ECSS Conference 2025
17.-21.11.25 – Utrecht, NL
P95: ECSS2025-184

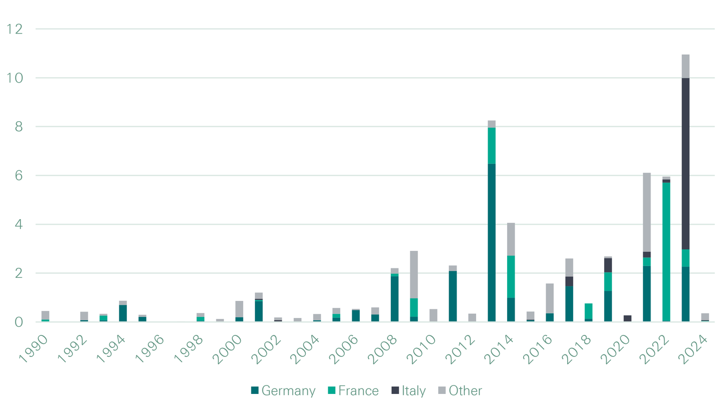
From Observation to Simulation: Building a Continental-Scale Hail Model for Re/Insurance Applications

Joel Zeder (Joel_Zeder@swissre.com), Elisa Spreitzer, Rogier de Jong, and Elisabeth Viktor
Cat Perils, Swiss Re

Recent SCS Loss Trends in Europe

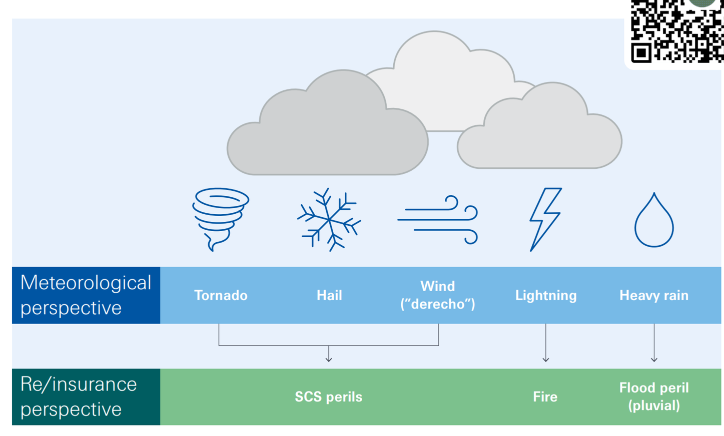
- **Rising Losses:** Severe convective storms (SCS) now cause event losses of several billion Euros, comparable to peak Nat Cat perils.
- **Defining Recent Extreme Events:**
  - *Volker in DE/AT 2021*: Multi-day outbreak with high accumulation.
  - *Qiara/Maya in FR 2022*: Record hail losses nationwide.
  - *Unai/Tisna in IT 2023*: Giant hail (19 cm) and unprecedented loss creep.
- **Loss Drivers:** Regionally varying changes in large hail frequency driven by climate change, new vulnerabilities (e.g. solar PV) and market soft factors.
- **Early Model Deficiencies:** Previous SCS models often miss spatio-temporal event clustering and cross-border correlation.
- **Market Need:** Advanced models enable better underwriting and risk selection and reduce the likelihood of unpleasant surprises.

Annual insurance losses from severe convective storms (SCS), considering major events in Europe, in USD billions, at 2025 prices based on US consumer price index (not corrected for exposure changes and other macro trends).



