https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

Exploiting radar polarimetry for nowcasting thunderstorm hazards using deep learning

Nathalie Rombeek[1], Jussi Leinonen[1], and Ulrich Hamann[1]

[1]Federal Office of Meteorology and Climatology MeteoSwiss, Locarno-Monti, Switzerland

Correspondence: Ulrich Hamann (ulrich.hamann@meteoswiss.ch)

Abstract. Severe convective weather events, such as hail, lightning and heavy rainfall pose a great threat to humans and cause a considerable amount of economic damage. Nowcasting convective storms can provide warning signals and mitigate the impact of these storms. Dual-polarization weather radars are a crucial source of information for nowcasting severe convective events; nevertheless, they are most often not considered in nowcasting. These radars provide signatures of different hydrometeors. This work presents the importance of polarimetric variables as an additional data source for nowcasting thunderstorm hazards using an existing neural network architecture with convolutional and recurrent layers. This network has a common framework, which enables nowcasting of hail, lightning and heavy rainfall for lead times up to 60 min with a 5 min resolution. The study area is covered by the Swiss operational radar network, which consists of five operational polarimetric C-band radars. Results indicate that including polarimetric variables and quality indices improve the accuracy of nowcasting heavy precipitation and lightning, with the largest improvement found for heavy precipitation.

1 Introduction

Severe convective weather events, such as hail, lightning and heavy precipitation, are likely to increase across Europe during this century (Rädler et al., 2019; Raupach et al., 2021; Taszarek et al., 2021). These convective storms can turn into flash floods and consequently be a great threat to humans (Holle et al., 1993; Lynn and Yair, 2010). Additionally, a considerable part of the total weather-related economic losses are caused by these weather phenomena (Hoeppe, 2016). Therefore, accurate short-term predictions of convective events are of interest, as they allow to send out warnings in order to reduce societal and economic impact.

Numerical weather prediction (NWP) models are useful in predicting stratiform long-duration storms. However, forecasting convective activity is challenging in weather modeling. These models typically have a low update frequency (e.g. every 3 hours for COSMO-1E, the NWP model used in Switzerland by MeteoSwiss) and, due to high demand in computation time results, are not available in real time (Pierce et al., 2012). In addition, NWP often have a lower predictability for shorter lead times (i.e. 1 h) and smaller scales (Simonin et al., 2017) as the initial state of the atmosphere in NWP assimilation is based on previous model predictions rather than the latest available observations, which makes it less suited for accurately predicting the time and location of convective storms. For this reason, very short-term forecasts, also called nowcasts, can play a crucial role in early warning systems, as they exploit the most recent observations in near real time.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

Weather radars are often utilized for nowcasting purposes as they provide high resolution near real time input data with a broad spatial coverage. Conventional nowcasting techniques extrapolate these latest observations in time, based on either estimation of the motion field such as used in Pysteps (Pulkkinen et al., 2019), NowPrecip (Sideris et al., 2020) or rainymotion (Ayzel et al., 2019), or identifying and tracking individual storms (e.g. Thunderstorms Radar Tracking (TRT; Hering et al., 2004) or Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN; Dixon and Wiener, 1993)). However, these methods often have difficulties to take all growth and dissipation processes into account and consequently result in relatively short skilful lead times for convective weather (Imhoff et al., 2020; Foresti et al., 2016; Wilson et al., 1998).

In the recent years there have been significant advances in using deep learning for generating nowcasts of heavy precipitation using radar as input, such as done by Guastavino et al. (2022), Han et al. (2021), Ritvanen et al. (2023) and Yin et al. (2021), or in the case of Leinonen et al. (2023) including multiple data sources. In addition, radar is also used for nowcasting lightning, such as done by Leinonen et al. (2022b) and Zhou et al. (2020). However, these studies primarily focus on using the rainfall fields obtained from radar, despite the fact that weather radars can provide a much broader range of information. Adding polarimetric radar variables help considerably to reduce ambiguities concerning the hydrometeor classes and drop size distributions.

Dual-polarization radars have two orthogonally polarized beams, making it possible to derive additional properties such as particle shape and to some extent the size, which are useful for meteorological applications (Fabry, 2018; Kumjian, 2013b). Hydrometeor classification algorithms such as those developed by Besic et al. (2016) and Vivekanandan et al. (1999) use this extra information to identify different hydrometeors. Other studies showed the potential of polarimetric variables for providing information on other convective hazards, such as hail and lightning (Figueras i Ventura et al., 2019; Lund et al., 2009), or the evolution of convective storms (Snyder et al., 2015). However, interpretation of polarimetric signatures for convective weather forecasts remains challenging (Kumjian, 2013a, b), and requires more advanced data processing techniques such as machine learning.

For that reason this research will investigate the additional value of polarimetric variables for nowcasting severe convective weather, which includes hail, lightning and severe precipitation. This research uses the model from Leinonen et al. (2023), as it is able to utilize multiple data sources and predict, with a slight modification, multiple hazard types.

One of the first successful attempts to incorporate polarimetric variables for nowcasting convective precipitation using deep learning was done by Pan et al. (2021). However, that work only uses data of 3 km altitude Constant Altitude Plan Projection Indicator (CAPPI). In this study, we retrieve relevant information on microphysics from multiple altitudes. In addition, we investigate the potential for nowcasting hail and lightning by using polarimetric variables.

This paper introduces the data used for training in section 2, while section 3 describes the model architecture. Results are described and discussed in section 4, and section 5 concludes the article.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

2 Data

For training purposes, part of the dataset from Leinonen et al. (2022b) was used. This was extended with polarimetric variables retrieved from the Swiss operational radar network, and quality indices from Feldmann et al. (2021). The data were collected from April to September 2020.

