Operational implementation of a random forest approach to derive quantitative precipitation estimations with measurements from the Swiss polarimetric radar network

Rebecca Gugerli[1,2] , Daniel Wolfensberger[2], Marco Gabella[2], Marco Boscacci[2], Urs Germann[2], Alexis Berne[1]

[1] Environmental Remote Sensing Laboratory, LTE, École Polytechnique Fédéral de Lausanne EPFL, Lausanne, Switzerland
[2] Radar,Satellite and Nowcasting Department, Federal Office of Meteorology and Climatology MeteoSwiss, Locarno-Monti, Switzerland

Motivation

Quantitative precipitation estimation (QPE) is crucial for hydrological, climatological and meteorological studies. Weather radars provide remote observations of hydrometeors with a wide-range coverage. An accurate retrieval of precipitation intensity, however, remains a challenge.

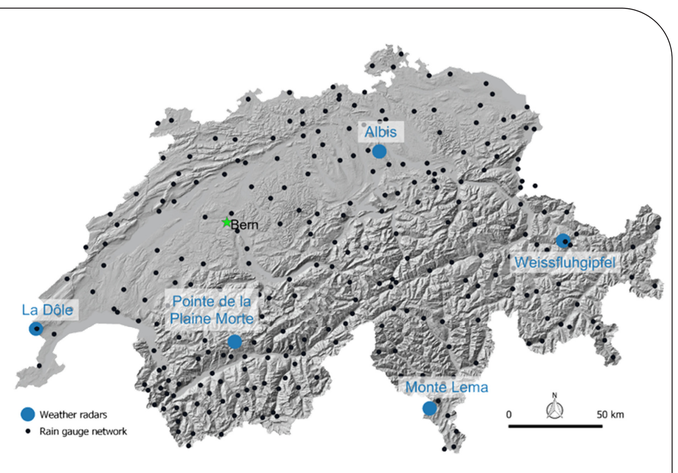
We present an approach using random forest regression to derive QPE from operational polarimetric radar observations. This approach is based on Wolfensberger et al. (2021). The model has since then been improved and implemented in an operational environment. Currently, it provides Swiss-wide QPE maps at a 5 minute resolution.

Swiss radar network

- 5 C-band dual-polarized Doppler radars
- PPI scans at 20 elevations (-0.2° to 40°)
- Radial resolution: 500 m (six range gates)
- Temporal resolution: full-volume scan every 5 minutes

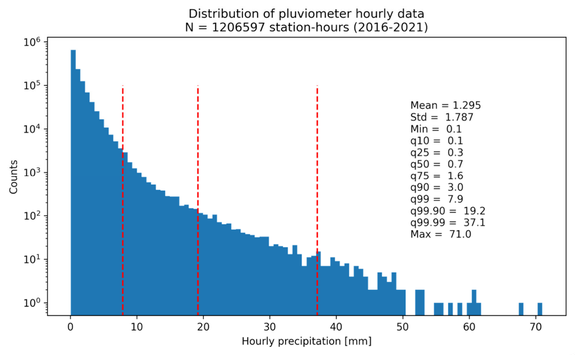
Swiss gauge network (SwissMetNet)

- 288+ automatic weather stations
- Ott Pluvio2 weighing gauges or Lambrecht tipping buckets
- Gauges heated, but mostly not wind shielded
- Temporal resolution: 10 minutes



Database

- Polarimetric radar data, station observations and COSMO-1 analysis (0°isothermal altitude).
- Temporal and spatial resolutions: 10min and 1x1 km^2
- January 2016 – December 2021



RainForest – an operational QPE model

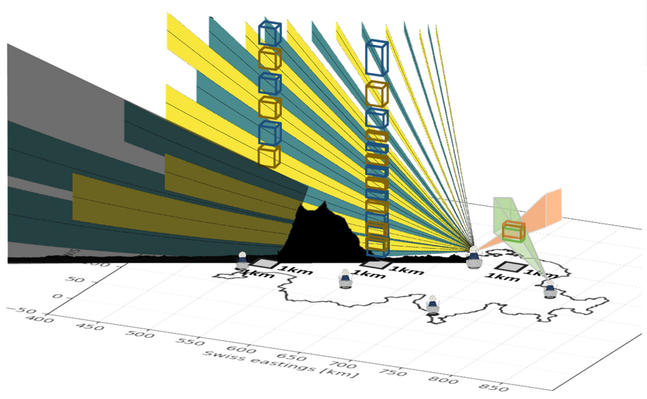
1. Data preparation: Mapping 3D polar data to a 2D Cartesian grid