Challenges in SCS Loss Modelling

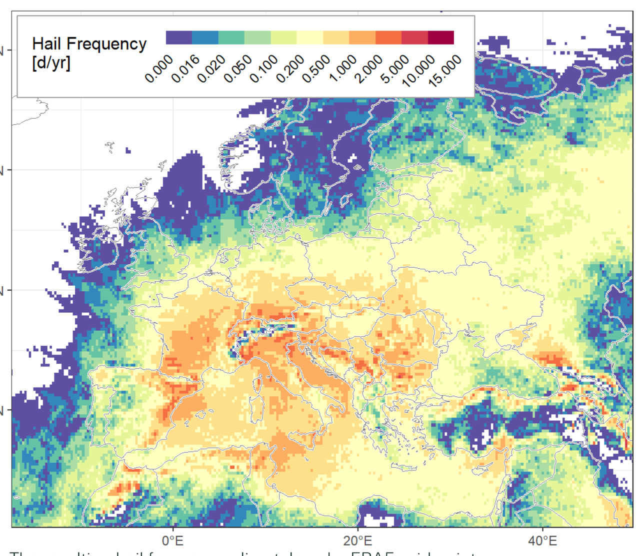
- **Small Scale & Short Duration**: Most storms are localised but very intense, resulting in a large hit-or-miss potential
- **Complex Clustering**: Under certain conditions, storms merge into long-lived systems which can last for multiple days, highlighting the importance of loss occurrence clauses.
- **Different Sub-Perils**: Hail dominates insured losses in Europe, however severe winds and water ingress can aggravate the losses.
- **Data Availability**: Monitoring and reporting standards vary substantially between European countries (hazard & loss).
- **Loss Drivers**: The relative contribution of different loss drivers – such as inflation, economic growth, urbanization, climate change, changing vulnerabilities and societal and behavioural trends – remains highly uncertain.



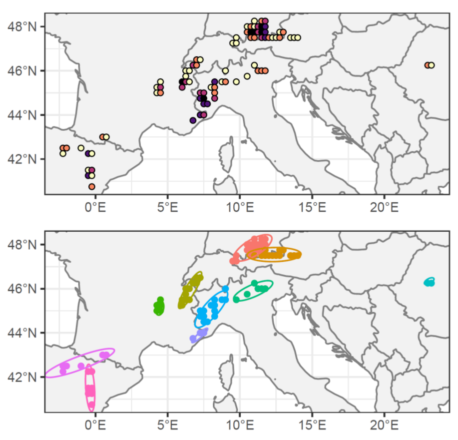
Hail Climatology and Event Statistics

The basis of this hail model is a spatially and temporally consistent dataset of daily hail probabilities by Blanc et al. (in prep.), using machine learning to derive hail probabilities from daily ERA5 fields.

- **Input Datasets**: Hail reports (≥25 mm diameter) from NOAA (USA), BoM (Australia), and ESSL (Europe) are spatially and temporally matched to ERA5 reanalysis data (0.25° grid).
- **Statistical Model**: An XGBoost (gradient boosting trees) model is trained regionally (USA, Australia, Europe) and globally (combined data). Eleven atmospheric predictors are used (CAPE, FLH, CIN, VS, SRH, Td, RH, TT index). Bayesian optimisation tunes hyperparameters of the model.
- **Output**: The model outputs daily hail probabilities (0 to 1) for each grid cell. A threshold (≥0.5) defines modelled hail events. The temporal aggregate is a hail frequency map for Europe (right).
- **Evaluation Approach**: Performance is assessed by comparing modelled vs. observed hail frequencies in major cities, using metrics like RMSE and correlation.



The resulting hail frequency climatology by ERA5 grid point



Visualisation of the GMM fit

The dataset of daily hail probabilities can be further leveraged to derive statistical parameters describing the geometry, orientation, and clustering of hail environments, which are approximated by ellipses.

