calibration approach, which informs about the shape of the impact function and provides detailed uncertainty estimates, we calibrate impact functions for buildings and cars depending on the available single-polarization radar metrics: MESHS, reflectivity, hail kinetic energy, and vertically integrated liquid. Spatially explicit hail damages are estimated for each day from April–September 2013–2021 by combining the impact function with exposure and hazard data. The MESHS-based model successfully estimates the correct order of magnitude for most daily hail damages across four cantons in Switzerland, which allows deriving first-guess estimates of expected damages that can serve as valuable information for, e.g., insurances in taking immediate action after large hail events. In particular, the largest observed 26 car damage events and the largest 33 building damage events are modelled with the correct order of magnitude. However, considerable uncertainty remains, especially in the spatial structure. The model has a bias towards large extent and lower local impact of hail damages, reflecting the fact that MESHS (or any other available radar variable) does not reliably distinguish medium-sized hail from extreme hail which causes roof-penetrating hail damages. The uncertainty in hail damage predictions with the proposed radar-based model remain large, warranting careful assessment of its intended use. While the model is suitable to derive approximate hail damage estimates to buildings and cars in near real-time, accurately quantifying tail risk using a probabilistic event set of MESHS footprints remains challenging given limitations in the radar-based identification of extreme hail.

Crowdsourced data shows promising results regarding the modelling of locally intense hail damages, but longer time series and improved pre-processing are required to outperform radar-based estimates. Furthermore, the development of dual-polarization hail intensity measures could potentially overcome some limitations regarding the spatial structure of intense hail streaks. Such innovations will have instantaneous effect, as new variables can easily be included in our hail damage model using the presented calibration approach. Lastly, the use of hail size estimates from numerical weather predictions as hazard variable would allow for new applications of the model. For example, ensemble weather predictions may be used to provide probabilistic hail impact estimates as a basis for impact-based warnings. Furthermore, application of the model to high-resolution convection resolving climate simulations could quantify changing hail risk in a warming climate.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

Code and data availability. CLIMADA is an open-source and -access software (https://doi.org/10.5281/zenodo.7691855) and can be used with any user-provided portfolio under the General Public Licence gpl-3.0. The code for the conducted analysis and creation of all figures is written in Python 3.9 and can be found in the CLIMADA paper repository under the following link: https://github.com/CLIMADA-project/climada_papers. Radar data is available from MeteoSwiss upon request (https://www.meteoschweiz.admin.ch/service-und-publikationen/service.html), with a licencing requirement for commercial use. Hail damage data for both buildings and cars from insurance companies are only available within the scClim project (https://scclim.ethz.ch/).