2.1 Operational radar network

The study area is covered by the Swiss operational radar network, which consists of five operational polarimetric C-band radars (Germann et al., 2022). Operational available products have a resolution of 500 m, comprising 20 elevation scans from -0.2°to 40°within 5 min per radar. The maximum range of observations is 246 km.

Radar-derived precipitation fields are subject to a sophisticated data-processing chain which includes bias correction, removal of ground clutter and non-weather echoes, visibility correction and vertical profile correction (Germann et al., 2006), to get an improved precipitation estimate at the surface (RZC).

The final radar products that are used as input for the deep learning algorithm have a resolution of 1 km.

2.2 Data sources and processing

The model was trained based on all possible combinations of the data sources mentioned below:

Weather radar (R) observations from the Swiss operational network (Germann et al., 2022) were used. This source was already used in Leinonen et al. (2023), and contains information about the rain-rate, vertically integrated water content, echo top height and maximum echo.

Polarimetric variables (P) were also obtained from the Swiss operational radar network. The considered polarimetric variables in this research are the reflectivity factor at vertical polarization (Z_V), differential reflectivity (Z_{dr}), co-polar cross-correlation coefficient (ρ_{hv}), and specific differential phase (K_{dp}).

Z_{dr} is an indicator for shape, with positive values indicating targets that are larger in the horizontal than the vertical dimension. Such targets include large raindrops, which are flattened by aerodynamic forces while falling, but not solid hailstones, which tend to be round and therefore have values close to 0 (Seliga and Bringi, 1976).

K_{dp} is an indication for concentration and shape, and is used as a measure for rain intensity (Sachidananda and Zrnic, 1986). Positive values can be an indicator for heavy rain, while negative values means that targets are more elongated vertically than horizontally (e.g. graupel) and values close to zero indicate nearly round or randomly oriented particles (Rinehart, 2010). One advantage of K_{dp} over Z_{dr} is that it is unaffected by attenuation.

ρ_{hv} indicates homogeneity, with smaller values indicating more heterogeneity among the shape, size and orientation of the detected particles (Fabry, 2018).

To reduce the dimensionality and estimate values at the ground level, the polarimetric variables at various altitudes are aggregated following the method of Wolfensberger et al. (2021), which uses a weighted sum that takes both static radar visibility and height above the ground level of each point into account. Radar visibility is determined by the fraction of the

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

radar beam that is not blocked due to partial and total beam shielding by the complex mountainous terrain. The weight is determined using a linear relationship with visibility and an exponential relationship with height:

$$w(h) = \exp(\beta \frac{h}{1000}) \frac{\mathrm{VIS}}{100} \tag{1}$$

Here, h represents the height above the ground of the observation in meters, β indicates the slope of the exponential and VIS is the visibility. After hyper-parameter turning, a value of -0.5 for β was selected, as Wolfensberger et al. (2021) found that this resulted in the best parameter value. Next, the data was transformed by first normalizing it by bringing the mean close to 1. Second, to reduce presence of noise, fields were compared with RZC, and set to zero where RZC does not contain precipitation.

Quality indices (Q) were obtained from Feldmann et al. (2021). Quality of radar observations in mountainous terrain fluctuates over elevations and is influenced by the scanning strategy. Especially at low levels, visibility is reduced as a consequence of radar beam blockage. The quality of the observations at every location is influenced by multiple properties. The quality index combines the following factors into a single index: visibility, minimum altitude of observation, maximum altitude of observation and numerical noise.

2.3 Targets

The same target variables from Leinonen et al. (2023) and Leinonen et al. (2022b) are used, and summarized in the following:

Lightning occurrence is obtained from the observations by the EUCLID lightning network (Schulz et al., 2016; Poelman et al., 2016), delivered to MeteoSwiss by Météorage. The point-data was transformed to a gridded binary map, with 1 indicating lightning within a vicinity of 8 km of the grid point that took place in the last 10 min, and 0 otherwise.

Probability of hail (POH) is the probability of hail reaching the ground. This a product from the operational MeteoSwiss radar network, using the formula from Foote et al. (2005) based on Waldvogel et al. (1979). It utilizes the difference between the 45 dBZ echo top level and the freezing level.

CombiPrecip is an operational product at MeteoSwiss for estimating precipitation, which combines real-time radar and rain-gauge observations to adjust the biases that often are observed in radar measurements (Sideris et al., 2014a, b). This product is transformed to make probabilistic estimates in four pre-defined classes, based on warning levels used by MeteoSwiss forecasters. The thresholds are $R_0 = 0$, $R_1 = 10$ mm, $R_2 = 30$ mm and $R_3 = 50$ mm over 60 min aggregated precipitation at 1 km^2 grid point. Probabilities q_c are assigned to each class $c \in [0, 3]$ as

$$q_c = \int_{R_c}^{R_{c+1}} p(R)\,\mathrm{d}R \tag{2}$$

where p is a lognormal probability distribution function. CombiPrecip estimations are considered as the expected value $\mathrm{E}[R]$. The standard deviation is approximated as $\mathrm{Std}[R] = 0.33\mathrm{E}[R]$ by separating the error due to the lack of raingauge representation from the uncertainty in the radar measurement, using the method from Ciach and Krajewski (1999).

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

3 Methods

3.1 Neural network

The recurrent-convolutional deep learning model from Leinonen et al. (2023) is used, adding the newly introduced sources described in Sect. 2.2. This model has a common framework, which makes it possible to predict the probabilities of hail, lightning and heavy precipitation, by only changing the target. The main difference between the predicted thunderstorm hazards is that heavy precipitation is trained for only predicting the accumulated precipitation over 1 hour for predefined warning levels, while hail and lightning are produced at a 5-min resolution for 12 time steps (1 hour). For a more detailed description we refer to the publications of Leinonen et al. (2022b) and Leinonen et al. (2023).