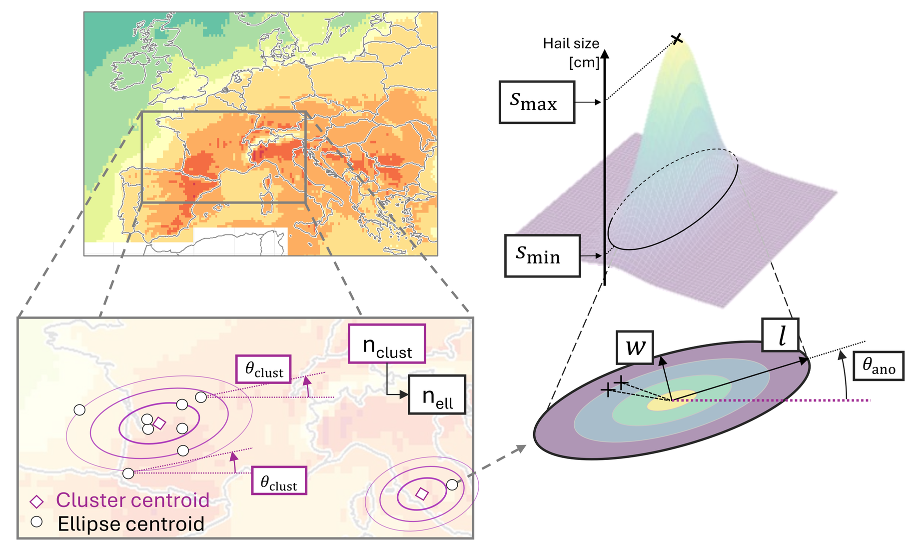
- **Gaussian Mixture Models** (GMM): For each event date, spatial clusters of high hail probability grid points are identified by fitting a flexible GMM to the respective coordinates (left), selecting the number of clusters via Bayesian Information Criterion (BIC).
- **Ellipse Extraction**: Distribution parameters (centroid, orientation, and spread) are derived from the GMM covariance matrices via eigen-decomposition, enabling each hail cluster to be represented as an ellipse with quantified location, size, and angle.
- **Event Statistics**: Additional metrics derived from the fitted ellipses include cluster size and count, spatial distribution of ellipse centroids, orientation angles, length and width statistics, and the fraction of each ellipse area affected by hail.

Sampling Hail Footprints

The stochastic hail footprints are stored on so-called Calculation Units, a high-resolution grid (3 km in densely populated areas, 9 km elsewhere), which serve as the spatial reference for mapping exposure and local event intensities.

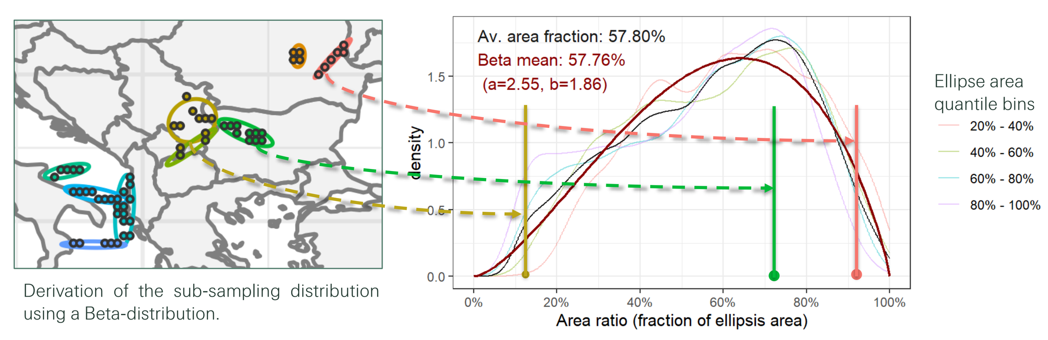
The sampling is based on the following steps:

- **Centroid Sampling**: For each event, the number and location of clusters and ellipses are stochastically sampled using geometric distributions and spatial Poisson processes, with parameters learned from the GMM fit.
- **Footprint Sampling**: Ellipse geometry (length, width, orientation) is sampled from Gamma distributions and Gaussian copulas, with intensity correlations set according to Punge et al. (2014).
- **Hail Intensity**: Hail intensity at each CU is assigned by combining the sampled maximum intensity and a Gaussian decay from the ellipse centroid, ensuring realistic spatial severity patterns.



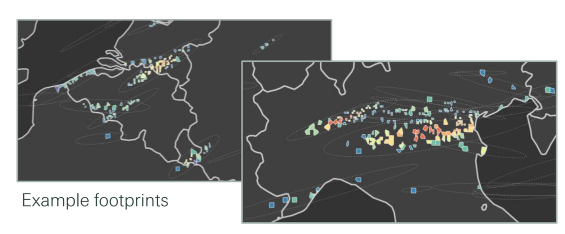
Visualisation of the various sampling steps

- **Subsampling**: Subsampling steps adjust the footprint area to reflect localized hail occurrence and observational limitations, using beta and logistic functions for retention rates.