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

Appendix A: Building damages

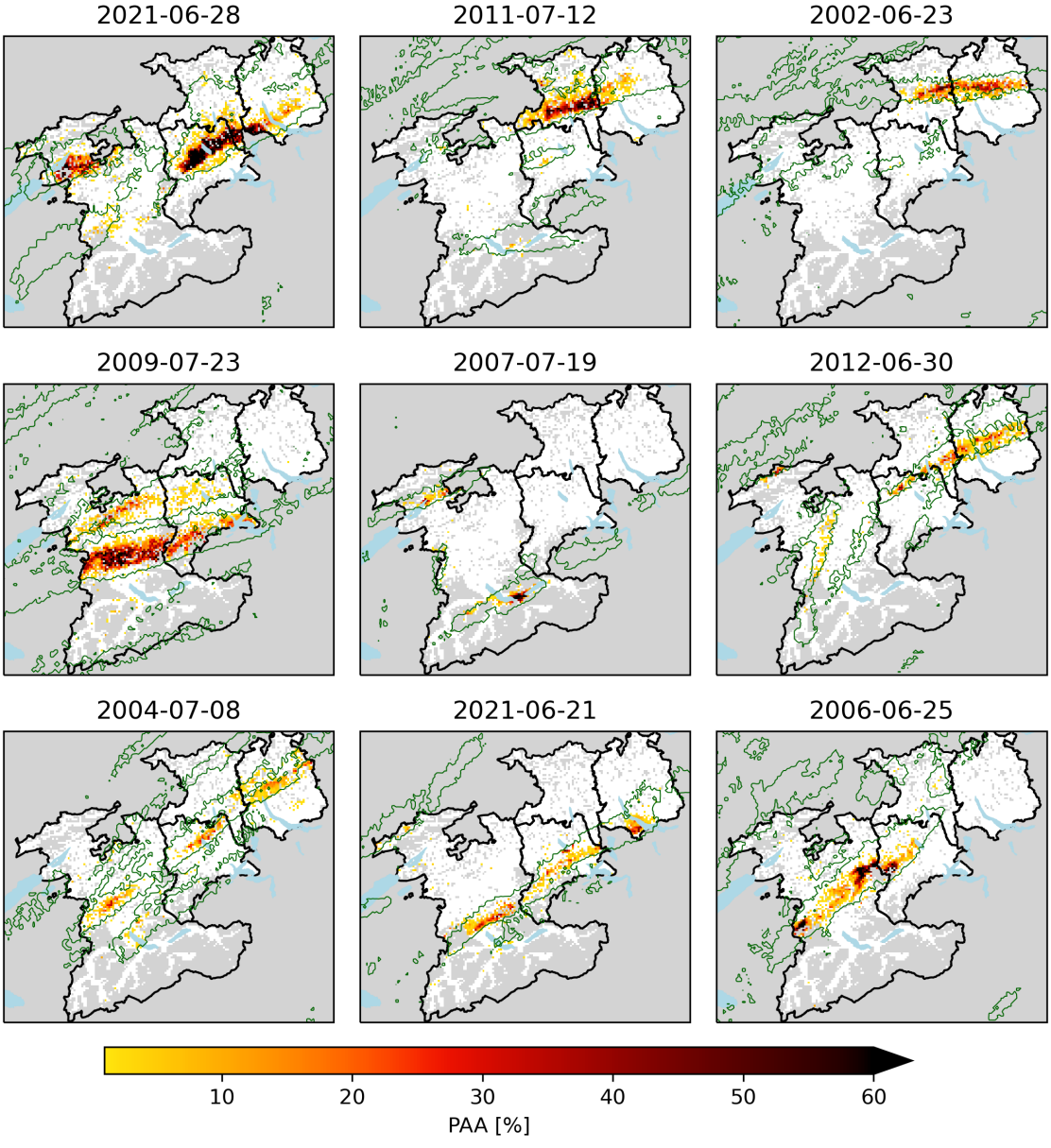

Figure A1. Percentage of damaged buildings for the 9 strongest hail events between 2002 and 2021. 1 km gridcells with less than 10 buildings are marked in grey and green lines outline the 20 mm MESHS contour.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023


Natural Hazards and Earth System Sciences
Discussions

Appendix B: Crowdsourced data

MeteoSwiss launched a crowdsourcing function in 2015 via the MeteoSwiss app with 500'000 active daily users. Since 2017, users can select the following hail size categories in their report (Barras et al., 2019):

- Smaller than coffee bean: >5 mm
- Coffee bean: 8mm (5-15 mm)
- One Swiss franc coin: 23 mm (15-27 mm)
- Five Swiss francs coin: 32 mm (27-37 mm)
- Golf ball: 43 mm (37-55 mm)
- Tennis ball: 68 mm (>55 mm)

Multiple plausibility filters are applied to remove implausible reports. Firstly, reports outside a conservative threshold of 35 dBZ radar reflectivity within a buffer zone of 4 km are removed. Furthermore, entries from user ID's with an unusual reporting pattern are removed (Barras et al., 2019).

From the data we created a gridded dataset, with the following approach: For each 1 km grid point, the average size and total number of crowdsourcing reports are computed. Gridcells with a population of >2000 people and no crowdsourcing reports are assumed to have no hail (0 mm) because observations show that within known hail footprints, gridcells with >2000 people consistently contain hail reports. For grid points with a population of less than 2000 people and no hail report, the hail size is estimated by considering the average of all cells within 3.5 km distance, given that this area contains at least three reports. Lastly, a noise filter is applied which estimates the hail size at every gridcell as an average over the surrounding 3×3 km box (i.e. 9 gridcells in total). As the largest hail size report is "Tennis ball" with 68 mm, higher values cannot be reached.

