RainQuest: Precipitation Estimation from Weather Radar Data Challenge

Felipe Donoso, Jonathan Hellwig.

September 2, 2024

Abstract

Accurately estimating rainfall using radar presents a significant challenge, as radars measure reflectivity rather than directly capturing precipitation. This conversion is further complicated by environmental variations that impact the relationship between reflectivity and rainfall intensity. The RainQuest hackathon aims to address this problem by developing models that integrate precise point measurements from rain gauge data with radar reflectivity, which offers high resolution spatial and temporal coverage. By combining these data sources, we strive to enhance the precision of precipitation estimates for weather forecasting. This event is organized by machine learning enthusiasts from KCDS, with support from MathSEE and Triangle.

1 Introduction

Measuring rainfall with high accuracy is crucial for meteorology, hydrology, and various environmental applications. Traditional rain gauges, while quite reliable for precise point measurements, lack the spatial coverage necessary to fully capture the variability of precipitation over large areas. On the other hand, weather radars provide high resolution spatial and temporal coverage, but they measure reflectivity (Z), which must be converted to precipitation intensity (R). The direct conversion depends on the drop size distribution, which is unknown and influenced by various environmental factors that also affect radar measurements, complicating precise precipitation estimates.

The RainQuest hackathon was conceived to tackle this complex issue by leveraging the strengths of both rain gauge data and radar reflectivity. By integrating these two data sources, we aim to develop models that significantly improve the accuracy of precipitation estimates. This effort is particularly important for enhancing short-term weather forecasting and flood prediction, which are critical for public safety and resource management.

Organized by machine learning enthusiasts from the KIT Graduate School Computational and Data Science (KCDS), with support from MathSEE and Triangle, the RainQuest hackathon brings together experts and enthusiasts from diverse fields. Participants will have the opportunity to work on real-world data, collaborate in a dynamic environment, and contribute to advancements in meteorological science.

2 Scientific background

The radar reflectivity is defined as the sum of backscattering cross-sections of small particles per cubic meter, proportional to the sixth power of the diameter of a spherical object.

In the optical regime, where the wavelength is much smaller than the objects, the cross-section corresponds to the object's area, aligning with common experience. However, in the radar regime, particularly within the Rayleigh scattering domain where the wavelength is much larger than the objects, the cross-section scales with D^6.

Reflectivity is determined by the drop size distribution $n(D)$ [5], which represents the number of drops per unit volume with diameter D. This is expressed as:

$$Z = \int_0^\infty n(D) D^6 \, dD. \tag{1}$$

Rainfall intensity is related to $n(D)$ and the terminal fall velocity $v(D)$ of raindrops by:

$$R = \frac{\pi}{6} \int_0^\infty n(D)v(D)D^3 \, dD. \tag{2}$$

Both radar reflectivity and rainfall intensity depend on the drop size distribution, indicating a mathematical link between them. However, this relationship does not lend itself to a straightforward expression. Instead, radar reflectivity and rainfall intensity are empirically related through specific functional forms that encapsulate the influence of local conditions and precipitation types. This empirical relationship [4, 6, 3, 1, 9] is commonly expressed as:

$$Z = aR^b \quad or \quad R = \left(\frac{Z}{a}\right)^{1/b} \tag{3}$$

Here, Z is the radar reflectivity factor measured in $\mathrm{mm}^6/\mathrm{m}^3$, and R is the rainfall rate in mm/h. The constants a and b are empirical parameters that vary depending on the type of precipitation, local climatic conditions and other factors.

Since radar measures reflectivity rather than direct rainfall, accurately converting Z to R requires the calibration of the constants a and b using ground-based rain gauge measurements. This calibration is essential to improve the accuracy of rainfall estimates derived from radar data, especially given the spatial and temporal limitations of rain gauges.

Different types of precipitation (stratiform, convective, pre- and post-frontal, orthographically induced) might exhibit more predictable $n(D)$ patterns, and reflectivity patterns may provide clues about the type of precipitation.

The challenge of this approach lies in developing and calibrating algorithms to derive accurate Z-R relationships for different climatic regions and precipitation types. For example, different a and b values have been proposed for various regions and conditions [7], as shown in the following studies:

- $Z = 250R^{1.2}$ for tropical convective systems in Guyana [3].
- $Z = 50R^{1.02}$ at the Upper Blue Nile Basin in Ethiopia [1].
- $Z = 41.45R^{1.90}$ wet season in South Korea [4].
- $Z = 1.17R^{2.99}$ dry season in South Korea [4].