Derivation of the sub-sampling distribution using a Beta-distribution.

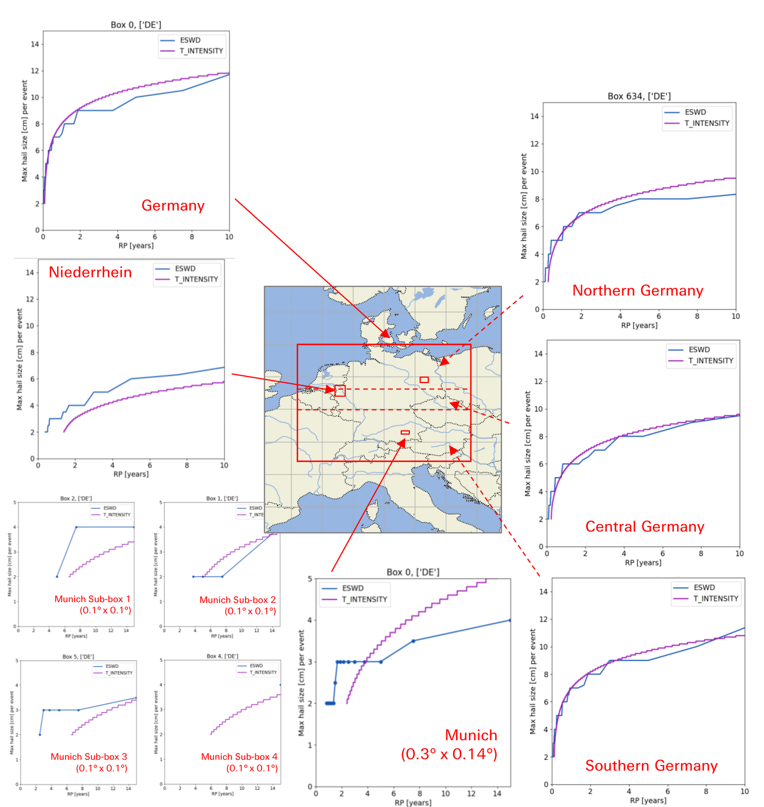
Resulting Event Set: As the algorithm generates several millions of stochastic hail footprints across continental Europe, the resulting event set can even be used to assess the hail risk for individual sites.



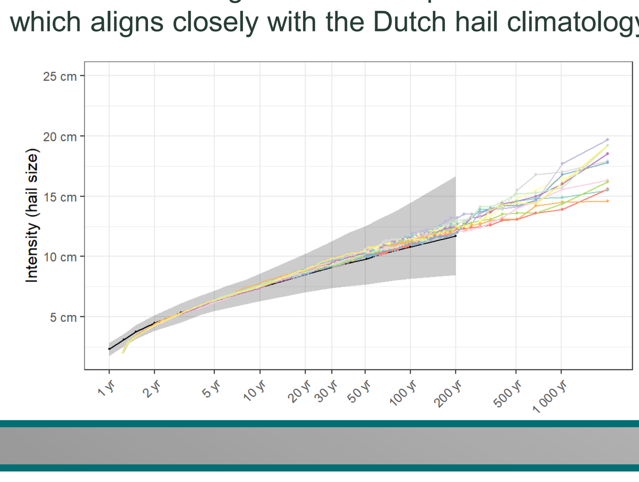
Example footprints

Evaluating the Frequency – Severity Hazard Relationship

The severity-frequency relationship is assessed by aggregating quality-checked ESWD hail observations (Dotzek et al., 2009) within various spatial boxes across Germany, from national to metropolitan scales. For each area, empirical return level curves are derived, representing the frequency of severe hail events of different magnitudes. These curves can be compared to the respective return level curves from the event set, as they provide lower-bound estimates of event frequency, based on the observed data.