Figure B1 shows all hail reports of 28 June 2021 filtered as described in Barras et al. (2019). A comparison with areas of high reported hail damages shows that only within areas of high PAA "Tennis ball" hail stones are consistently reported, although scattered "Tennis ball" reports in other regions also exist. These may be individual large hail stones that did not hit any building or, more likely, reports where a user overestimated the hail stone size. Note that the category "Tennis ball" was originally introduced to filter our "joke" reports of users who choose the largest possible category for fun (Barras et al., 2019). Applying the calibration approach described in Sect. 4 to the gridded crowdsource dataset, corresponding empirical impact functions are derived. Figure B2 shows the resulting modelled hail damages for 28 June 2021, which represent the observed locally intense hail damages well.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

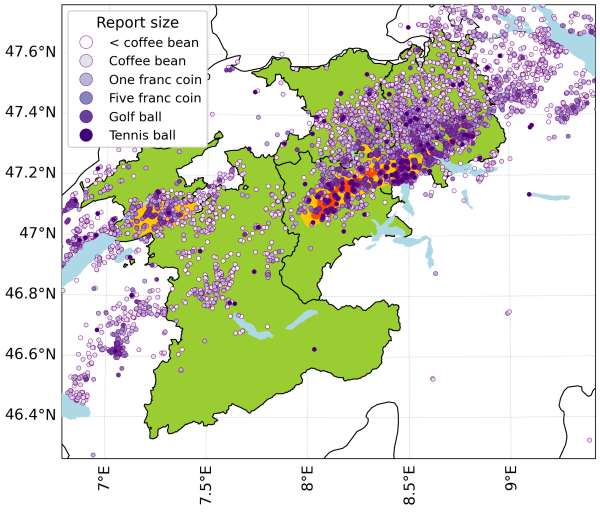

Figure B1. Filtered crowdsourced reports from the extreme hail event on 28 June 2021, with shading marking areas of PAA>10%(orange) and >50% (red). Cantons with available building damage data are marked in green.

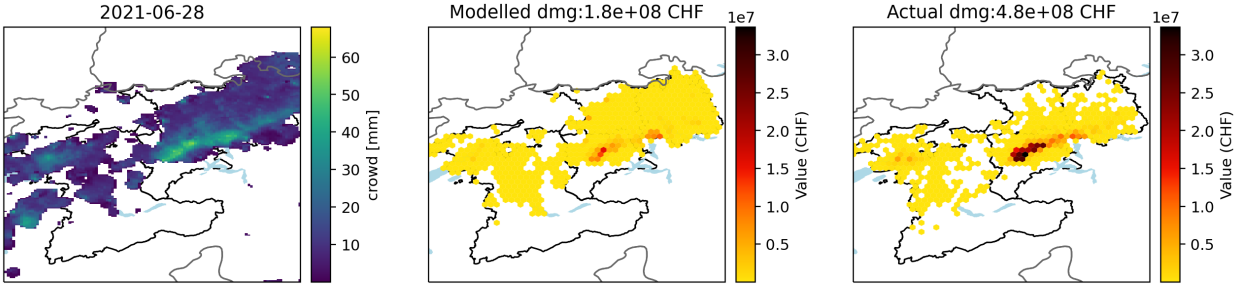

Figure B2. Crowdsource-based hail size estimate derived with the gridding approach described in Appendix B and corresponding modelled and observed damages for 28 June 2021. Modelled damages are calculated with an impact function derived as shown in Sect. 4 and crowdsourced data as hazard variable.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

Appendix C: Model evaluation

Table C1. Skill scores for calculating hail damages to buildings (cars) with different variable combinations for the period 2013-2021 (2017-2021). All values are calculated with the sigmoidal impact function and, as comparison, values calculated with a smooth-empirical fit are shown in parentheses. The first half (a) refers to predictions for the number of damage claims where POD_M, FAR_M, FH_M refer to events with >100 building damage claims (for car threshold refer to text) and the second half (b) to predictions of the damage sums in CHF where POD_M, FAR_M, FH_M refer to events with >100'000 CHF building damages (for car threshold refer to text). The column "No. of events" refers to the number of events that exceed the damage threshold and are considered for the skill metrics.

