Improving nowcasts of hail in Switzerland using deep neural networks

George Pacey[1], Ulrich Hamann[2], Ophélia Miralles[3] and Olivia Martius[1]
[1]University of Bern, Switzerland
[2]MeteoSwiss, Locarno, Switzerland
[3]Norweigan Meteorological Institute, Oslo, Norway

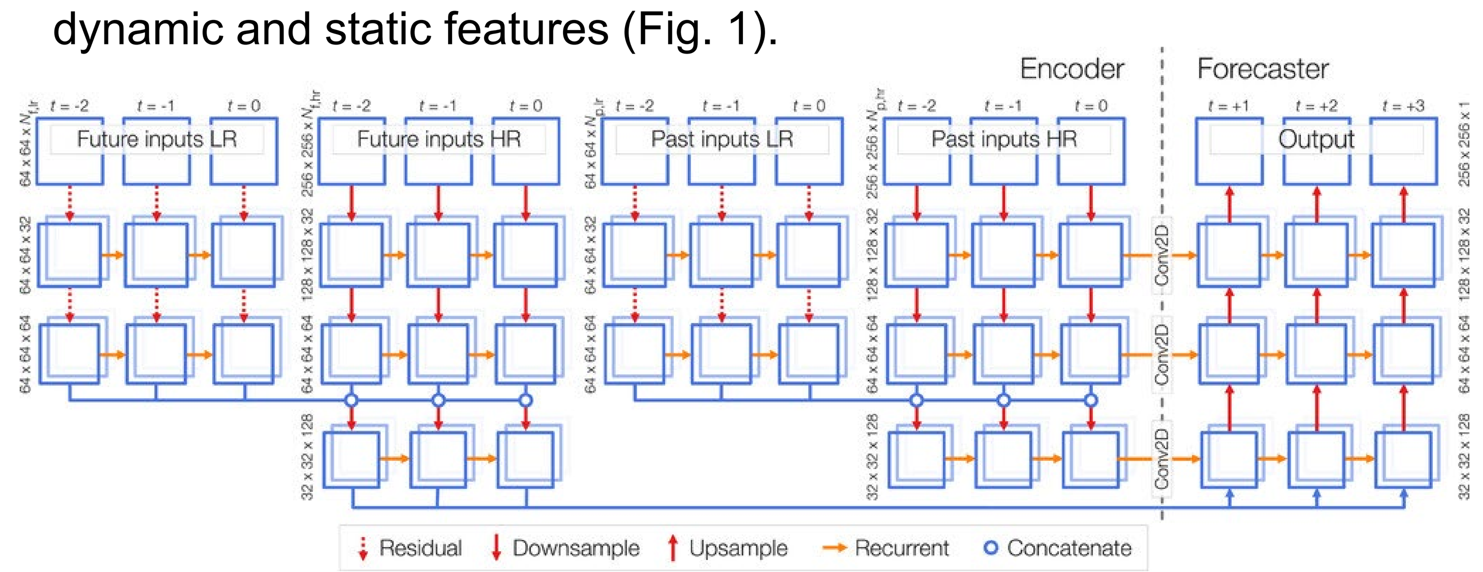
UNIVERSITÄT BERN
OESCHGER CENTRE CLIMATE CHANGE RESEARCH
Mobiliar Lab für Naturrisiken
MeteoSwiss

Introduction and motivation

- Nowcasting convective storm hazards is vital for several sectors including aviation, emergency services and the public.
- Machine learning, especially deep learning (DL), has emerged as a promising alternative to traditional nowcasting approaches (e.g., Lagrangian extrapolation).
- MeteoSwiss will use a deep learning model (COALITION4) operationally next year for thunderstorm warnings. [1,2].
- COALITION4 produces nowcasts of precipitation, lightning and hail at a 1 km spatial resolution for the next hour.
- Despite encouraging advancements in recent years, there is still large potential for future development of DL nowcasting models.
- Convective hazards are particularly difficult to model due to the complex non-linearity of convective initiation and the evolution of convective storms.
- Here, we present an approach that has made significant improvements over the original model for the hail hazard.

Model architecture

- The model combines a U-Net architecture with gated recurrent units (GRU) to capture spatial and temporal relationships [3].
- The architecture is flexible for inputs with different resolutions as well as both dynamic and static features (Fig. 1).



Loss function: Binary cross entropy (pointwise)

Figure 1 from [3] – showing the UNetGRU used by MeteoSwiss (COALITION4). LR = low resolution inputs, HR = high resolution inputs.