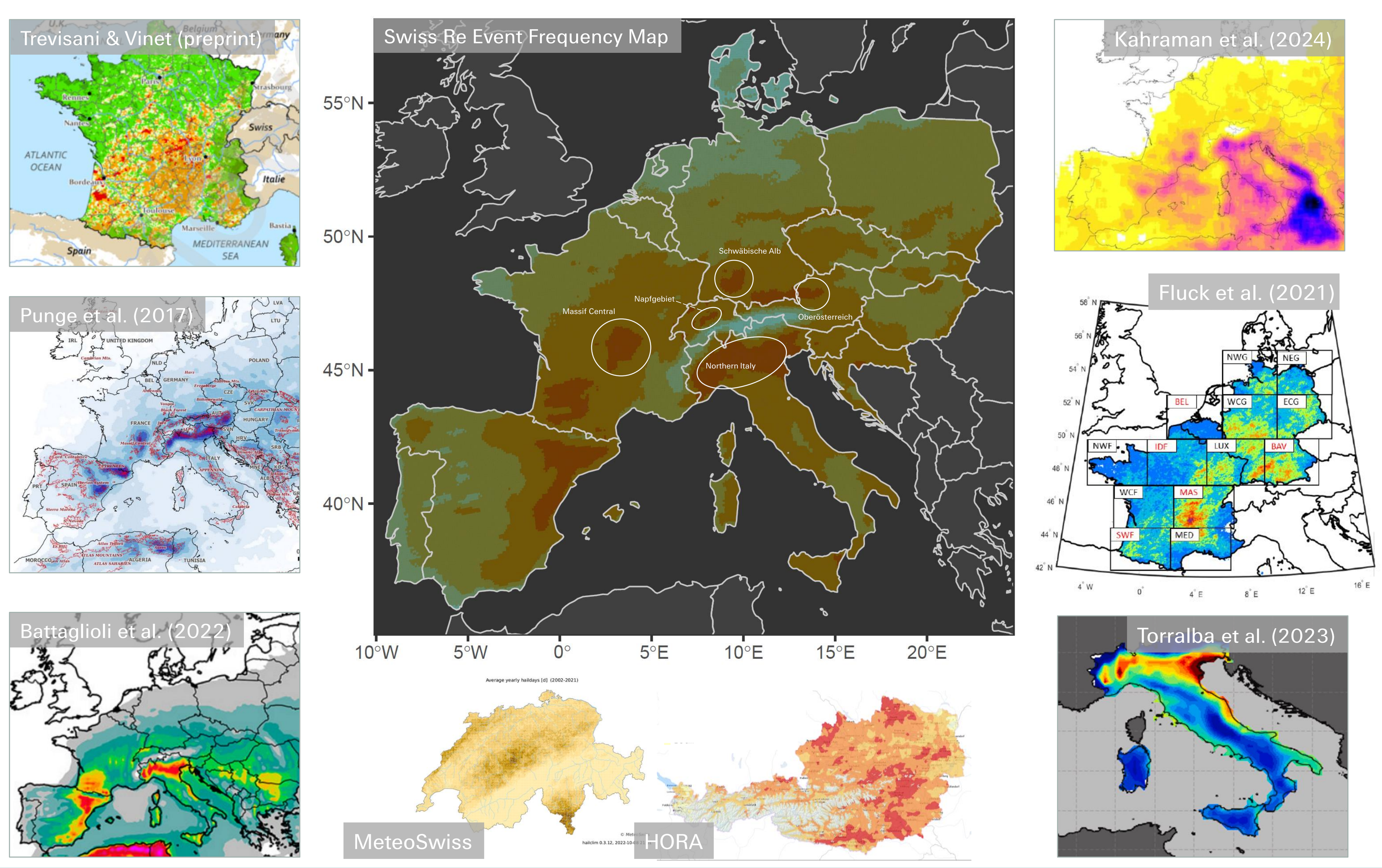


Another reference dataset to assess the intensity-frequency relationship is provided by Wouters et al. (2019), offering return period curves for maximum hailstone size across the Netherlands and its subregions. These area-integrated curves enable a straightforward comparison with the event set, which aligns closely with the Dutch hail climatology.