Exposure	**Hazard**	**FAR_M**	**POD_M**	**FH_M**	**No. of events**
(a) No. of damage claims					
buildings	E_{kin}	33% (53%)	87% (89%)	64% (46%)	112 (163)
buildings	VIL	43% (43%)	94% (94%)	58% (58%)	125 (125)
buildings	dBZ	36% (43%)	95% (96%)	63% (57%)	120 (134)
buildings	MESHS	9% (9%)	94% (96%)	91% (91%)	78 (80)
buildings (scaled)	MESHS	11% (11%)	96% (96%)	90% (90%)	81 (80)
cars	MESHS	7% (7%)	74% (77%)	74% (76%)	121 (122)
(b) Damages					
buildings	E_{kin}	65% (65%)	78% (74%)	34% (34%)	284 (263)
buildings	VIL	57% (61%)	78% (81%)	41% (39%)	229 (261)
buildings	dBZ	65% (67%)	71% (70%)	34% (32%)	241 (239)
buildings	MESHS	14% (21%)	80% (88%)	77% (77%)	116 (132)
buildings (scaled)	MESHS	18% (24%)	81% (88%)	75% (74%)	122 (138)
cars	MESHS	8% (8%)	60% (66%)	60% (66%)	233 (237)

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

Appendix D: Impact function parameters

Table D1. Parameters of best-fit sigmoidal impact function for building and car damages as described in Sect 4.

Exposure	**Hazard**	$\mathbf{V}_{thresh}$	$\mathbf{V}_{half}$	**scale**
(a) No. of damage claims (PAA)				
buildings	MESHS [mm]	8.6	66.9	0.118
buildings	Max. reflectivity [dBZ]	53.3	74.5	0.318
buildings	E_{kin} [J m^{-2}]	0	465	0.037
buildings	VIL [g m^{-2}]	7.2	42.5	0.072
cars	MESHS [mm]	8.1	55.0	0.046
(b) Damages (MDR)				
buildings	MESHS [mm]	20	73.9	1.47e-3
buildings	Max. reflectivity [dBZ]	52.1	74.5	1.80e-3
buildings	E_{kin} [J m^{-2}]	0	480	2.50e-4
buildings	VIL [g m^{-2}]	15.8	29.3	1.80e-3
cars	MESHS [mm]	13.7	53.3	3.96e-3

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Author contributions. **TS:** Conceptualization, data curation, methodology, software, visualization, writing - original draft, writing - review & editing. **RP:** Methodology, software, writing - review & editing. **LV:** Supervision, writing - review & editing, **KS:** Supervision, writing - review & editing, **DB:** Conceptualization, funding acquisition, supervision, writing - review & editing

Competing interests. The authors declare no competing interests.

Acknowledgements. We thank Alessandro Hering, Cornelia Schwierz, and Urs Germann from MeteoSwiss for useful inputs regarding the meteorological data. Further, we are greatful to the scClim project partners from the insurance industry for the provided data and explanations thereof. Specifically, we thank Mirco Heidemann, Jan Wüthrich, and Daniel Steinfeld from GVZ, Markus Wigger from GVL, Daniel Wey and Manuel Vonarb from AGV, and Hannes Suter from GVB. Finally, we thank all scClim (https://scclim.ethz.ch/) researchers and Jérôme Kopp for valuable inputs throughout the project. This study was funded by the Swiss National Science Foundation (SNSF) Sinergia grant CRSII5_201792.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

References

Allen, J. T., Giammanco, I. M., Kumjian, M. R., Jurgen Punge, H., Zhang, Q., Groenemeijer, P., Kunz, M., and Ortega, K.: Understanding Hail in the Earth System, Reviews of Geophysics, 58, e2019RG000 665, https://doi.org/10.1029/2019RG000665, 2020.

Amburn, S. A. and Wolf, P. L.: VIL Density as a Hail Indicator, Weather and Forecasting, 12, 473–478, https://doi.org/10.1175/1520-0434(1997)012<0473:VDAAHI>2.0.CO;2, 1997.

Atlas, D., Harper, W. G., Ludlam, F. H., and MacKlin, W. C.: Radar scatter by large hail, Quarterly Journal of the Royal Meteorological Society, 86, 468–482, https://doi.org/10.1002/qj.49708637004, 1960.