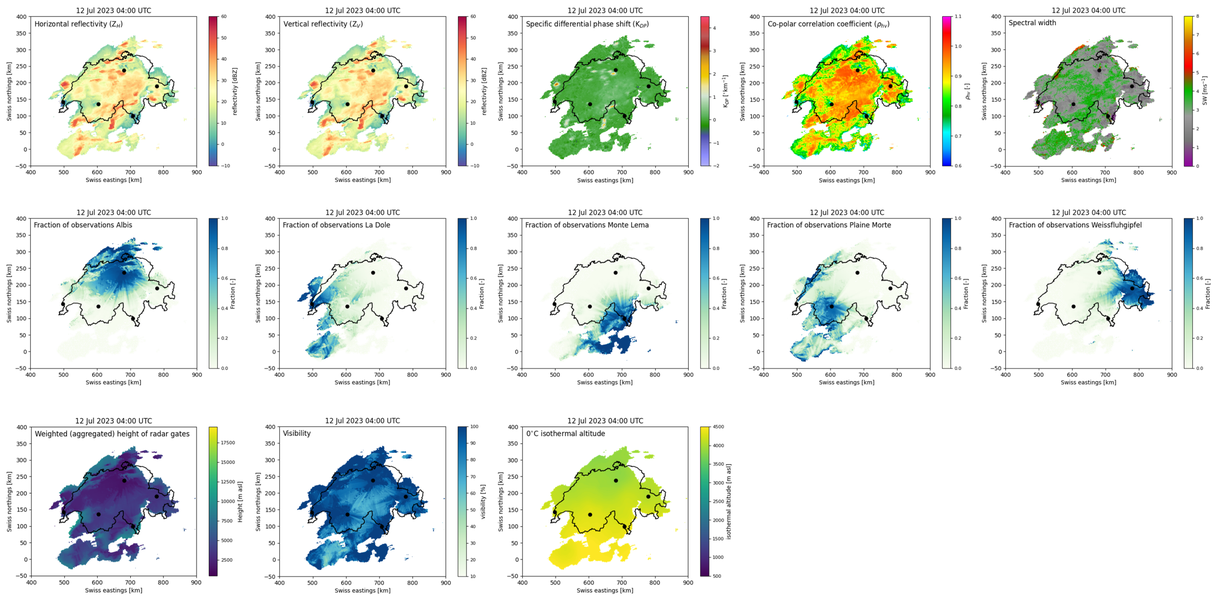
- Clutter rejection and compensation of partial beam shielding (see Germann et al., 2022)
- Mapping calibrated polarimetric data at 500m radial resolution to a Cartesian grid with 710x640 grid cells (1x1 km^2)
- Vertically weighted aggregation:

$$w(h) = 10^{-\beta h}\frac{VIS}{100}$$

- Temporal average over two radar volumes (consistency with gauge resolution):

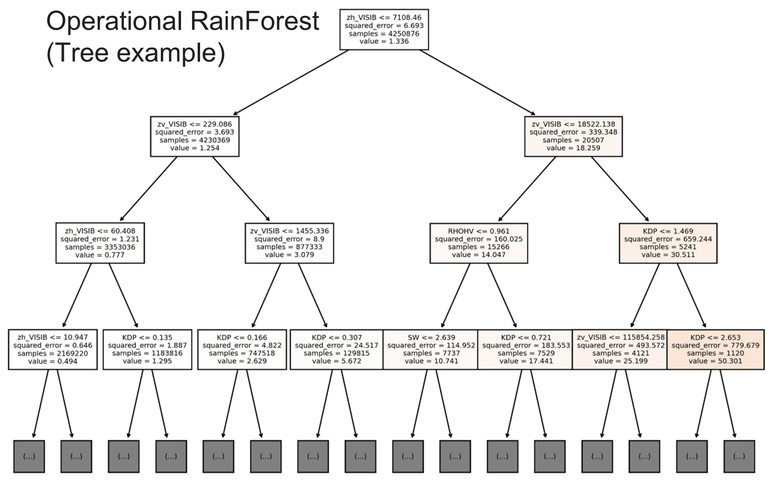
$$\overline{X_t} = 0.5\,(X_{t-5m} + X_t)$$





