Title: Improving radar-based severe precipitation nowcasting with machine learning methods

Allocation Period: 01.01.2024 – 31.12.2024

Principal Investigator: Boran Frank (s5136492@stud.uni-frankfurt.de)

Reliable nowcasts of heavy convective rainfall events during summer become more and more important, as they are expected to intensify in a changing global climate and thus may cause urban pluvial floods with severe damages and even human losses. Due to their local spatial and short temporal scales, the nowcast of such events is quite challenging. In current nowcasting systems, observed radar images are typically simply extrapolated into the future by the Lagrangian persistence. In reality, convective storms are highly dynamic systems which undergo initiation, growth, translation, and decay processes on very short time ranges. Promising methods for properly nowcasting these dynamic systems are machine learning (ML) based approaches, which use historical radar observations as training data to predict rainfall intensity up to two hours ahead. Up to now, most ML based rainfall nowcasts exclusively use 2D radar images as input data (e.g., Shi et al. 2015, Ravuri et al. 2021, Ritvanen et al. 2023), and thus do not include any information about the stability of the atmosphere, the atmospheric moisture content, and the initiation of clouds.

In this project, which is part of the Master's thesis of Boran Frank at the Insitute for Atmospheric and Environmental Sciences Frankfurt, we aim to develop a novel ML based nowcast system for strong convective rainfall events. The project is supported by PD Dr. Bernhard Thomas (https://zoologie.uni-koeln.de/arbeitsgruppen/habilitierte/mitarbeiterinnen/pd-dr-bernd-thomas), who is an expert for Machine Learning and Artificial Intelligence, and by TV meteorologist Karsten Schwanke, who is an expert for synoptic meteorology and extreme weather events. In the project, two different ML architectures will be tested: (i) variational autoencoder (Pinheiro Cinelli et al. 2021), and (ii) encoder-decoder Convolutional Neural Networks (Badrinarayanan et al. 2017). Further, in addition to radar data, we will use satellite, lightning, and station data as well as NWP model data as input for the neural networks. Radar data is available from the German Weather Service (DWD). To save computing time and disk space, the approach will be trained for a single radar site (Essen or Offenthal; see https://www.dwd.de/DE/derdwd/mess-netz/atmosphaerenbeobachtung/_functions/Teasergroup/radarverbund_teaser5.html). As satellite data we will use High Rate SEVIRI infrared images from EUMETSAT. Lightning data is measured at 160 stations across Europe and is available from 1991 onwards. Station observations at hundreds of sites in Germany are also provided by the DWD. NWP model data from ICON-EU and ICON-D2 is used to derive stability and moisture quantities, like e.g. CAPE and integrated water vapour, which are essential to predict the life cycle of convective storms. Altogether, we plan to use data of 10 summer half-years (2014-2023), splitted into a training dataset (80%) and a test dataset (20%). The above mentioned ML architectures will be designed with components of the Python packages scikit-learn and KERAS.

Due to the large amount of input data, the training of the neural networks is expected to be the most time-consuming part of the project. An assessment for the required computing time is given in Ayzel et al. (2020), who uses a Convolutional Neural Network and 7 years of radar data, and found that the average training time is 72-76 hours on a single GPU. As we aim to test different setups of the neural networks and different combinations of the datasets as input, it is worthwhile to speed up the training procedure, which can be achieved by parallelising the ML architecture. Hence, we think the HPC architecture of the DKRZ is particularly suited to reach the goals for our project.

References:

Ayzel et al. (2020): RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting. Geosci. Model Dev., 13, 2631–2644. https://doi.org/10.5194/gmd-13-2631-2020.

Badrinarayanan et al. (2017). “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation.” In: IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12). doi:10.1109/TPAMI.2016.2644615

Pinheiro Cinelli et al. (2021). Variational Autoencoder. Variational Methods for Machine Learning with Applications to Deep Networks. Springer. pp. 111–149. doi:10.1007/978-3-030-70679-1_5

Ravuri, S., Lenc, K., Willson, M. et al. (2012): Skilful precipitation nowcasting using deep generative models of radar. Nature 597, 672–677. doi:10.1038/s41586-021-03854-z

Ritvanen et al. (2023): Advection-Free Convolutional Neural Network for Convective Rainfall Nowcasting. IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, 16. doi:10.1109/JSTARS.2023.3238016

Shi, X. et al. (2015): Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In Proceedings of the 28th International Conference on Neural Information Processing Systems Vol. 1, 802–810 (MIT Press, 2015).