Auer, A. H.: Hail Recognition through the Combined Use of Radar Reflectivity and Cloud-Top Temperatures, Monthly Weather Review, 122, 2218–2221, https://doi.org/10.1175/1520-0493(1994)122<2218:HRTTCU>2.0.CO;2, 1994.

Aznar-Siguan, G. and Bresch, D. N.: CLIMADA v1: a global weather and climate risk assessment platform, Geoscientific Model Development, 12, 3085–3097, https://doi.org/10.5194/gmd-12-3085-2019, 2019.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T.: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities, Monthly Weather Review, 139, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1, 2011.

Barras, H., Hering, A., Martynov, A., Noti, P.-A., Germann, U., and Martius, O.: Experiences with >50,000 Crowdsourced Hail Reports in Switzerland, Bulletin of the American Meteorological Society, 100, 1429–1440, https://doi.org/10.1175/BAMS-D-18-0090.1, 2019.

Besic, N., Figueras i Ventura, J., Grazioli, J., Gabella, M., Germann, U., and Berne, A.: Hydrometeor classification through statistical clustering of polarimetric radar measurements: a semi-supervised approach, Atmospheric Measurement Techniques, 9, 4425–4445, https://doi.org/10.5194/amt-9-4425-2016, 2016.

Besic, N., Gehring, J., Praz, C., Figueras i Ventura, J., Grazioli, J., Gabella, M., Germann, U., and Berne, A.: Unraveling hydrometeor mixtures in polarimetric radar measurements, Atmospheric Measurement Techniques, 11, 4847–4866, https://doi.org/10.5194/amt-11-4847-2018, 2018.

Betschart, M. and Hering, A.: Automatic Hail Detection at MeteoSwiss, Tech. rep., MeteoSwiss, 2012.

Biedermann, F.: Mobilitätsverhalten der Bevölkerung, 24165261, Bundesamt für Statistik (BFS), Neuchâtel, https://dam-api.bfs.admin.ch/hub/api/dam/assets/24165261/master, 2023.

Blair, S. F., Deroche, D. R., Boustead, J. M., Leighton, J. W., Barjenbruch, B. L., and Gargan, W. P.: A Radar-Based Assessment of the Detectability of Giant Hail, E-Journal of Severe Storms Meteorology, 6, 1–30, https://doi.org/10.55599/ejssm.v6i7.34, number: 7, 2011.

Bohren, C. F. and Battan, L. J.: Radar Backscattering of Microwaves by Spongy Ice Spheres, Journal of the Atmospheric Sciences, 39, 2623–2628, https://doi.org/10.1175/1520-0469(1982)039<2623:RBOMBS>2.0.CO;2, 1982.

Brown, T. M., Pogorzelski, W. H., and Giammanco, I. M.: Evaluating Hail Damage Using Property Insurance Claims Data, Weather, Climate, and Society, 7, 197–210, https://doi.org/10.1175/WCAS-D-15-0011.1, 2015.

Cecchini, M. A., Heymsfield, A. J., Honeyager, R., Field, P., Machado, L. A. T., and Dias, M. A. F. D. S.: Revisiting the Hail Radar Reflectivity–Kinetic Energy Flux Relation by Combining T-Matrix and Discrete Dipole Approximation Calculations to Size Distribution Observations, Journal of the Atmospheric Sciences, 79, 1927–1940, https://doi.org/10.1175/JAS-D-20-0373.1, 2022.

Delobbe, L. and Holleman, I.: Uncertainties in radar echo top heights used for hail detection, Meteorological Applications, 13, 361–374, https://doi.org/10.1017/S1350482706002374, 2006.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

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

Eberenz, S., Stocker, D., Röösli, T., and Bresch, D. N.: Asset exposure data for global physical risk assessment, Earth System Science Data, 12, 817–833, https://doi.org/10.5194/essd-12-817-2020, 2020.

Eberenz, S., Lüthi, S., and Bresch, D. N.: Regional tropical cyclone impact functions for globally consistent risk assessments, Natural Hazards and Earth System Sciences, 21, 393–415, https://doi.org/10.5194/nhess-21-393-2021, 2021.

Emanuel, K.: Global Warming Effects on U.S. Hurricane Damage, Weather, Climate, and Society, 3, 261–268, https://doi.org/10.1175/WCAS-D-11-00007.1, 2011.