2. Machine Learning Model: Random forest regression model

Operational RainForest (Tree example)



Hyperparameters

Random Forest Regressor [LB2001, FP2011]:

Number of estimators (trees): 15
Number of nodes (depth): 20
Number of features randomly selected for node split: 7
Minimum number of samples in a node for a split: 2
Minimum number in a leaf to accept a split: 1

Pre- and Postprocessing:

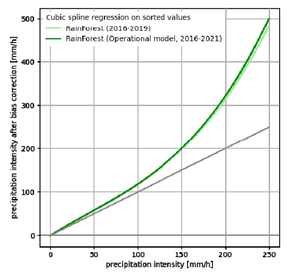
Exponent of exponential altitude weighing (β) : 0.5
Bias correction: histogram matching with cubic spline

3. Post-Processing

A) Bias correction

To avoid an overestimation of low values and an underestimation of high values, several a posteriori bias correction approaches were included in the search for the optimal hyperparameters.

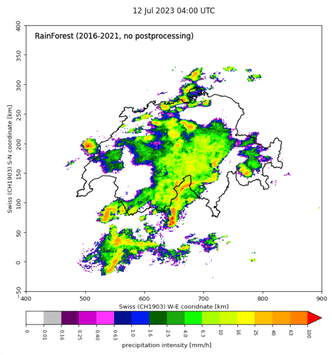
- Polynominal regression
- Histogram matching with polynominal regression
- Histogram matching with cubic spline (best performance)



B) Temporal disaggregation (10min -> 5min)

The native temporal resolution of the model is 10 minutes (limited by gauge observations). To produce QPE at a 5 min resolution, we apply an approach inspired by Joss et al. (1998).

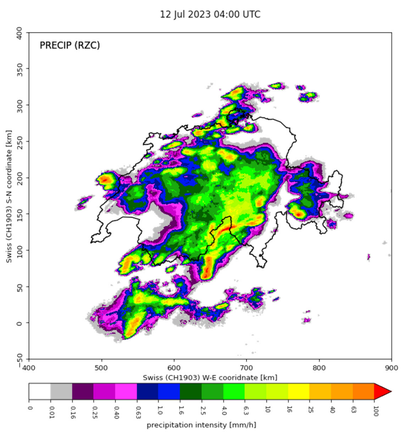
$$R_{ZR}^{(t)} = \frac{1}{316}^{\frac{2}{3}} z_{H,t}^{\frac{2}{3}} \quad \begin{cases} z_{H,t}^{(5)} \\ z_{H,t}^{(10)} = 0.5\,(z_{H,t-5m}^{(5)} + z_{H,t}^{(5)}) \end{cases} \quad R_{RF}^{(5)} = \frac{R_{ZR}^{(5)}}{R_{ZR}^{(10)}}\, RFModel(\overline{X_t})$$



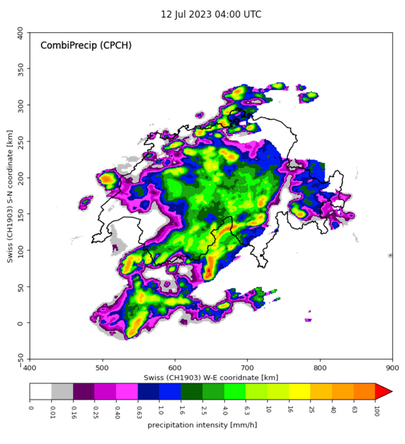
C) Local outlier removal and low-pass filtering

- A 3×3 pixel local outlier removal is applied. If individual pixels show significantly larger or lower values than their neighborhood, the values are replaced by the mean of the 3x3 pixel area.
- A 2D Gaussian filtering with standard deviation of σ=0.5 km (0.5 pixel) is applied.