To improve the accuracy of precipitation estimates, neural networks can be used to model the relationship between radar reflectivity and rainfall [10, 8, 6]. This approach leverages the ability of neural networks to capture complex, non-linear relationships.

The study by Zhang et al. [10] proposes using a one-dimensional convolutional neural network (1D-CNN) to estimate precipitation by merging radar reflectivity data with various meteorological factors. The neural network model, referred to as RM-1DCNN, includes multiple layers to extract features and learn the relationship between reflectivity and rainfall.

The model was trained using radar reflectivity data and meteorological factors such as temperature, relative humidity, air pressure, and wind speed. The input data were processed as three-dimensional tensors, and the output was the estimated precipitation rate.

The performance of the RM-1DCNN model was evaluated against traditional methods such as Ordinary Kriging interpolation [2], Z-R relationships, and Back Propagation Neural Networks (BPNN). The RM-1DCNN model showed significant improvement in accuracy, with lower root mean square error (RMSE) and higher correlation coefficients.

The study by Neuper and Ehret [6] presents an information-theoretic framework to integrate any kind of data to the problem of quantitative precipitation estimation (QPE). Information theory plays a crucial role in this study by enhancing the neural network's ability to handle the complexity and uncertainty inherent in the relationship between radar reflectivity and rainfall.These are important key aspects to be considered:

1. Entropy and Mutual Information: Entropy measures the uncertainty in the data, while mutual information quantifies the amount of information shared between variables (e.g., between radar reflectivity

and rainfall). By maximizing mutual information, the neural network can better capture the relevant features that influence rainfall estimation.

2. Feature Selection: Information-theoretic methods help in selecting the most informative features from the radar and meteorological data. This reduces the dimensionality of the input data and improves the efficiency and accuracy of the neural network model.
3. Data-driven Approach: The use of information theory allows for a data-driven approach to model the complex, nonlinear relationship between radar reflectivity and rainfall. This approach is more flexible and adaptive compared to traditional empirical methods, which rely on fixed parameters.

The integration of neural networks and information theory provides a robust method for estimating rainfall from radar reflectivity.

3 Objective

The task is to develop a model to estimate rainfall intensity (R) from radar reflectivity (Z), leveraging a composite dataset from the German Weather Service (DWD). The dataset includes radar reflectivity on a 1 km² grid over Germany at 5-minute intervals and ground truth measurements from rain gauges with 1-minute resolution from several hundred stations. The model might consider:

- Variations in $n(D)$ due to different precipitation types and environmental conditions.
- Temporal factors such as time of day and seasonality.
- Any other innovative features or methods that could improve the estimation accuracy.

One of the key objectives of the hackathon is to ensure that participants can clearly articulate their methodology and findings in scientific terms. This involves not only developing a model to estimate rainfall intensity from radar reflectivity but also effectively communicating the underlying reasoning, processes, and results.

4 Data

1. Radar Reflectivity (Z): Available on a 1 km² grid over Germany, updated every 5 minutes.

   It is important to note that radar systems typically store reflectivity values in a logarithmic scale rather than in a linear form. Specifically, the stored values are in decibels relative to Z, or dBZ, which is calculated as:

   $$dBZ = 10 \cdot \log_{10}\left(\frac{Z}{1\ \mathrm{mm}^6/\mathrm{m}^3}\right)$$

   where Z represents the linear reflectivity in units of $\mathrm{mm}^6/\mathrm{m}^3$. For instance, a reflectivity of 1000 $\mathrm{mm}^6/\mathrm{m}^3$ corresponds to 30 dBZ, and a reflectivity of 2000 $\mathrm{mm}^6/\mathrm{m}^3$ corresponds to 33 dBZ. Notably, an increase of 3 dB indicates a doubling of the reflectivity value.
2. Ground Truth Rainfall Intensity (R): High-resolution (1-minute) data from rain gauges at various stations across Germany.
3. Temporal Data:Time of day and day of the year to account for diurnal and seasonal variations in precipitation patterns.

5 Evaluation Criteria

1. Goodness-of-Fit measures: How closely the model's estimations match the ground truth data.
2. Science communication: The participants' ability to articulate and justify the methodology used to solve the problem.