Hail and precipitation targets have a probabilistic output, and for that reason cross entropy (CE; Goodfellow et al., 2016) was used as a loss function. CE measures the difference in the probability distributions between the true distribution and predicted distribution of the target classes. For lightning the focal loss (Lin et al., 2017) is used. The focal loss is an adaptation of the CE and focuses more on the difficult cases.

In order to estimate the influence of the random weights used for initialization and the consistency of the model, we trained each possible combination of data sources three times. As the sample size is rather small, we used the unbiased sample standard deviation for calculating the standard deviation between these runs.

3.2 Importance of data sources

The importance from individual data sources can be assessed using the Shapley value (Shapley, 1951), which distributes the total score among its predictors, by assigning a value that represents their marginal contribution. For more information on calculating the Shapley value we refer to the description of Molnar (2022) (chapter 9.5).

3.3 Model evaluation

Before calculating different metrics to evaluate the models, the ground truth for hail and precipitation were transformed to binary fields. For hail a threshold of 0.5 was selected, meaning that a $\mathrm{POH} \geq 50\%$ is considered as hail and set to 1, otherwise 0. For precipitation the predictability per class is analysed by summing all probabilities in and above the selected class. Second, a threshold of 0.5 is used, with setting probabilities ≥ 0.5 to 1.

The models are evaluated based on the critical success index (CSI), precision recall (PR) curve and the fractions skill score (FSS). The CSI and PR curve are based on contingency tables, containing true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN).

CSI indicates the amount of events that were correctly predicted:

$$\mathrm{CSI} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}. \tag{3}$$

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

When there is a imbalance between two classes (no event and event), the PR curve is a useful tool for interpretation of probabilistic forecasts. Precision indicates how good the model is at predicting an event:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

Recall gives the fraction of events that were predicted:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

The PR curve is obtained by computing both precision and recall at all threshold levels ranging from 0 to 1. The information of the PR-curve can be summarized by the area under the curve (AUC). A larger AUC indicates a better-performing model over the whole range of thresholds.

The FSS is a measure for neighbourhood verification, which measures the skill of the forecast in predicting the occurrence of an event at a selected spatial scale (Roberts and Lean, 2008). The FSS is the mean-square error of the observed and forecast fractions for a neighbourhood of length n, relative to a low-skill reference forecast. Values range between 0 and 1, with higher values indicating a more skilled forecast.

4 Results and discussion

4.1 Example cases

This section presents examples that illustrate the difference of adding polarimetric variables on top of non-polarimetric radar data for hazard prediction purposes. However, unlike for lightning and heavy precipitation, no significant differences were observed in the hail prediction, so no example is provided here.

Figure 1 serves as an example of the model output for lightning for several time steps. Both data source combinations (R: radar and RPQ: radar, polarimetric variables and quality indices) are able to accurately predict the location of the lightning. However, the difference is in the certainty of the predicted lightning over all lead times, with higher probabilities seen in RPQ. Locations where RPQ is more certain compared to R are also at locations with higher K_{dp} values.

In Fig. 2 an example for the prediction of rain exceeding 10 mm is shown. Both R and RPQ are able to accurately predict the location of the rainfall. However, compared to R, RPQ is more certain about the precipitation in the lower area of the rainfall field, which corresponds with the observed probability. These locations also have higher K_{dp} values, which can be an indication of heavy rain.

Overall, we see similar spatial patterns in the predictions for lightning and precipitation when using RPQ compared to R, but RPQ tends to give higher confidence in the predictions.

4.2 Predictor importance

The average loss of lightning and the unbiased sample standard deviation based on the test dataset are shown in Fig. 3a, indicating that incorporating polarimetric variables with the radar source improves the overall outcome. While including all

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

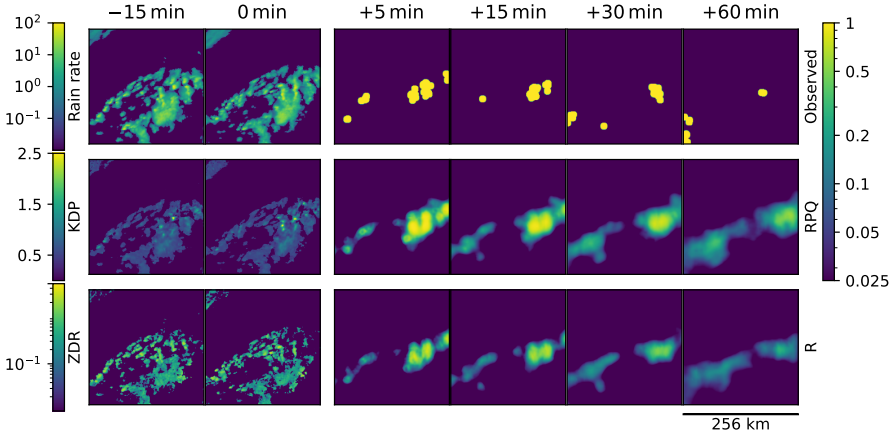

Figure 1. Results of the lightning prediction. On the left three input variables are shown (Rain rate, K_{dp} and Z_{dr}), and on the right the observed lightning occurrence and the predicted lightning probability according to the input sources RPQ and R at different lead times (indicated at the top of each column) are shown.

sources (RPQ) for the lightning model results in the highest skill, it is within the spread of RP or RQ, indicating that multiple runs are necessary to verify the robustness of the results, avoiding that coincidental convergence resulted in slightly better or worse results.