Foote, B., Makitov, V., and Krauss, T.: Hail metrics using conventional radar., in: Proceedings of the 16th Conference on Planned and Inadvertent Weather Modification, San Diego, CA USA, 2005.

Gao, F. and Han, L.: Implementing the Nelder-Mead simplex algorithm with adaptive parameters, Computational Optimization and Applications, 51, 259–277, https://doi.org/10.1007/s10589-010-9329-3, 2012.

Germann, U., Marco, B., Marco, G., and Maurizio, S.: Peak Performance: Radar design for prediction in the Swiss Alps, Meteorological Technology International, pp. 42–45, 2015.

Greene, D. R. and Clark, R. A.: Vertically Integrated Liquid Water—A New Analysis Tool, Monthly Weather Review, 100, 548–552, https://doi.org/10.1175/1520-0493(1972)100<0548:VILWNA>2.3.CO;2, 1972.

Grieser, J. and Hill, M.: How to Express Hail Intensity—Modeling the Hailstone Size Distribution, Journal of Applied Meteorology and Climatology, 58, 2329–2345, https://doi.org/10.1175/JAMC-D-18-0334.1, 2019.

GVL: Hagelereignis überschattet das Geschäftsjahr 2021, media release from 05 May 2022, https://www.gvl.ch/unternehmen/medien/, 2022.

Haklay, M. and Weber, P.: OpenStreetMap: User-Generated Street Maps, IEEE Pervasive Computing, 7, 12–18, https://doi.org/10.1109/MPRV.2008.80, 2008.

Heymsfield, A. and Wright, R.: Graupel and Hail Terminal Velocities: Does a "Supercritical" Reynolds Number Apply?, Journal of the Atmospheric Sciences, 71, 3392–3403, https://doi.org/10.1175/JAS-D-14-0034.1, 2014.

Hohl, R., Schiesser, H.-H., and Aller, D.: Hailfall: the relationship between radar-derived hail kinetic energy and hail damage to buildings, Atmospheric Research, 63, 177–207, https://doi.org/10.1016/S0169-8095(02)00059-5, 2002a.

Hohl, R., Schiesser, H.-H., and Knepper, I.: The use of weather radars to estimate hail damage to automobiles: an exploratory study in Switzerland, Atmospheric Research, 61, 215–238, https://doi.org/10.1016/S0169-8095(01)00134-X, 2002b.

Holleman, I., Wessels, H., Onvlee, J., and Barlag, S.: Development of a hail-detection-product, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25, 1293–1297, https://doi.org/10.1016/S1464-1909(00)00197-0, 2000.

Joe, P., Burgess, D., Potts, R., Keenan, T., Stumpf, G., and Treloar, A.: The S2K Severe Weather Detection Algorithms and Their Performance, Weather and Forecasting, 19, 43–63, https://doi.org/10.1175/1520-0434(2004)019<0043:TSSWDA>2.0.CO;2, 2004.

Kopp, J., Schröer, K., Schwierz, C., Hering, A., Germann, U., and Martius, O.: The summer 2021 Switzerland hailstorms: weather situation, major impacts and unique observational data, Weather, p. wea.4306, https://doi.org/10.1002/wea.4306, 2022.

Kopp, J., Manzato, A., Hering, A., Germann, U., and Martius, O.: How observations from automatic hail sensors in Switzerland shed light on local hailfall duration and compare with hailpad measurements, Atmospheric Measurement Techniques, 16, 3487–3503, https://doi.org/10.5194/amt-16-3487-2023, 2023.

Kumjian, M. R. and Lombardo, K.: A Hail Growth Trajectory Model for Exploring the Environmental Controls on Hail Size: Model Physics and Idealized Tests, Journal of the Atmospheric Sciences, 77, 2765–2791, https://doi.org/10.1175/JAS-D-20-0016.1, 2020.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Kunz, M. and Kugel, P. I. S.: Detection of hail signatures from single-polarization C-band radar reflectivity, Atmospheric Research, 153, 565–577, https://doi.org/10.1016/j.atmosres.2014.09.010, 2015.

Kunz, M. and Puskeiler, M.: High-resolution assessment of the hail hazard over complex terrain from radar and insurance data, Meteorologische Zeitschrift, 19, 427–439, https://doi.org/10.1127/0941-2948/2010/0452, 2010.

