

EGU26-10856, updated on 01 Jul 2026

<https://doi.org/10.5194/egusphere-egu26-10856>

EGU General Assembly 2026

© Author(s) 2026. This work is distributed under the Creative Commons Attribution 4.0 License.



Calibration of a Long Short-Term Memory (LSTM) rainfall-runoff model using remote sensing soil water content estimations

Tibor Rapai¹, Petra Baják², István Gábor Hatvani⁴, András Lukács³, and Balázs Székely¹

¹ELTE Eötvös Loránd University, Faculty of Science, Institute of Geography and Earth Sciences, Department of Geophysics and Space Science, Budapest, Hungary

²Università degli Studi di Milano, Department of Earth Sciences "Aldo Moro" (DiSTAD), Milan 20133, Italy

³ELTE Eötvös Loránd University, Institute of Mathematics, Department of Computer Science, Budapest, Hungary

⁴HUN-REN Research Centre for Astronomy and Earth Sciences, Institute for Geological and Geochemical Research, Budapest, Hungary

Long Short-Term Memory (LSTM) neural networks have proven their excellence in basin-level discharge prediction, provided there is an adequate amount of high-quality time series data available for training, including meteorological forcings and streamflow gauge measurements. Such data-driven black-box models can successfully learn the complex behavior of delayed hydraulic responses; however, they cannot yet be easily applied in water management practice, and model transfer attempts to ungauged catchments have not been entirely successful.

In our previous work, we explored an approach to characterizing near-surface flow regimes, starting from a full catchment model and then applying a single LSTM network layer within a semi-distributed subbasin setup reflecting downstream topography. Application to the Tarna River catchment area in Hungary (2,116 km²) showed that transfer learning from the full catchment model (achieving an NSE of 0.91 on the training set and 0.66 on an independent test set) to a downstream chain of gauged Hydrological Response Units (HRUs) is a powerful tool for investigating a semi-distributed HRU network. The entire setup, however, involves a much higher level of complexity, and the available detailed meteorological data and gauge measurements in only two-thirds of the subbasins did not provide sufficient information for the single LSTM model to fully predict the HRU network processes.

Because these models apply “virtual water amounts” stored in the hidden cells of the LSTM network for discharge estimation, their internal variables lack direct physical interpretability. *In the present research, we investigate how data fusion during calibration with Gravity Recovery and Climate Experiment (GRACE) data, downscaled using soil water content and evapotranspiration products from the ECMWF Reanalysis (ERA5) database, can improve predictive performance, and help to verify our working hypothesis regarding the theoretical connection between Near Surface Water Content (NSWC) and LSTM cell states.*

These results can also validate interpretations derived from our model concerning baseflow contributions and recharge-discharge classification of subbasins, while promising realistic

transferability of the pre-trained lumped catchment model to all subbasins and broader general applicability of the proposed method. We hypothesize that the daily change dynamics of Terrestrial Water Storage (TWS) and NSW – the latter playing a decisive role in gravitational flows within the Critical Zone – are strongly correlated.

Accordingly, we propose using downscaled TSW estimates to (1) introduce a new term into the loss function based on our working hypothesis relating median LSTM cell state values to the normalized dynamics of NSW, and (2) add a new input dimension approximating total runoff as precipitation minus evapotranspiration and infiltration.

Furthermore, the current model extension, still based on 0.1 ° gridded input data, prepares the ground for future developments that incorporate high-spatial-resolution satellite remote sensing data, such as Sentinel-2 NDWI, to support local-scale hydrological applications efficiently. Integrating satellite data products with different temporal and spatial resolutions is not a straightforward calibration step for rainfall-runoff models, as pixel-wise normalization of measurements requires complex physically based geostatistical methods compatible with model logic to avoid performance deterioration.