6 Acknowledgements

We would like to express our sincere gratitude to the Institute of Meteorology and Climate Research (IMK-TRO) for their invaluable technical support. Additionally, we extend our acknowledgment to the German Weather Service (DWD) for providing the essential data that made this challenge possible.

References

[1] Megbar W Birhan, U Jaya Prakasha Raju, and Samuel T Kenea. "Estimation of Rainfall Intensity From First Observation Weather Radar Reflectivity Data Over Upper Blue Nile Basin, Ethiopia". In: *Transactions on Science and Technology* 5(4) (2018), pp. 223–232.

[2] Thaís Da Silva Charles et al. "Estimating Average Annual Rainfall by Ordinary Kriging and TRMM Precipitation Products in Midwestern Brazil". In: *Journal of South American Earth Sciences* 118 (Oct. 2022), p. 103937. ISSN: 08959811. DOI: 10.1016/j.jsames.2022.103937. URL: https://linkinghub.elsevier.com/retrieve/pii/S0895981122002267 (visited on 06/07/2024).

[3] Komalchand Dhiram and Zhenhui Wang. "Evaluation on Radar Reflectivity-Rainfall Rate (Z-R) Relationships for Guyana". In: *Atmospheric and Climate Sciences* 06.04 (2016), pp. 489–499. ISSN: 2160-0414, 2160-0422. DOI: 10.4236/acs.2016.64039. URL: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/acs.2016.64039 (visited on 05/21/2024).

[4] Tae-Jeong Kim, Hyun-Han Kwon, and Kue Bum Kim. "Calibration of the Reflectivity-Rainfall Rate (Z-R) Relationship Using Long-Term Radar Reflectivity Factor over the Entire South Korea Region in a Bayesian Perspective". In: *Journal of Hydrology* 593 (Feb. 2021), p. 125790. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2020.125790. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022169420312518 (visited on 05/21/2024).

[5] J. S. Marshall and W. Mc K. Palmer. "THE DISTRIBUTION OF RAINDROPS WITH SIZE". In: *Journal of Meteorology* 5.4 (Aug. 1948), pp. 165–166. ISSN: 0095-9634, 0095-9634. DOI: 10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2. URL: http://journals.ametsoc.org/doi/10.1175/1520-0469(1948)005%3C0165:TDORWS%3E2.0.CO;2 (visited on 06/11/2024).

[6] Malte Neuper and Uwe Ehret. "Quantitative Precipitation Estimation with Weather Radar Using a Data- and Information-Based Approach". In: *Hydrology and Earth System Sciences* 23.9 (Sept. 16, 2019), pp. 3711–3733. ISSN: 1607-7938. DOI: 10.5194/hess-23-3711-2019. URL: https://hess.copernicus.org/articles/23/3711/2019/ (visited on 05/21/2024).

[7] "Radar Observation of the Atmosphere. L. J. Battan (The University of Chicago Press) 1973. PP X, 324; 125 Figures, 21 Tables. £7·15". In: *Quarterly Journal of the Royal Meteorological Society* 99.422 (Oct. 1973), pp. 793–793. ISSN: 0035-9009, 1477-870X. DOI: 10.1002/qj.49709942229. URL: https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.49709942229 (visited on 06/11/2024).

[8] Stephan Rasp and Sebastian Lerch. "Neural Networks for Postprocessing Ensemble Weather Forecasts". In: *Monthly Weather Review* 146.11 (Nov. 1, 2018), pp. 3885–3900. ISSN: 0027-0644, 1520-0493. DOI: 10.1175/MWR-D-18-0187.1. URL: http://journals.ametsoc.org/doi/10.1175/MWR-D-18-0187.1 (visited on 01/26/2024).

[9] Marc Schleiss et al. "The Accuracy of Weather Radar in Heavy Rain: A Comparative Study for Denmark, the Netherlands, Finland and Sweden". In: *Hydrology and Earth System Sciences* 24.6 (June 19, 2020), pp. 3157–3188. ISSN: 1607-7938. DOI: 10.5194/hess-24-3157-2020. URL: https://hess.copernicus.org/articles/24/3157/2020/ (visited on 05/21/2024).

[10] Yonghong Zhang et al. "Radar Reflectivity and Meteorological Factors Merging-Based Precipitation Estimation Neural Network". In: *Earth and Space Science* 8.10 (Oct. 2021), e2021EA001811. ISSN: 2333-5084, 2333-5084. DOI: 10.1029/2021EA001811. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021EA001811 (visited on 05/21/2024).