Evaluating the Geographical Risk Differentiation

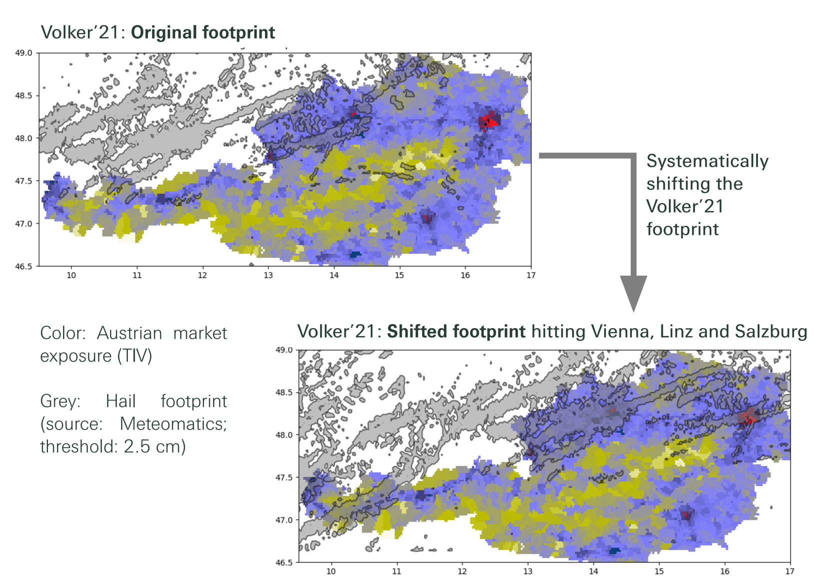
The hail frequency map derived from the event set captures all major hail hotspots consistent with other established climatologies. In addition to reanalysis-based approaches, these reference datasets rely on sources such as remote sensing (radar or satellite overshooting top detections), convection-resolving models, and ground-based observations.



Calibration of the Tail

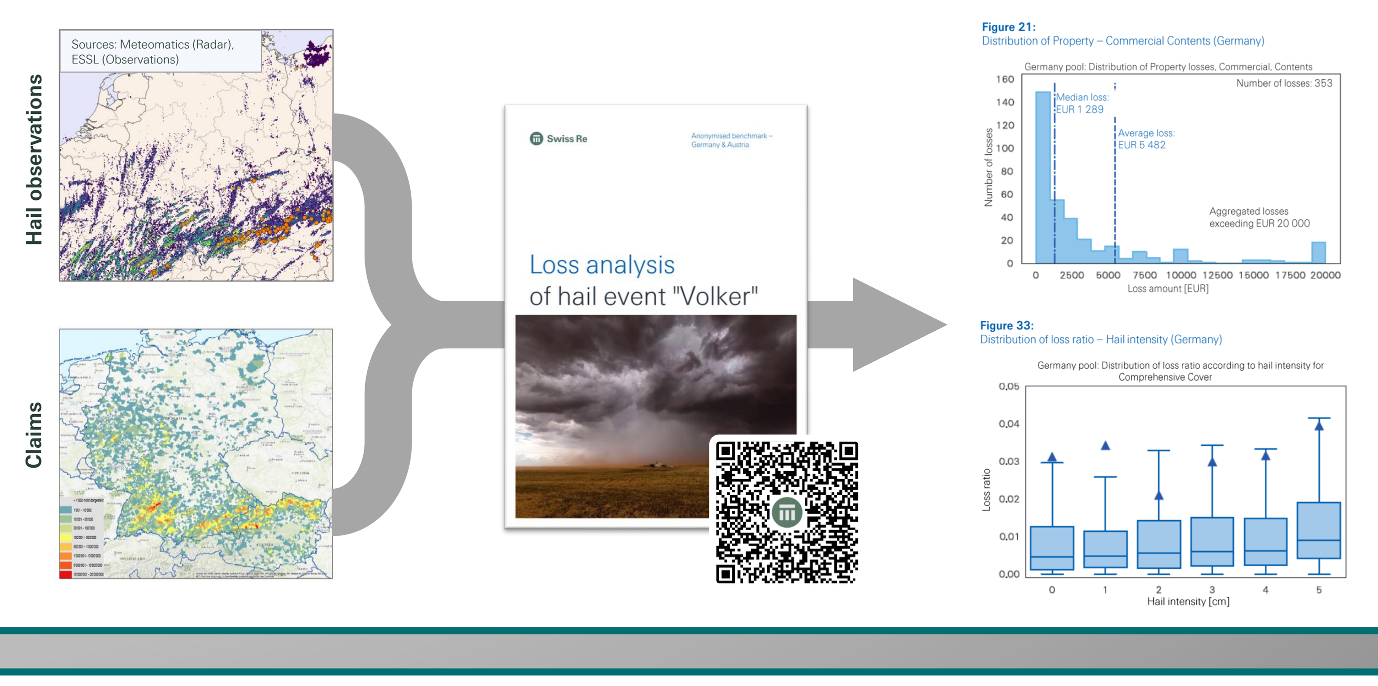
For high-frequency hail events, model calibration is possible using historical loss experience; however, for the tail of the loss distribution, few (if any) relevant benchmarks have been observed to date.

