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

Tibor Rapai[1], **Petra Baják**[2], **István Gábor Hatvani**[4], **András Lukács**[3] and **Balázs Székely**[1]

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

[2]Università degli Studi di Milano, Department of Earth Sciences "Ardito Desio" (DiSTAD), Milan 20133, Italy

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

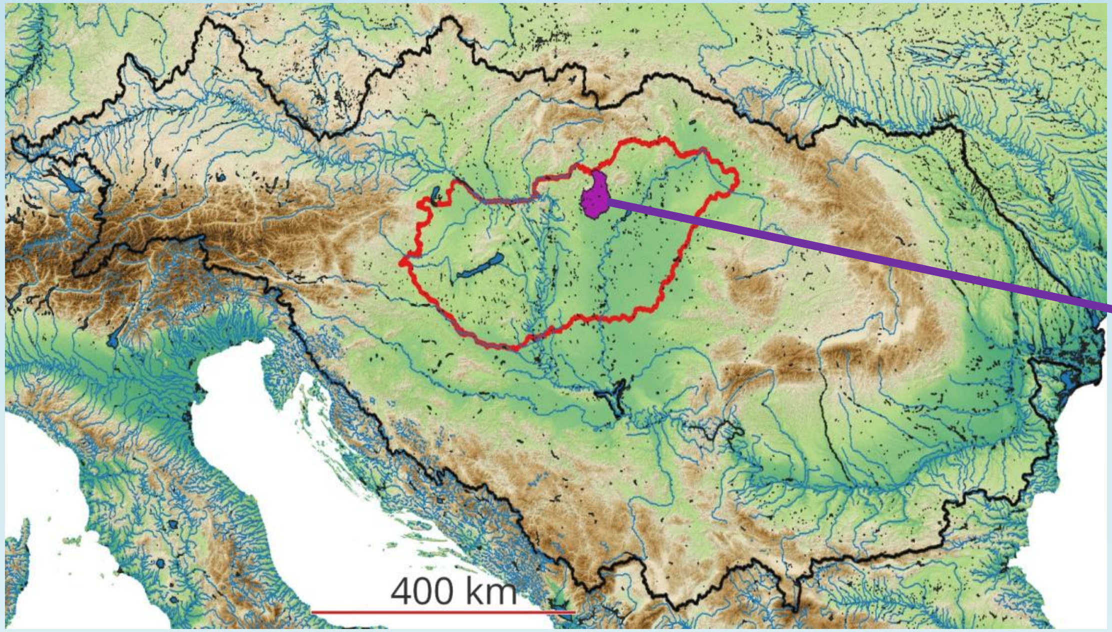
[4]HUN-REN Research Centre for Astronomy and Earth Sciences, Institute for Geological and Geochemical Research, Budapest, Hungary

Acknowledgements: *BSz received travel support from journal Remote Sensing of publisher MDPI Basel.*

Physical landscape – starting points

Topography is the main driver not only of surface runoff but also of interflow/baseflow!

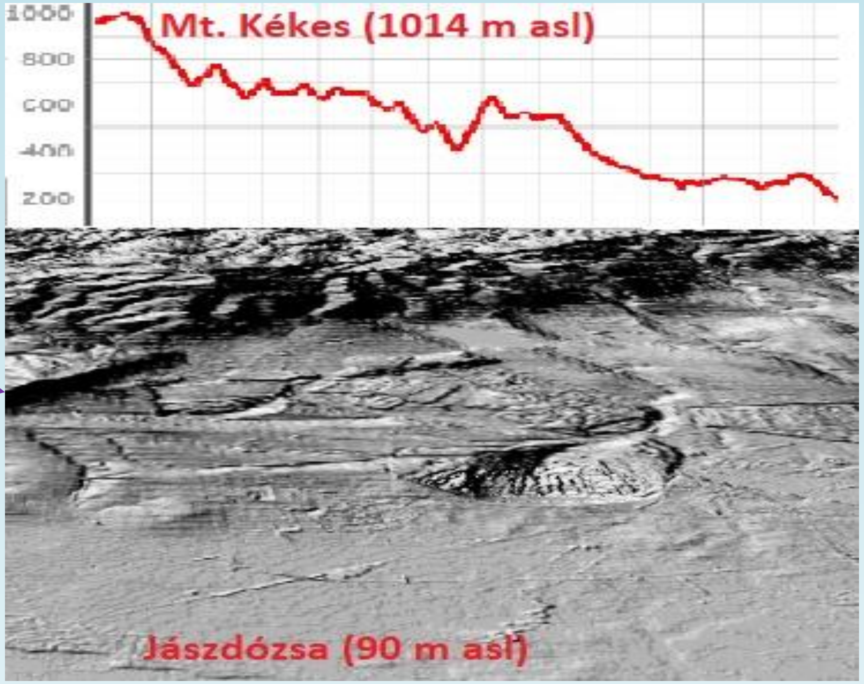
JIANG, X-W. et al. (2009): hydraulic conductivity decreases exponentially with depth -> **remote sensing may reveal near-surface water content patterns**!