- Target variable: probability of hail (**POH**) – a radar proxy for hail [4]
- Model operates at a 5-minutely and 1 km resolution.

Retraining setup and verification

Primary differences between the original and retrained model

Configuration	Original model	Retrained model
Training years	2020	2019–2025 (excluding 2021)
Tile size	256 x 256	160 x 160
Input predictors	7 radar, 2 lightning, 3 static	3 radar and 3 static
Previous timesteps input	6	8
Data augmentation	Rotations and mirroring	None
Tile section	> 10 mm hr^{-1} precipitation (10 pixels)	> 0.25 POH (15 pixels)

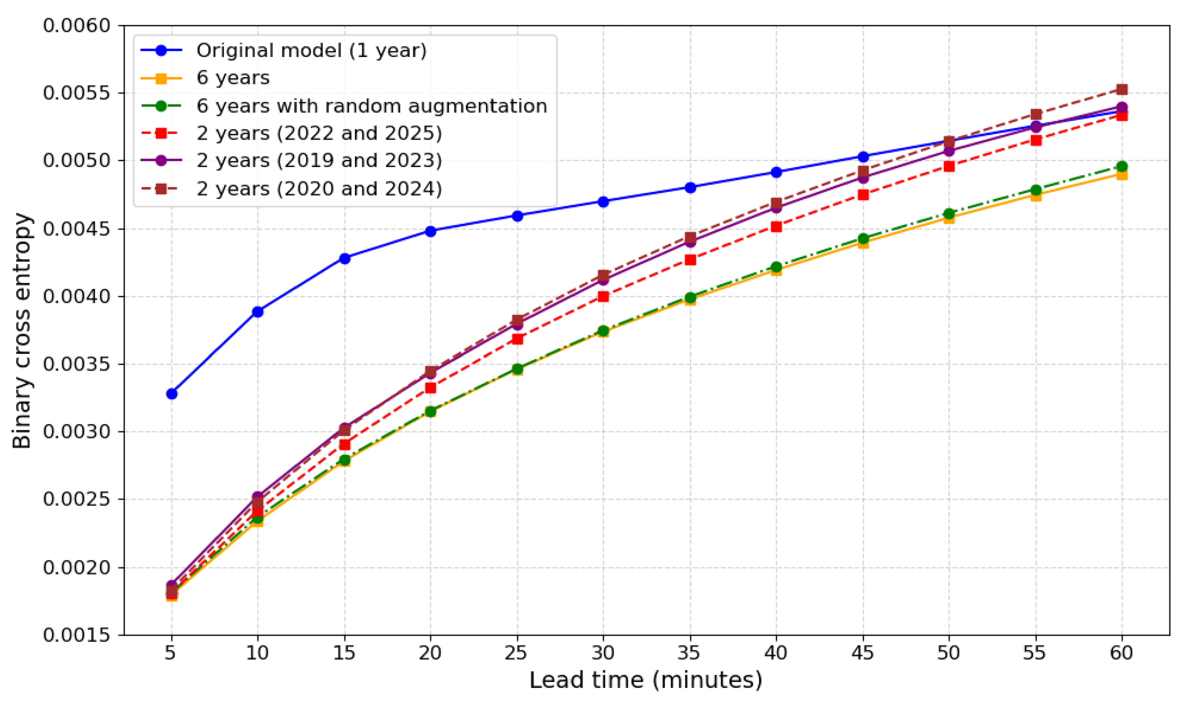


Figure 2: Binary cross entropy as a function of lead time during the 2021 convective season. The yellow and blue lines show the retrained and original models, respectively. The green line shows a model where random transformations (flipping/rotation) are applied in the batch generator. The remaining three lines show models using just two years of training data instead of six.

- Radar archives go back over a decade in Switzerland offering the possibility to train with much more data.
- The tile size decrease reduces the number of pixels with zero potentially enabling the model to better focus on how hail cells evolve in time.
- Feature importance analysis revealed several low importance features in the original model; these were removed to reduce redundancy.
- The resulting retrained model is superior to the original model at all lead times, especially at earlier lead times (Fig. 2).