- **Shifting Historic Events**: Tail benchmarks are created by taking the footprints of major historic hail events (e.g. Andreas 2013, Volker 2021, or Qiara 2022) and systematically shifting them across the region to simulate their impact on different high-exposure areas.
- **Maximum Plausible Loss**: This method estimates the maximum plausible losses if a severe event were to strike a more vulnerable or densely insured location.
- **Tail Benchmarks**: The resulting scenarios provide realistic, stress-test benchmarks for assessing portfolio risk from rare, extreme hail events.



Vulnerabilities

- **Vulnerability Base Curves**: Vulnerability parameters were calibrated for seven base occupancies: mixed residential, commercial, industrial, agriculture, motor hull, greenhouse, and solar power plant. Each base curve reflects typical damage levels for its occupancy class, depending on the intensity of the event.
- **Detailed Sub-Occupancies**: Sub-occupancy relativities (e.g., single-family vs. multi-family housing) were derived from detailed claims analyses, especially the Volker'21 event, which provided occupancy-specific mean damage degrees (MDD).
- **Calibration by Market**: Country- and city-level adjustments were applied to account for differences in exposure concentration, insured values (e.g. varying TIV assumptions for cars across markets), and cultures in coverage encoding (e.g. "Content" in Switzerland), with calibration based on recent loss experience.



References

Battaglioli, F., et al. (2023): Modeled Multidecadal Trends of Lightning and (Very) Large Hail in Europe and North America (1950–2021), Journal of Applied Meteorology and Climatology, 1627–1653.

Blanc, B., et al. (in prep.): Improving our Understanding of Large Hail Hazards Using Machine Learning, in prep.

Dotzek, N., et al. (2009): Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. Atmos. Res., 93, 575–586.

Fluck, E., et al. (2021): Radar-based assessment of hail frequency in Europe, Nat. Hazards Earth Syst. Sci., 21, 683–701.

Kahraman, A., et al. (2024): Climatology of severe hail potential in Europe based on a convection-permitting simulation. Climate Dynamics, 62, 6625–6642

Punge, H. J., et al. (2014): A new physically based stochastic event catalog for hail in Europe, Nat. Hazards, 73, 1625–1645.

Punge, H. J., et al. (2017): Hail frequency estimation across Europe based on a combination of overshooting top detections and the ERA-INTERIM reanalysis, Atmospheric Research, 198, 34–43.

Torralba, V., et al. (2023): Modelling hail hazard over Italy with ERA5 large-scale variables, Weather and Climate Extremes, 39, 100535.

Trevisani, M., & Vinet, F. (preprint): A 10-Year climatology of hail in France: towards a new estimate of damaging hail.

Wouters, L., et al. (2019): A Hail Climatology of the Netherlands. 37 p.