Tarna River catchment Hungary

Inside the Danube basin

source: OpenStreetMap



Research goal: could we construct a data-driven emulation model for hydrological characterization of a given Region Of Interest?

Understanding droughts – Landscape of models

How to predict it

or at least understand?

Source: local media

Classical **hydrological** (SWAT, HEC-HMS) and new generation (eg. HydroGeoSphere, STE) **physical simulation systems** are powerful tools in an expert hand …

Recent hybrid systems have the highest predictive performance, but these still need manual parameterisation and substantial teamwork to develop …

Time series are **best understood by Long Short-Term Memory models**, but how to make these results explainable? *Hint*: soil moisture probe, in *Lees, T. et al. (2022)*

Full catchment model – input data sources

Meteorological forcings: 365-days sequences of min. and max. temperatures, precipitation, length of day, radiation, vapor pressure deficit +

ERA5 Land: Potential EvapoTranspiration and Soil Water Content (0-28 cm)

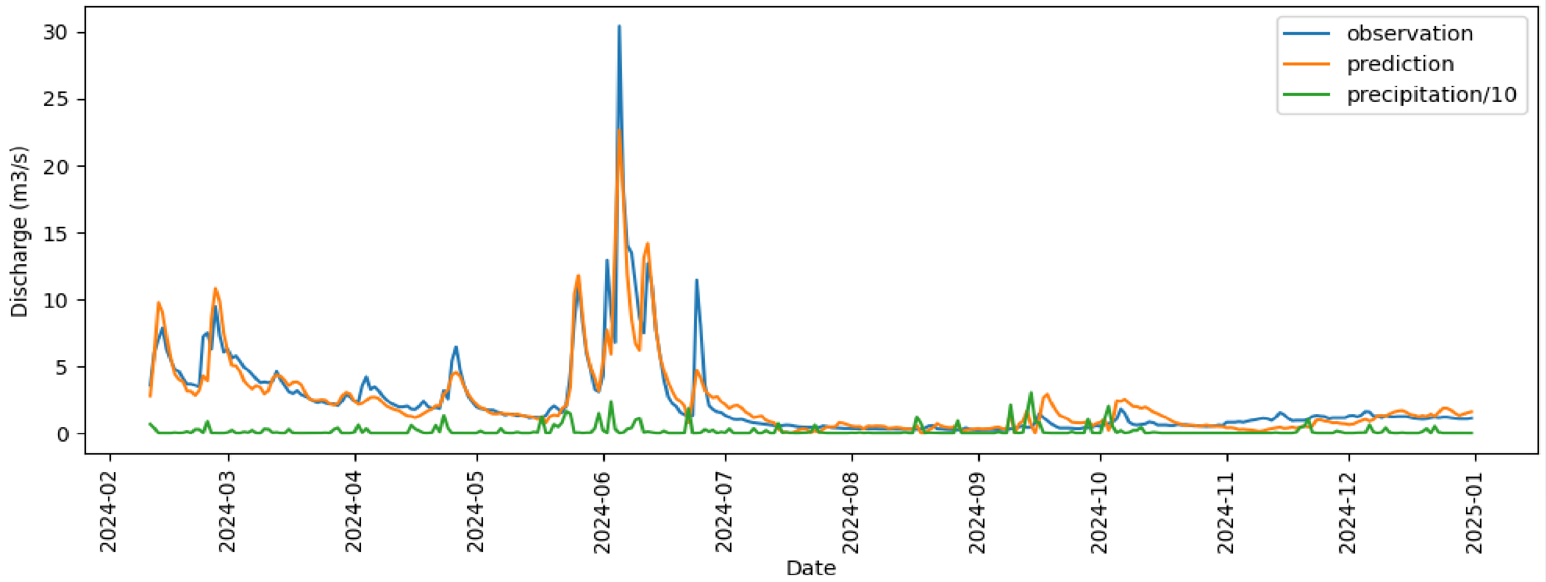
Important model features

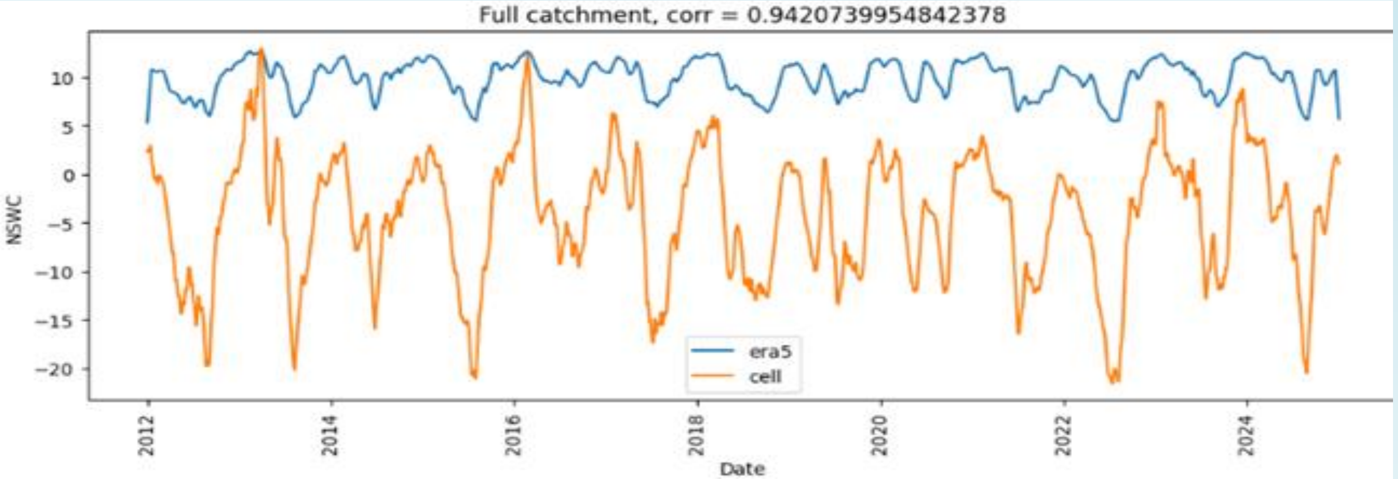
- **single LSTM cell, Cross Validation, output is generated only from input data sequence**
- **SWC data for calibration is compared with a simple differentiable output from the cell state vector, with the Pearson correlation coefficient subtracted from the cost function**
- ***Frederick Kratzert* shares the original model** at: https://github.com/neuralhydrology/, which also serves as a base of CAMELS-type **large sample studies for many countries** (US, DE)

Full catchment model – performance, results

The LSTM model learns successfully even extremes, and for some extent the occurence of snowmelts

<u>Performance on training + validation set:</u> **Nash-Sutcliffe Efficiency = 0.92**, on the test set: NSE 0.3–0.6