Lüthi, S., Aznar-Siguan, G., Fairless, C., and Bresch, D. N.: Globally consistent assessment of economic impacts of wildfires in CLIMADA v2.2, Geoscientific Model Development, 14, 7175–7187, https://doi.org/10.5194/gmd-14-7175-2021, 2021.

Macdonald, J. R. and Stack, M. M.: Some Thoughts on Modelling Hail Impact on Surfaces, Journal of Bio- and Tribo-Corrosion, 7, 37, https://doi.org/10.1007/s40735-020-00458-4, 2021.

Martius, O., Hering, A., Kunz, M., Manzato, A., Mohr, S., Nisi, L., and Trefalt, S.: Challenges and Recent Advances in Hail Research, Bulletin of the American Meteorological Society, 99, ES51–ES54, https://doi.org/10.1175/BAMS-D-17-0207.1, 2018.

Mason, B. J.: The physics of clouds, Oxford classic texts in the physical sciences, Clarendon Press, Oxford, 2nd ed edn., https://doi.org/10.1002/qj.49709841723, 1971.

Miller, R.: RMS obtains hail risk model, Business Insurance, p. 1, https://www.businessinsurance.com/article/20070109/STORY/20009195/RMS-obtains-hail-risk-model, last access: 29 August 2023, 2007.

Miralles, O., Davison, A. C., and Schmid, T.: Bayesian modeling of insurance claims for hail damage [in review], The Annals of Applied Statistics, https://doi.org/10.48550/arXiv.2308.04926, 2023.

Nisi, L., Martius, O., Hering, A., Kunz, M., and Germann, U.: 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, 1590–1604, https://doi.org/10.1002/qj.2771, 2016.

Punge, H. and Kunz, M.: Hail observations and hailstorm characteristics in Europe: A review, Atmospheric Research, 176-177, 159–184, https://doi.org/10.1016/j.atmosres.2016.02.012, 2016.

Puskeiler, M., Kunz, M., and Schmidberger, M.: Hail statistics for Germany derived from single-polarization radar data, Atmospheric Research, 178-179, 459–470, https://doi.org/10.1016/j.atmosres.2016.04.014, 2016.

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegrìa, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., and Rama, B., eds.: Climate Change 2022: Impacts, Adaptation and Vulnerability, Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, https://report.ipcc.ch/ar6/wg2/IPCC_AR6_WGII_FullReport.pdf, 2022.

Púčik, T., Groenemeijer, P., Rädler, A. T., Tijssen, L., Nikulin, G., Prein, A. F., Meijgaard, E. v., Fealy, R., Jacob, D., and Teichmann, C.: Future Changes in European Severe Convection Environments in a Regional Climate Model Ensemble, Journal of Climate, 30, 6771–6794, https://doi.org/10.1175/JCLI-D-16-0777.1, 2017.

Púčik, T., Castellano, C., Groenemeijer, P., Kühne, T., Rädler, A. T., Antonescu, B., and Faust, E.: Large Hail Incidence and Its Economic and Societal Impacts across Europe, Monthly Weather Review, 147, 3901–3916, https://doi.org/10.1175/MWR-D-19-0204.1, 2019.

Raupach, T. H., Martius, O., Allen, J. T., Kunz, M., Lasher-Trapp, S., Mohr, S., Rasmussen, K. L., Trapp, R. J., and Zhang, Q.: The effects of climate change on hailstorms, Nature Reviews Earth & Environment, 2, 213–226, https://doi.org/10.1038/s43017-020-00133-9, 2021.

Röösli, T., Appenzeller, C., and Bresch, D. N.: Towards operational impact forecasting of building damage from winter windstorms in Switzerland, Meteorological Applications, 28, e2035, https://doi.org/10.1002/met.2035, 2021.

Saltikoff, E., Tuovinen, J.-P., Kotro, J., Kuitunen, T., and Hohti, H.: A Climatological Comparison of Radar and Ground Observations of Hail in Finland, Journal of Applied Meteorology and Climatology, 49, 101–114, https://doi.org/10.1175/2009jamc2116.1, 2010.

https://doi.org/10.5194/nhess-2023-158
Preprint. Discussion started: 19 September 2023

Natural Hazards and Earth System Sciences
Discussions

Schmidberger, M.: Hagelgefährdung und Hagelrisiko in Deutschland basierend auf einer Kombination von Radardaten und Versicherungsdaten, PhD Thesis, Karlsruher Institut für Technologie (KIT), https://doi.org/10.5445/KSP/1000086012, iSBN: 978-3-7315-0846-5, 2018.