Figure 3b indicates that despite incorporating only polarimetric variables or quality indices on top of the radar source improves on average the results for hail, this does not hold when including all three sources. Reducing the results when including all sources can be an indication of redundancy, exposing the training process to overfitting. In addition, the results for hail are inconsistent, with a larger spread in the loss between the model runs compared to the results for lightning (Fig. 3a) and rain (Fig. 3c). As a result, we are unable to find a significant benefit of the polarimetric variables in predicting hail.

From the losses for heavy precipitation (Fig. 3c), it is evident that incorporating polarimetric variables benefits the results, and produces the most significant improvement compared to lightning and hail. While the pattern of the standard deviation are somewhat similar to that of lightning, the average loss between the model combinations lie more apart.

Another method to quantify the importance of the data sources is by computing the Shapley score. This was calculated for the model runs with the optimal loss score (i.e. the model with the lowest loss out of three runs). The Shapley values for all thunderstorm hazards indicate the same, that radar is the most important source, followed by polarimetric variables (Table 1). The radar source is relatively more dominant for hail compared to lightning and heavy precipitation. While previous results (Fig. 3) showed that including Q improves the results, the Shapley score indicates that the importance of Q is small. However, the effect of Q is relatively independent of R and P, whereas R and P contain redundant information, and consequently, one does not add that much over the other.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

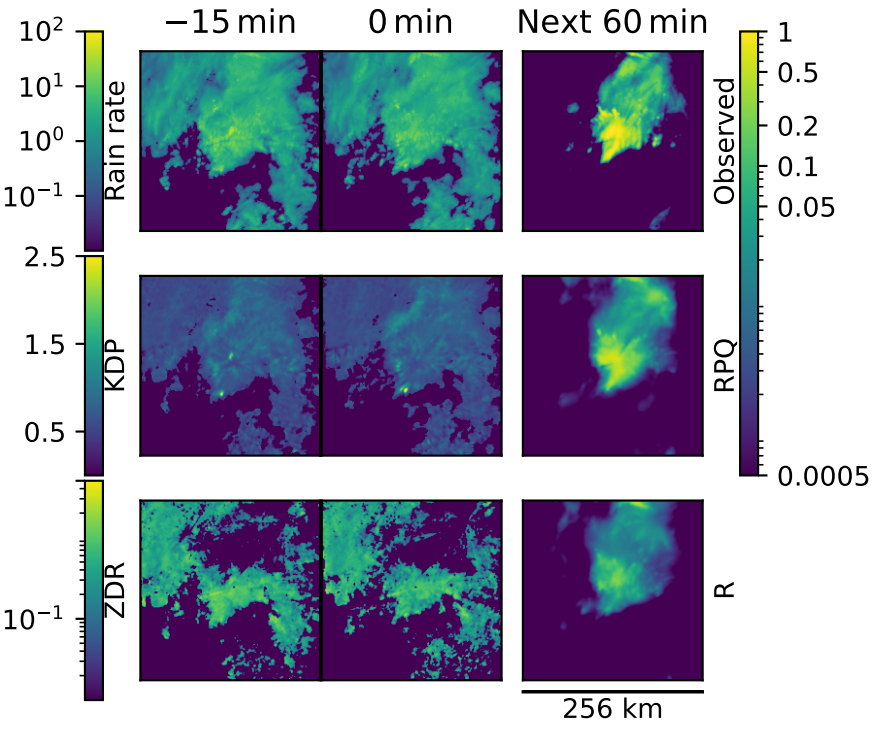

Figure 2. Same as Fig. 1, but for heavy precipitation. Only one output is shown as the precipitation is predicted as the accumulation over the next 60 min.

Table 1. Normalized Shapley values in the test dataset for the input sources (R: radar, P: polarimetric variables and Q: quality indices), and the prediction of lightning, hail and heavy precipitation.

	R	**P**	**Q**
Lightning	0.508	0.481	0.012
Hail	0.537	0.463	0.000
Precipitation	0.508	0.475	0.018

4.3 Performance of the forecasts

195 To get a more complete understanding of the skill of the model to predict the different variables, it is also important to see how it performs using other metrics. In Table 2 the average PR AUC and unbiased sample standard deviation are given. These values align with the loss, indicating that for both lightning and rain the model improves by incorporating all sources, with the largest improvement seen in precipitation. Meanwhile, for hail the RPQ model results in a slightly lower skill when including all sources, instead of the radar source alone. In addition, the least consistency is seen in the results of RPQ for hail.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