INTERPRETATION No. 1

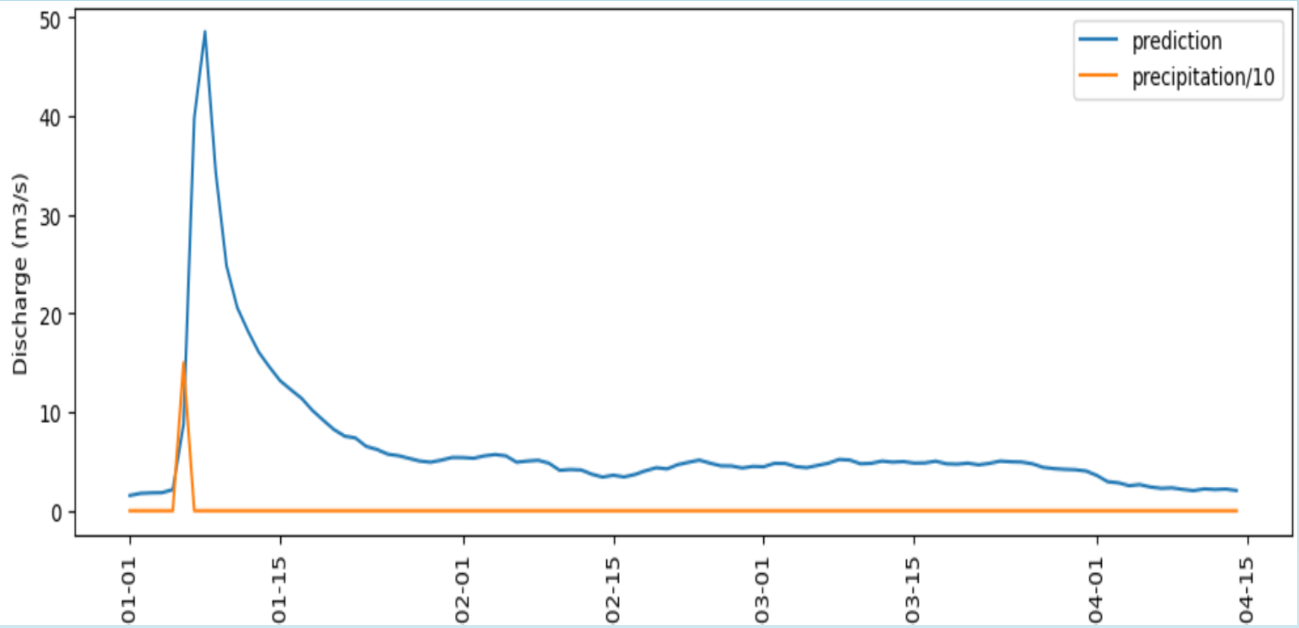
Calibration by the Soil Water Content estimation **raised the performance further, and correlation with cell state outputs reached high 0.94**!

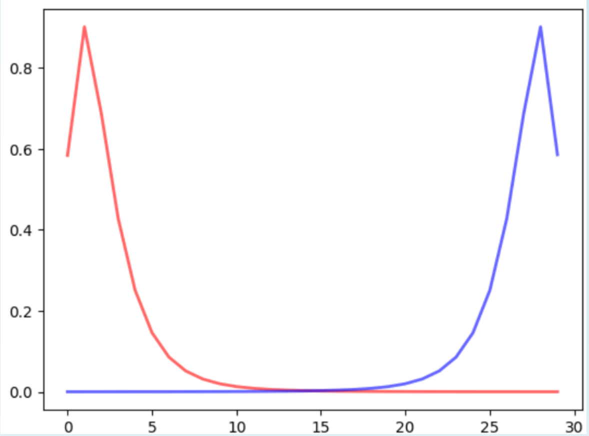
Further hydrographs – analysis and applications

INTERPRETATION No. 2

A trained LSTM model supports well investigating seasonal what-if scenarios!

With a single rain event inside an artificial dry test period: **THE DISCHARGE CURVE GENERATED BECOMES A LEARNED HYDROGRAPH!**





It's validity can be compared with outputs from hydrological simulation systems, or checked by a simple, discrete, 30-days reverse-hydrograph weighted summation module of rain amounts!

Semi-distributed setup, how to handle the ungauged

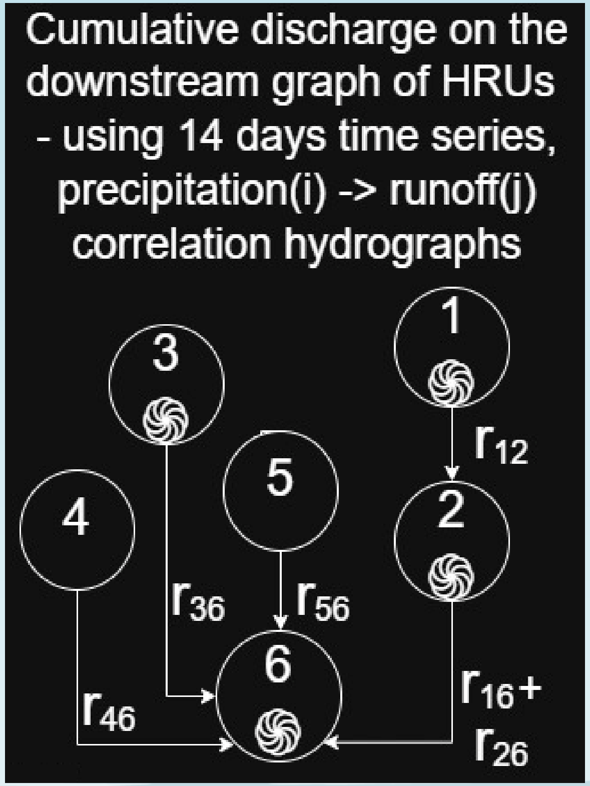
Streamflow routing is used in semi-distributed systems, like Hydro-GraphNet, or in HydroPy for the Amazon basin in *WANG, C. et al. (2024)*