Model calibration plots → relevant for warnings

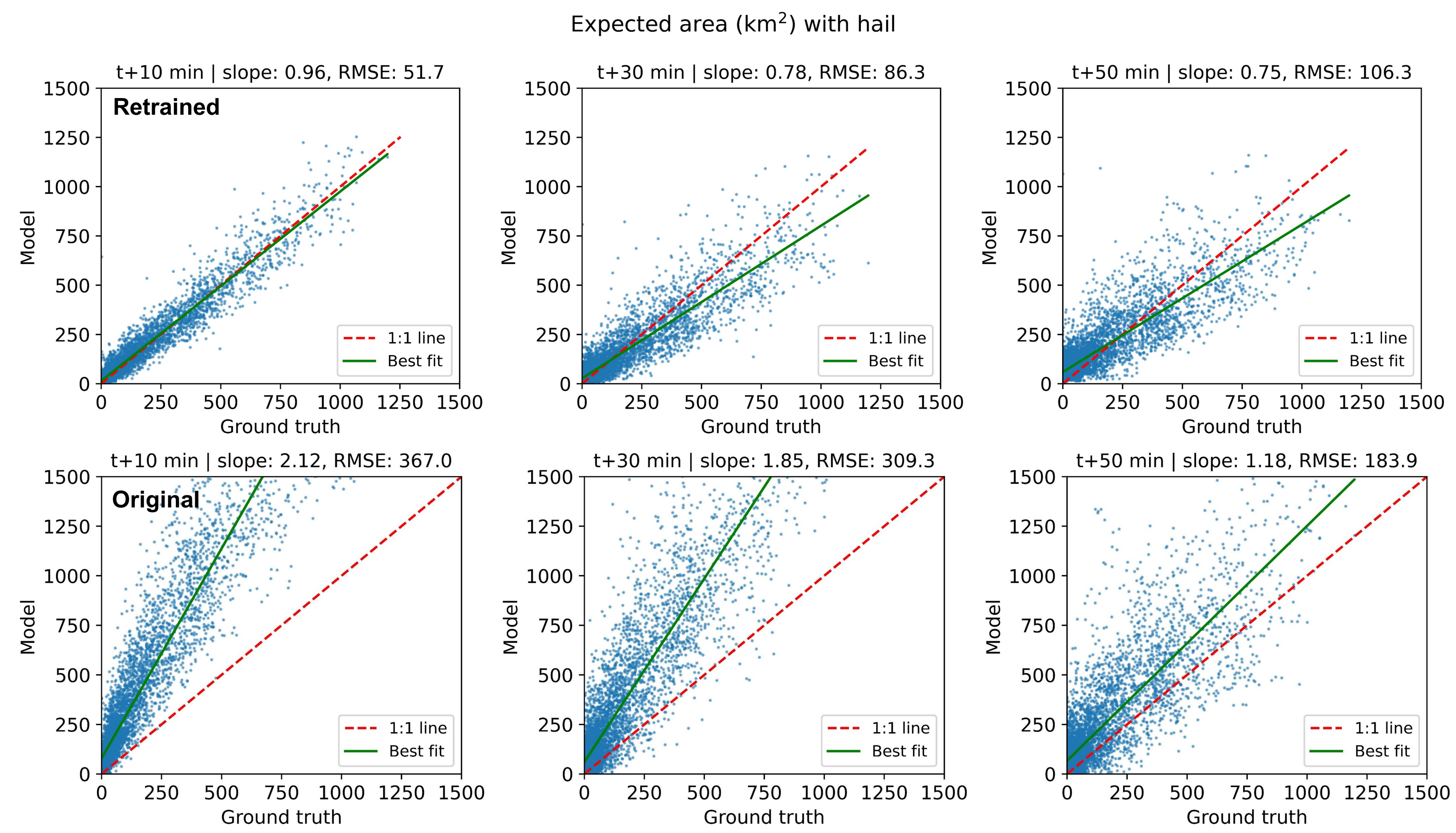


Figure 3: Each point on the scatter plot is a nowcast initialised from a different time in 2021. Going from left to right are nowcasts at lead times 10, 30 and 50 minutes. The top and bottom row show the retrained and original model, respectively. Probabilities are accumulated across the entire COALITION4 domain (see Fig. 4) to give the expected area in km² with hail. A perfectly calibrated nowcast would be along the 1:1 line. In operations, probabilities will instead be integrated over individual warning regions.

- The retrained model overall also shows a better calibration of model probabilities although the calibration is notably worse at later lead times (Fig. 3).