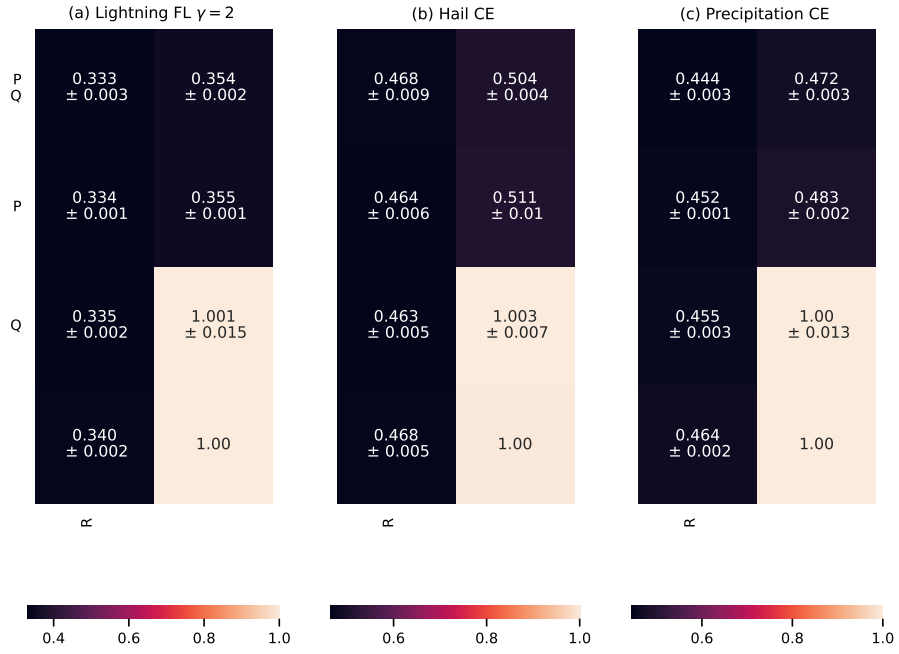

Figure 3. The average loss in the test dataset for the prediction of (a) lightning, (b) hail and (c) heavy precipitation using different combinations of the selected data sources. The mean loss and the spread of the three runs is shown here. Each panel shows a matrix where the data sources corresponding to each element can be found by combining the row and column labels. Values are normalised with the same loss value (obtained by not including inputs) from Leinonen et al. (2023) to make it comparable. With "R" indicating radar, "P" the polarimetric variables and "Q" quality indices. All loss scores are normalized, such that the baseline model (model without any input) is set to 1.

Table 2. Comparison of the average PR AUC and standard deviation of different model configurations (R: radar and RPQ: radar, polarimetric variables and quality indices) with the test set. For hail and lightning the average over all lead times is shown; for precipitation, the score is given for the accumulated precipitation in 1 hour exceeding 10 mm.

	PR AUC	
	R	RPQ
Lightning	0.626 ± 0.001	0.632 ± 0.003
Hail	0.239 ± 0.005	0.235 ± 0.010
Rain	0.466 ± 0.003	0.482 ± 0.005

200 We also investigated the effect of different thresholds and lead time on the skill of the forecasts. In Fig. 4 the CSI was calculated for different thresholds. For hail and lightning this was done for lead times of 5, 15, 30 and 60 min. With increasing lead times skill of the forecasts decrease. The decrease in skill is more gradual for lightning, while for hail values drop quickly, with having almost half of the maximum CSI (indicated value in the legend) after 15 min compared to 5 min.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

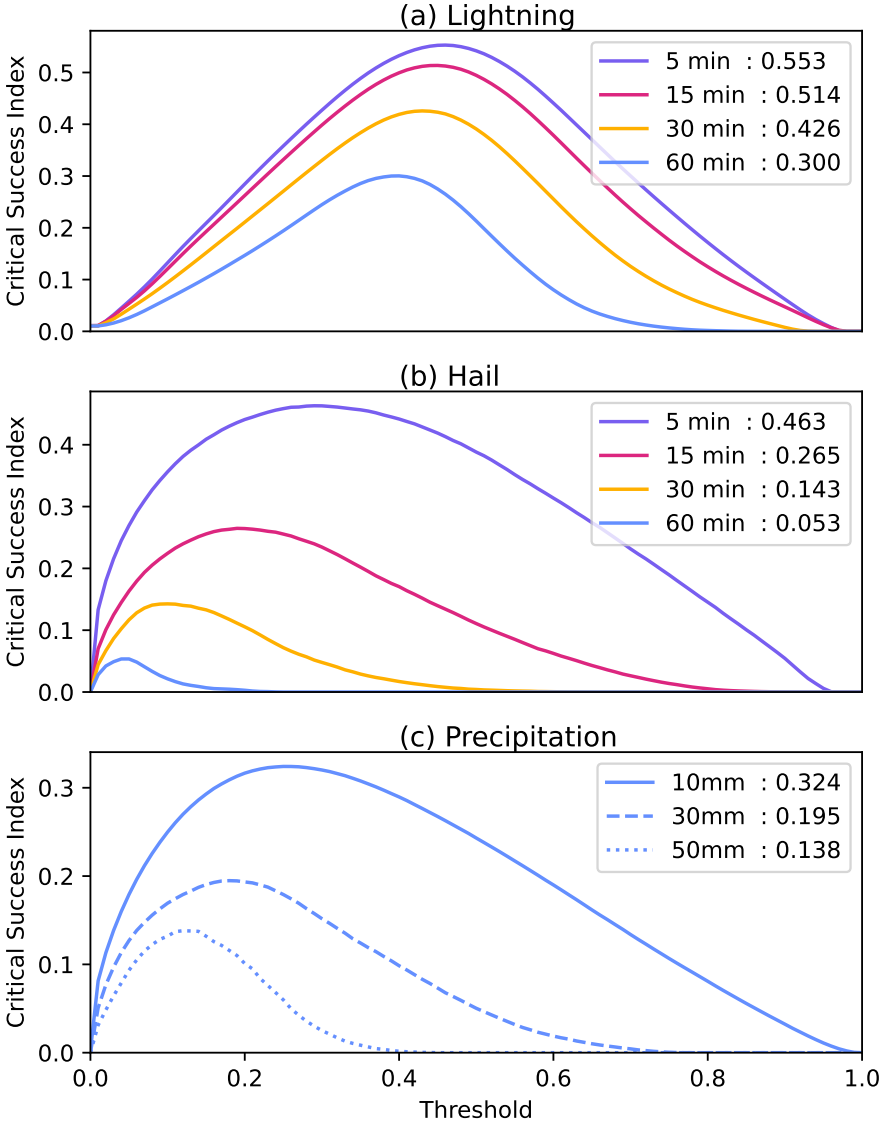

Figure 4. Critical Success Index over the test dataset at different thresholds for (a) lightning and (b) hail for different lead times and (c) the accumulated precipitation in 1 hour exceeding 10 mm, 30 mm or 50 mm, using the source combination "RPQ" (radar, polarimetric variables and quality indices). The value behind the lead time or class in the legend indicates the optimal CSI.

