Project: **1546**

Project title: **Groundwater Time Series Prediction using Deep Learning**

Principal investigator: **Annette Rudolph**

Report period: **2026-01-01 to 2026-10-31**

Summary

In 2026, project 1546 developed and benchmarked deep learning (DL) models for groundwater level (GWL) prediction in Brandenburg (WP 1) and investigated atmospheric drivers of extreme precipitation: atmospheric blocking, large-scale weather patterns and convection-resolving CM1 simulations (WP 2). The work led to three conference contributions and a manuscript about to be submitted; in 2027 the project moves to regional LSTM models for Berlin–Brandenburg.

1 Use of resources

Resource	Requested	Granted	Used, Jan–Oct 2026
Levante GPU [node hours]	5000	2500	59 (590 A100 GPU-hours)
Levante CPU [node hours]	5120	2048	454 (342 for weather regimes)
Levante storage [TiB]	1	1	0.79
Archive [TiB]	0	0	0

Usage stayed well below the grant for four reasons:

1) **Unit conversion in the 2026 request.** We estimated the GPU and CPU demand in single-GPU hours and CPU core-hours but entered these numbers as node hours, which considerably overestimated our demand.
2) **CM1 experiments not carried out.** CM1 test runs for Brandenburg cost about 0.3 node hours per case instead of the 100 CPU-hours we had assumed. More importantly, the tests showed that CM1 was less helpful for our research question than anticipated. We therefore pursued extreme precipitation through weather pattern analysis instead, and did not run the planned 20-case experiments.
3) **Wall-time limits.** For the weather-pattern recognition and clustering some preprocessing jobs ran as long single jobs that exceeded Levante's maximum wall time. We therefore did part of the preprocessing on Levante (342 CPU node hours, three quarters of our CPU use) and ran the rest of the pattern analysis on the TU Berlin HPC cluster. However, that cluster has limited capacity and has had frequent, long outages, so it is not a reliable basis for 2027. We will instead split long analyses into shorter, checkpointed job chains that run on Levante.
4) **Priorities and postponed work.** In 2026 we concentrated on WP 1 and 2 and on preparing the regional Deep Learning model for full Berlin–Brandenburg, mainly data harmonisation. The weather-pattern and extremes analyses needed more development and reprocessing than anticipated. Phases 3 and 4 of the 2026 plan (adding extreme-event features to the DL models; scenario projections in WP 3) were therefore postponed to 2027.

The 2027 request is based on the costs measured in 2026 and is substantially smaller. Until the end of the allocation period, we will use further GPU node hours for first tests of the regional transfer-learning LSTM planned for 2027; a large part of the 2026 grant will nevertheless remain unused. The following sections briefly summarize the results achieved in WP 1 and WP 2 and the preparations for the 2027 work.

2 WP 1 – Deep learning for groundwater level prediction

We built a containerized (Singularity) training, evaluation and explanation pipeline on the Levante GPU partition. On 363 Brandenburg monitoring wells (weekly data, 1991–2024; climate forcing, soil moisture and static site features), we compared convolutional and recurrent models (CNN, CNN-LSTM) with three Transformer architectures (Vanilla, PatchTST and Autoformer). Each model was trained as a 10-member ensemble on one A100. Model variants covered input scaling variants, well embeddings and FiLM conditioning, adjusted training periods and Optuna hyperparameter searches. SHAP values and

permutation importance were computed for all models.

Model	Time per ensemble member (1 × A100)	GPU-hours in 2026
CNN–LSTM	3.5–4.3 h	234
Vanilla Transformer	0.4–0.9 h	136
Autoformer	0.25–0.5 h	110
PatchTST	1.0–1.7 h	78
Cluster-wise LSTM, BB–BE (tests)	0.7–1.5 h (62–105 wells)	30

Results. The CNN–LSTM performed clearly best (median test $R^2 = 0.76$, NSE = 0.60), followed by the vanilla Transformer ($R^2 = 0.68$, NSE = 0.29), PatchTST ($R^2 = 0.63$, NSE = 0.52) and Autoformer ($R^2 = 0.62$, NSE = 0.43); a pure CNN reached $R^2 = 0.59$. In a sensitivity analysis with three precipitation scenarios (persistent drought, –20 % and +20 % precipitation in the 2019–2024 test period), the CNN–LSTM responded only weakly (normalized sensitivity index ≤ 0.01), and groundwater levels even rose under reduced precipitation. This shows that extreme conditions are not yet well represented, which motivates the regional and extreme-event work planned for 2027. These results were presented at the Tag der Hydrologie 2026 in Kassel.

3 WP 2 – Atmospheric drivers of extreme precipitation

We first computed atmospheric blocking indices from ERA5 data (1979–2025) following Detring et al. (2021) (21 CPU node hours). Blocking alone proved too restrictive and not informative enough to explain extreme precipitation, so we extended the analysis to large-scale weather regimes. We detected extreme convective precipitation events in ERA5 (71 CPU node hours) and preprocessed ERA5 fields for nested domains over Europe and North America on Levante (1979–2025; 342 CPU node hours). In stage 1, Z500 anomalies were clustered (PCA, Ward clustering) into 7 circulation regimes for Europe and 5 for North America. In stage 2, logistic regression on additional atmospheric fields (e.g. CAPE, moisture and wind) split each regime into a high-risk and a low-risk sub-regime for extreme precipitation. CM1 was compiled on Levante and tested with idealized cases and 18 convection-resolving Brandenburg runs (8 CPU node hours). The CM1 simulations offered less additional insight than expected, so we prioritized the weather-regime approach, which proved more informative for linking large-scale circulation to extreme precipitation.

Results. Z500 regimes alone explained the occurrence of extreme precipitation only weakly (AUC ≈ 0.5–0.6). A small, regime-specific set of additional fields raised the AUC above 0.8, and almost every regime in both regions contained a significantly wetter high-risk sub-regime, so the method transfers between regions. This work led to two conference contributions and a manuscript (Section 5).

4 Preparation of the 2027 work

We harmonized groundwater data from the Brandenburg Landesamt für Umwelt (1,923 wells) and the Berlin Senate (1,024 wells). After quality control, gap filling and resampling to weekly steps, 2,075 wells have complete climate, soil moisture and static inputs. First LSTM runs were tested on Levante (30 GPU-hours).

5 Publications and presentations using project 1546

- Eckert, M.-C. and Rudolph, A. (2026): Deep Learning für regionale Grundwasserprognosen unter Extremereignissen: Erkenntnisse aus Brandenburg. Poster, Tag der Hydrologie 2026, Kassel. https://www.researchgate.net/publication/401622860
- Eckert, M.-C., Rudolph, A. and Schielicke, L. (2026): From circulation regimes to convective precipitation extremes: a multi-scale perspective. Talk, CMOS Congress 2026 (online), 1 June 2026; oral presentation, 8th Int. Symposium on Computational Wind Engineering (CWE 2026), London, Ontario, 7–11 June 2026.
- Eckert, M.-C., Rudolph, A. and Schielicke, L. (in prep.): On the scale-dependence of weather regimes describing extreme precipitation events. To be submitted to Weather and Climate Dynamics.