Case study: record-breaking hail day of 28 June 2021

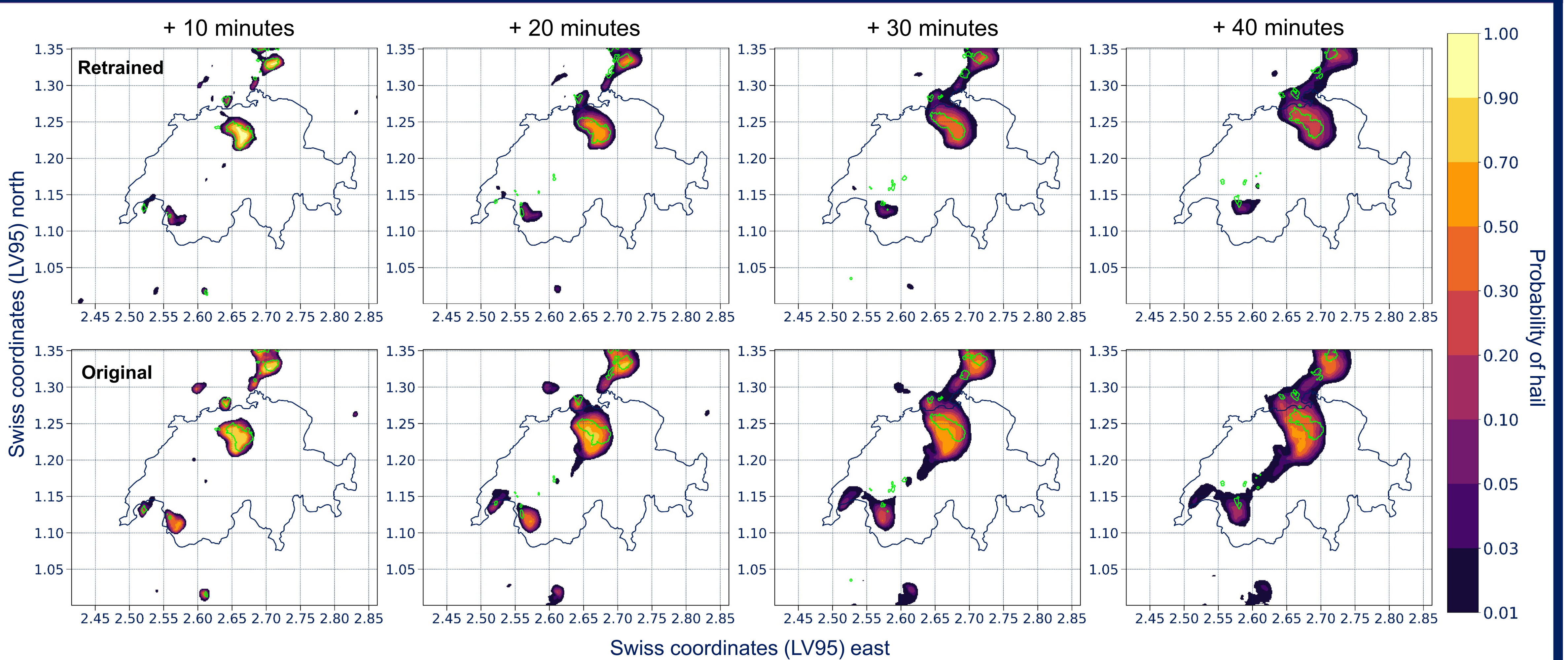


Figure 4: Nowcast initialised on 28 June 2021 at 1615 UTC. The top and bottom rows show the retrained and original model, respectively. The green contour shows the ground truth based on the 1% probability contour.

- On 28 June 2021, multiple insurance companies reported record damage claims from a single hail event [5].
- The mesoscale situation was complex with long-living supercells as well as cell mergers.
- For an example initialisation time (Fig. 4), the improved preciseness of the retrained model is apparent with the cell motion and structure being better represented. More initialisation times from this case and others can be viewed by scanning the QR code.

More cases (GIFs)

https://gpacey.github.io/C4_UNetGRU_output_GIF/

Summary and outlook

- The retraining experiments show significant improvement for the probability of hail model, particularly at earlier lead times.
- Sensitivity experiments are performed to highlight reasons for the increased skill.
- The additional years of training data are particularly important for skill at later lead times.
- Commonly applied random augmentation through flipping or rotating tiles did not significantly affect model skill.
- Providing the model information about the storm's environment (e.g. instability and wind shear) from the ICON forecast model should increase skill at later lead times.
- More complex loss functions are also a promising avenue for enhancing model performance [6].

References

[1]: Hamann et al. (2025), ECSS2025, P55
[2]: Leinonen et al. (2023), Geophysical Research Letters
[3]: Leinonen et al. (2022), AIES
[4]: Foote et al. (2005), 85th AMS Annual Meeting
[5]: GVL (2022), Hagelereignis überschattet das Geschäftsjahr, https://www.gvl.ch
[6]: Miralles et al. (2026), AIES

Contact: george.pacey@unibe.ch