For heavy precipitation CSI was calculated over the accumulated precipitation in 1 hour for the three classes. Figure 4c indicated that more extreme precipitation is more difficult to nowcast. The lifetime of precipitation events decreases with higher rain rates, affecting the skill of the forecasts.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

It is also evident that the threshold resulting in the highest CSI is not fixed over the lead times (for lightning and hail) or over the classes (for precipitation). Thresholds should be decided on by the end users, selecting values that fit their desired criteria.

Lower skill for precipitation and hail than lightning can be a consequence of the time and space scales of the target variables. This difference can be enhanced due to the definition of lightning occurrence that we inherit from (Leinonen et al., 2022b). This was set to the lightning occurrence within 8 km in the last 10 min, which assigns a larger spatial and temporal footprint to the lightnings. Both PR AUC and CSI are sensitive to any degree of error, i.e. it compares the occurrence of an event pixel-wise, resulting in penalization. Matching exactly high-resolution forecasts with observed small-scale features, such as thunderstorms, is rather difficult (Ebert, 2008). For that reason the FSS is calculated over multiple scales (Fig. 5). The differences between RPQ and R are marginal, especially for shorter lead times. RPQ is slightly better for predicting lightning, with increasing differences for larger lead times (Fig. 5a), which is in line with the previous results, while for hail it we find the opposite result, i.e. R is slightly more accurate compared to RPQ and differences decrease at longer lead times (Fig. 5b). For precipitation RPQ results in a higher skill for warning levels of 10 mm and 30 mm, while R is better for warning levels of 50 mm.

5 Conclusions

The objective of this work was to evaluate the benefits from including polarimetric radar observations as an additional data source for nowcasting thunderstorm hazards, compared to exploiting radar reflectivity alone, using the convolutional-recurrent neural network from Leinonen et al. (2022b). This model can nowcast the probability of lightning and hail occurrence up to 60 min with a 5 min resolution, as well as one-hourly accumulated precipitation above pre-defined threshold levels. Qualitative differences for lightning and heavy precipitation are visible when including the data about polarimetric variables (P) and quality indices (Q) on top of radar data (R). For all three hazards, radar is the most dominant data source according to Shapley values. Incorporating polarimetric variables in addition to radar results in a higher skill for lightning and heavy precipitation predictions. In addition, quality indices that take into account quality properties of the radar reflectivity fields have a positive impact on the results. However, the differences between the data source combinations (RP, RQ and RPQ) are small for lightning, and multiple runs indicate some overlap between the loss values of these combinations. This stresses out that minor differences should not be over-interpreted, and it is important to verify the robustness of the results. For hail, the results show that different input combinations are not significantly different from each other, but the differences are rather caused by random error. Consequently, we can not conclude that the polarimetric variables improve the hail predictions in the form that they were used in this study. For that reason, we recommend to investigate further how information of polarimetric variables, such as Z_{DR}-columns, can be used for improving hail predictions.

Code and data availability. The code used in this study can be found at https://github.com/MeteoSwiss/c4dl-polar. The datasets from the radar source are available for noncommercial use at https://doi.org/10.5281/zenodo.6802292 (Leinonen et al., 2022a). The additional datasets, models and results can be found at https://doi.org/10.5281/zenodo.7760740 (Rombeek et al., 2023).