Schroeer, K., Trefalt, S., Hering, A., Germann, U., and Schwierz, C.: Hagelklima Schweiz: Daten, Ergebnisse und Dokumentation, Fachbericht MeteoSchweiz 283, MeteoSchweiz, https://polybox.ethz.ch/index.php/s/VfW1BMBHdT0JrlW, [in prep.; available for reviewers at provided URL], 2023.

Schuster, S. S., Blong, R. J., and McAneney, K. J.: Relationship between radar-derived hail kinetic energy and damage to insured buildings for severe hailstorms in Eastern Australia, Atmospheric Research, 81, 215–235, https://doi.org/10.1016/j.atmosres.2005.12.003, 2006.

Schwierz, C., Köllner-Heck, P., Zenklusen Mutter, E., Bresch, D. N., Vidale, P.-L., Wild, M., and Schär, C.: Modelling European winter wind storm losses in current and future climate, Climatic Change, 101, 485–514, https://doi.org/10.1007/s10584-009-9712-1, 2010.

Skripniková, K. and Řezáčová, D.: Radar-based hail detection, Atmospheric Research, 144, 175–185, https://doi.org/10.1016/j.atmosres.2013.06.002, 2014.

Stucki, M. and Egli, T.: Synthesebericht: Elementarschutzregister Hagel, Tech. rep., Präventionsstiftung der Kantonalen Gebaeudeversicherungen, Bern, Switzerland, 2007.

Trefalt, S., Germann, U., Hering, A., Clementi, L., Boscacci, M., Schröer, K., and Schwierz, C.: Hail Climate Switzerland Operational radar hail detection algorithms at MeteoSwiss: quality assessment and improvement, Technical Report MeteoSwiss 284, MeteoSwiss, https://polybox.ethz.ch/index.php/s/WKoqBaSDWqc9d2k, [in prep.; available for reviewers at provided URL], 2023.

Treloar, A. B. A.: Vertically integrated radar reflectivity as an indicator of hail size in the Greater Sydney region of Australia., in: Proceedings of 19th Conference on Severe Local Storms, 14–18 September 1998, pp. 48–51, American Meteorological Society, Minneapolis, 1998.

Waldvogel, A., Schmid, W., and Federer, B.: The Kinetic Energy of Hailfalls. Part I: Hailstone Spectra, Cover Journal of Applied Meteorology and Climatology Journal of Applied Meteorology and Climatology, 17, 515–520, https://doi.org/10.1175/1520-0450(1978)017,0515:TKEOHP.2.0.CO;2., 1978.

Waldvogel, A., Federer, B., and Grimm, P.: Criteria for the Detection of Hail Cells, Journal of Applied Meteorology and Climatology, 18, 1521–1525, 1979.

Warren, R. A., Ramsay, H. A., Siems, S. T., Manton, M. J., Peter, J. R., Protat, A., and Pillalamarri, A.: Radar-based climatology of damaging hailstorms in Brisbane and Sydney, Australia, Quarterly Journal of the Royal Meteorological Society, 146, 505–530, https://doi.org/10.1002/qj.3693, 2020.

Wilks, D. S.: Chapter 9 - Forecast Verification, in: Statistical Methods in the Atmospheric Sciences (Fourth Edition), edited by Wilks, D. S., pp. 369–483, Elsevier, 4th edition edn., https://doi.org/10.1016/B978-0-12-815823-4.00009-2, 2019.

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

Yeo, S., Leigh, R., and Kuhne, I.: The April 1999 Sydney Hailstorm, The Australian Journal of Emergency Management, 14, 23–25, https://search.informit.org/doi/abs/10.3316/informit.391627197118013, 1999.

Yin, J., Ogane, Y., and Jinnai, K.: Modeling hail risk in the contiguous United States for insurance loss estimation, in: Proceedings of the 12th International Conference on Wind Engineering, vol. 2, pp. 1751–1758, Cairns, Australia, 2007.