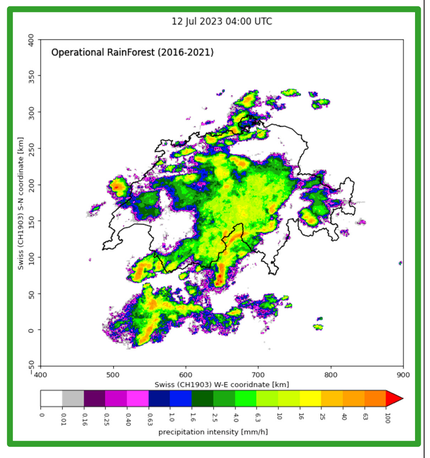
Performance of QPE's



PRECIP is the currently single-polarized radar-only operational QPE.



CombiPrecip is based on merging radar data with gauge observations.



The performance is assessed through a five-fold cross-validation on hourly precipitation.

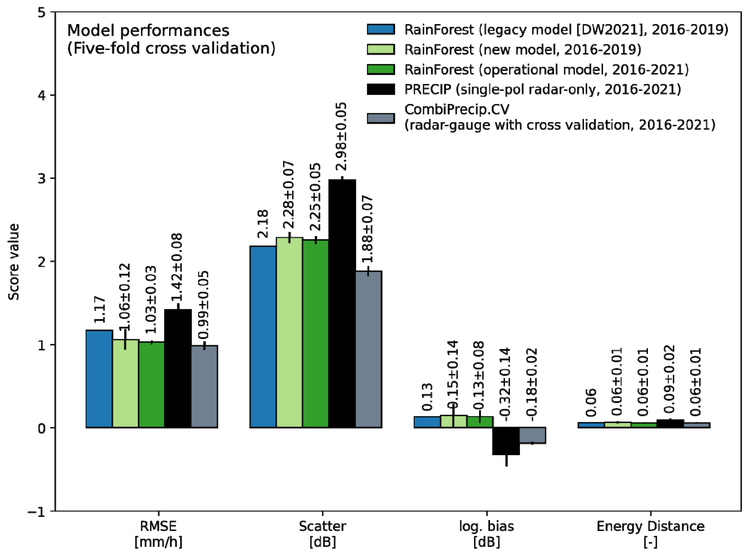
Note on the figure:

Included in the evaluation are also different model versions of RainForest:

- legacy model (Wolfensberger et al. 2021)
- the current model trained on data from 2016 to 2019 (builds a bridge between the legacy model and the operational model)
- the operationally implemented RainForest

See Germann et al. (2022) for more information on PRECIP and CombiPrecip.CV.

See Wolfensberger et al. (2021) for more information on the scores.



Conclusions

- The ML-based model RainForest outperforms the single-polarized radar-only QPE model (PRECIP) that is currently implemented at MeteoSwiss at 10- and 60-minutes precipitation accumulation.
- RainForest runs reliably in an operational environment, and provides a QPE at a 5 minute temporal and a 1x1 km^2 spatial resolution.

References

[JW1998] Joss, J. et al. (1998). Operational use of radar for precipitation measurements in Switzerland. vdf Hochschulverlag AG. ETH Zürich. Switzerland.
[LB2001] Breiman, L. (2001). Random Forests. Machine Learning. 45, 5–32
[FP2011] Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 12, 2825-2830
[DW2021] Wolfensberger, D. et al. (2021). RainForest: a random forest algorithm for quantitative precipitation estimation over Switzerland. Atmospheric Measurement Techniques. 14, 3169–3193
[RG2022] Gugerli, R. et al. (2022). Using a random forest approach to improve the quantitative precipitation estimation from the dual-polarized weather radar network in Switzerland. ERAD 2022 Book of abstracts. Locarno. Switzerland
[UG2022] Germann, U. et al. (2022) Weather radar in complex orography. Remote Sensing. 14(3), 503