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

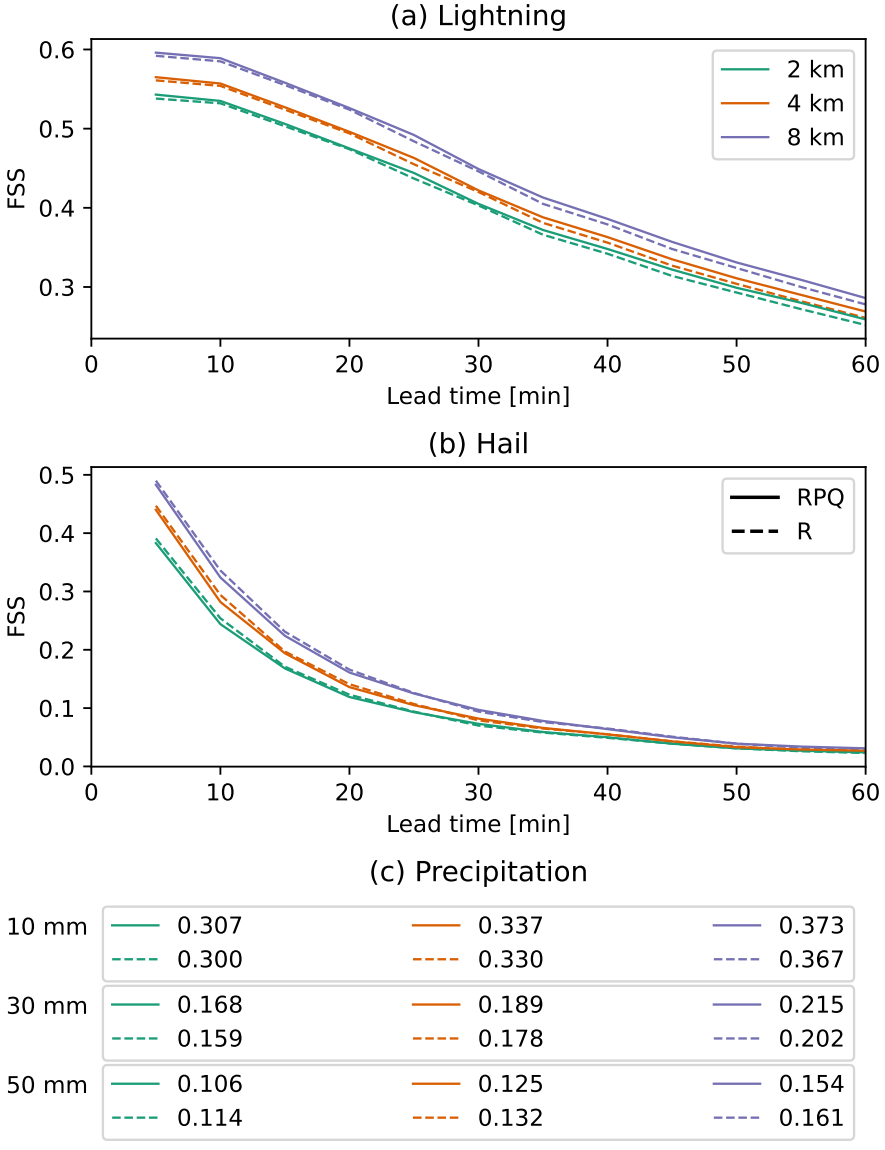

Figure 5. Fractions Skill Score (FSS) over the test dataset at different lead times for (a) lightning and (b) hail and (c) the accumulated precipitation in 1 hour exceeding 10 mm, 30 mm or 50 mm, using the source combination "RPQ" (radar, polarimetric variables and quality indices; solid lines) and "R" (radar; dashed lines).

Author contributions. All authors were involved in the design of the study. NR performed the data processing and analysis, with contributions from all co-authors. NR prepared the manuscript with contributions from all co-authors.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgements. We thank Monika Feldmann for proofreading the manuscript. JL was supported by the fellowship "Seamless Artificially Intelligent Thunderstorm Nowcasts" from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The hosting institution of this fellowship is MeteoSwiss in Switzerland.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

References

Ayzel, G., Heistermann, M., and Winterrath, T.: Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0. 1), Geoscientific Model Development, 12, 1387–1402, 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, 2016.

Ciach, G. J. and Krajewski, W. F.: On the estimation of radar rainfall error variance, Advances in water resources, 22, 585–595, 1999.

Dixon, M. and Wiener, G.: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology, Journal of atmospheric and oceanic technology, 10, 785–797, 1993.

Ebert, E. E.: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework, Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15, 51–64, 2008.

Fabry, F.: Radar meteorology: principles and practice, Cambridge University Press, 2018.

Feldmann, M., Germann, U., Gabella, M., and Berne, A.: A characterisation of Alpine mesocyclone occurrence, Weather and Climate Dynamics, 2, 1225–1244, 2021.

Figueras i Ventura, J., Pineda, N., Besic, N., Grazioli, J., Hering, A., van der Velde, O. A., Romero, D., Sunjerga, A., Mostajabi, A., Azadifar, M., et al.: Polarimetric radar characteristics of lightning initiation and propagating channels, Atmospheric Measurement Techniques, 12, 2881–2911, 2019.

Foote, G. B., Krauss, T. W., and Makitov, V.: 1.5 HAIL METRICS USING CONVENTIONAL RADAR, 2005.

Foresti, L., Reyniers, M., Seed, A., and Delobbe, L.: Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium, Hydrology and Earth System Sciences, 20, 505–527, 2016.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 132, 1669–1692, 2006.

Germann, U., Boscacci, M., Clementi, L., Gabella, M., Hering, A., Sartori, M., Sideris, I. V., and Calpini, B.: Weather radar in complex orography, Remote Sensing, 14, 503, 2022.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, Cambridge, Massachusetts, USA, http://www.deeplearningbook.org, 2016.

Guastavino, S., Piana, M., Tizzi, M., Cassola, F., Iengo, A., Sacchetti, D., Solazzo, E., and Benvenuto, F.: Prediction of severe thunderstorm events with ensemble deep learning and radar data, Scientific Reports, 12, 1–14, 2022.

Han, L., Zhao, Y., Chen, H., and Chandrasekar, V.: Advancing radar nowcasting through deep transfer learning, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–9, 2021.

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

Hoeppe, P.: Trends in weather related disasters–Consequences for insurers and society, Weather and climate extremes, 11, 70–79, 2016.

Holle, R., López, R., Ortiz, R., Paxton, C., Decker, D., and Smith, D.: The local meteorological environment of lightning casualties in central Florida, in: 17th Conference on Severe Local Storms and Conference on Atmospheric Electricity, pp. 779–784, 1993.

Imhoff, R., Brauer, C., Overeem, A., Weerts, A., and Uijlenhoet, R.: Spatial and temporal evaluation of radar rainfall nowcasting techniques on 1,533 events, Water Resources Research, 56, e2019WR026 723, 2020.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

Kumjian, M. R.: Principles and Applications of Dual-Polarization Weather Radar. Part III: Artifacts., Journal of Operational Meteorology, 1, 2013a.

Kumjian, M. R.: Principles and Applications of Dual-Polarization Weather Radar. Part I: Description of the Polarimetric Radar Variables., Journal of Operational Meteorology, 1, 2013b.

Leinonen, J., Hamann, U., and Germann, U.: Data archive for "Seamless lightning nowcasting with recurrent-convolutional deep learning", https://doi.org/10.5281/zenodo.6802292, 2022a.

Leinonen, J., Hamann, U., and Germann, U.: Seamless lightning nowcasting with recurrent-convolutional deep learning, Artificial Intelligence for the Earth Systems, 1, e220 043, 2022b.

Leinonen, J., Hamann, U., Sideris, I. V., and Germann, U.: Thunderstorm nowcasting with deep learning: a multi-hazard data fusion model, arXiv preprint arXiv:2211.01001, 2023.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P.: Focal Loss for Dense Object Detection, in: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2999–3007, https://doi.org/10.1109/ICCV.2017.324, 2017.

Lund, N. R., MacGorman, D. R., Schuur, T. J., Biggerstaff, M. I., and Rust, W. D.: Relationships between lightning location and polarimetric radar signatures in a small mesoscale convective system, Monthly Weather Review, 137, 4151–4170, 2009.

Lynn, B. and Yair, Y.: Prediction of lightning flash density with the WRF model, Advances in Geosciences, 23, 11–16, 2010.

Molnar, C.: Interpretable Machine Learning: A Guide For Making Black Box Models Explainable, Independently published, 2 edn., https://christophm.github.io/interpretable-ml-book/, 2022.

Pan, X., Lu, Y., Zhao, K., Huang, H., Wang, M., and Chen, H.: Improving Nowcasting of Convective Development by Incorporating Polarimetric Radar Variables Into a Deep-Learning Model, Geophysical Research Letters, 48, e2021GL095 302, 2021.

Pierce, C., Seed, A., Ballard, S., Simonin, D., and Li, Z.: Nowcasting, in: Doppler Radar Observations, edited by Bech, J. and Chau, J. L., chap. 4, IntechOpen, Rijeka, https://doi.org/10.5772/39054, 2012.

Poelman, D. R., Schulz, W., Diendorfer, G., and Bernardi, M.: The European lightning location system EUCLID – Part 2: Observations, 16, 607–616, https://doi.org/10.5194/nhess-16-607-2016, 2016.

Pulkkinen, S., Nerini, D., Pérez Hortal, A. A., Velasco-Forero, C., Seed, A., Germann, U., and Foresti, L.: Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1. 0), Geoscientific Model Development, 12, 4185–4219, 2019.

Rädler, A. T., Groenemeijer, P. H., Faust, E., Sausen, R., and Púčik, T.: Frequency of severe thunderstorms across Europe expected to increase in the 21st century due to rising instability, npj Climate and Atmospheric Science, 2, 1–5, 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, 2021.

Rinehart, R. E.: Radar for Meteorologists, Or, You Too Can be a Radar Meteorologist, Part III, Rinehart Publications Nevada, MO, USA, 2010.

Ritvanen, J., Harnist, B., Aldana, M., Mäkinen, T., and Pulkkinen, S.: Advection-Free Convolutional Neural Network for Convective Rainfall Nowcasting, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023.

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

Rombeek, N., Leinonen, J., and Hamann, U.: Data archive for "Exploiting radar polarimetry for nowcasting thunderstorm hazards using deep learning", https://doi.org/10.5281/zenodo.7760740, 2023.

Sachidananda, M. and Zrnic, D.: Differential propagation phase shift and rainfall rate estimation, Radio Science, 21, 235–247, 1986.

https://doi.org/10.5194/egusphere-2023-551
Preprint. Discussion started: 11 May 2023

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European lightning location system EUCLID – Part 1: Performance analysis and validation, 16, 595–605, https://doi.org/10.5194/nhess-16-595-2016, 2016.

Seliga, T. A. and Bringi, V.: Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation, Journal of Applied Meteorology and Climatology, 15, 69–76, 1976.

Shapley, L. S.: Notes on the n-Person Game – II: The Value of an n-Person Game, Tech. Rep. RM-670, The RAND Corporation, https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM670.pdf, 1951.

Sideris, I., Gabella, M., Erdin, R., and Germann, U.: Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland, Quarterly Journal of the Royal Meteorological Society, 140, 1097–1111, 2014a.

Sideris, I., Gabella, M., Sassi, M., and Germann, U.: The CombiPrecip experience: development and operation of a real-time radar-raingauge combination scheme in Switzerland, in: 2014 International Weather Radar and Hydrology Symposium, pp. 1–10, 2014b.

Sideris, I. V., Foresti, L., Nerini, D., and Germann, U.: NowPrecip: Localized precipitation nowcasting in the complex terrain of Switzerland, Quarterly Journal of the Royal Meteorological Society, 146, 1768–1800, 2020.

Simonin, D., Pierce, C., Roberts, N., Ballard, S. P., and Li, Z.: Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts, Quarterly Journal of the Royal Meteorological Society, 143, 2862–2873, 2017.

Snyder, J. C., Ryzhkov, A. V., Kumjian, M. R., Khain, A. P., and Picca, J.: A Z DR column detection algorithm to examine convective storm updrafts, Weather and Forecasting, 30, 1819–1844, 2015.

Taszarek, M., Allen, J. T., Brooks, H. E., Pilguj, N., and Czernecki, B.: Differing trends in United States and European severe thunderstorm environments in a warming climate, Bulletin of the American Meteorological society, 102, E296–E322, 2021.

Vivekanandan, J., Zrnic, D., Ellis, S., Oye, R., Ryzhkov, A., and Straka, J.: Cloud microphysics retrieval using S-band dual-polarization radar measurements, Bulletin of the american meteorological society, 80, 381–388, 1999.

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

Wilson, J. W., Crook, N. A., Mueller, C. K., Sun, J., and Dixon, M.: Nowcasting thunderstorms: A status report, Bulletin of the American Meteorological Society, 79, 2079–2100, 1998.

Wolfensberger, D., Gabella, M., Boscacci, M., Germann, U., and Berne, A.: RainForest: a random forest algorithm for quantitative precipitation estimation over Switzerland, Atmospheric Measurement Techniques, 14, 3169–3193, 2021.

Yin, J., Gao, Z., and Han, W.: Application of a Radar Echo Extrapolation-Based Deep Learning Method in Strong Convection Nowcasting, Earth and Space Science, 8, e2020EA001 621, 2021.

Zhou, K., Zheng, Y., Dong, W., and Wang, T.: A deep learning network for cloud-to-ground lightning nowcasting with multisource data, Journal of Atmospheric and Oceanic Technology, 37, 927–942, 2020.