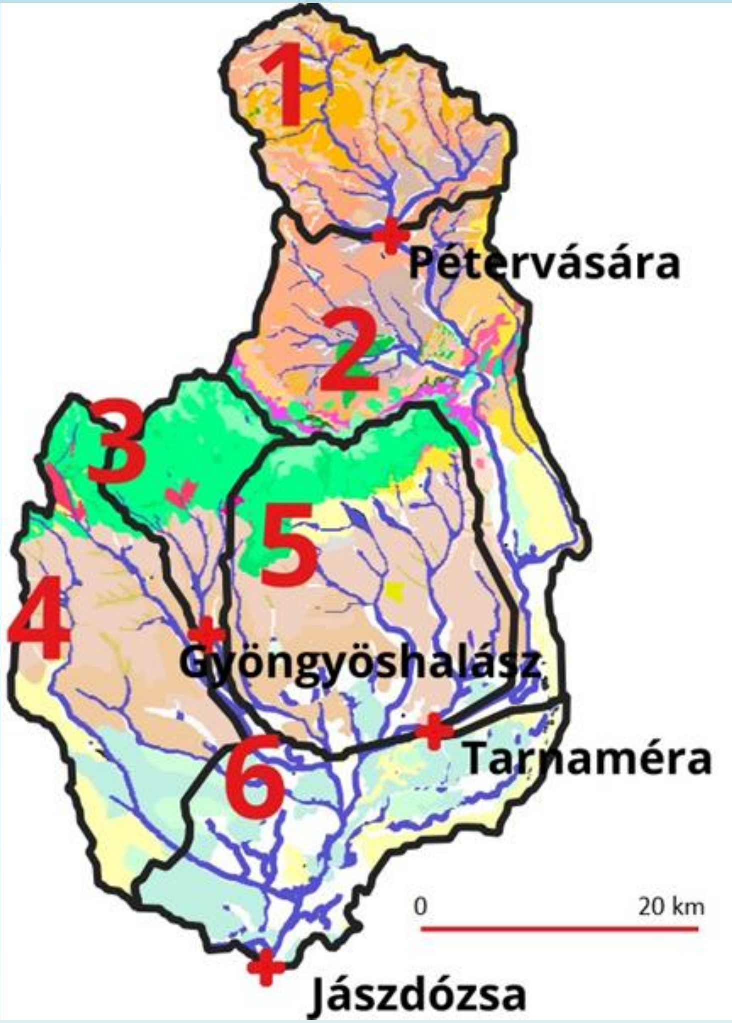
Transfer learning: area proportional weights and biases from the trained full-catchment model

Cost function: sum of MSE from the gauged HRUs, with added SWC calibration terms

Results confirmed these basic principles!

14 days correlations from shifted full time series data														
0.05	**0.17**	**0.16**	0.12	0.11	0.08	0.06	0.05	0.04	0.04	0.03	0.03	0.03	0.02	Rain1->Flow2
0.03	0.08	**0.13**	**0.12**	0.11	0.09	0.08	0.07	0.06	0.05	0.05	0.04	0.04	0.04	Rain1->Flow6
0.02	0.08	**0.13**	**0.12**	0.11	0.09	0.08	0.07	0.06	0.05	0.05	0.05	0.05	0.04	Rain2->Flow6
0.03	0.08	**0.13**	**0.12**	0.11	0.09	0.08	0.06	0.06	0.06	0.05	0.05	0.05	0.04	Rain3->Flow6
0.02	0.07	**0.13**	**0.11**	0.11	0.09	0.08	0.06	0.06	0.06	0.05	0.05	0.05	0.04	Rain4->Flow6
0.02	0.07	**0.13**	**0.11**	0.10	0.09	0.08	0.07	0.06	0.06	0.05	0.05	0.05	0.04	Rain5->Flow6





Colors: surface geology

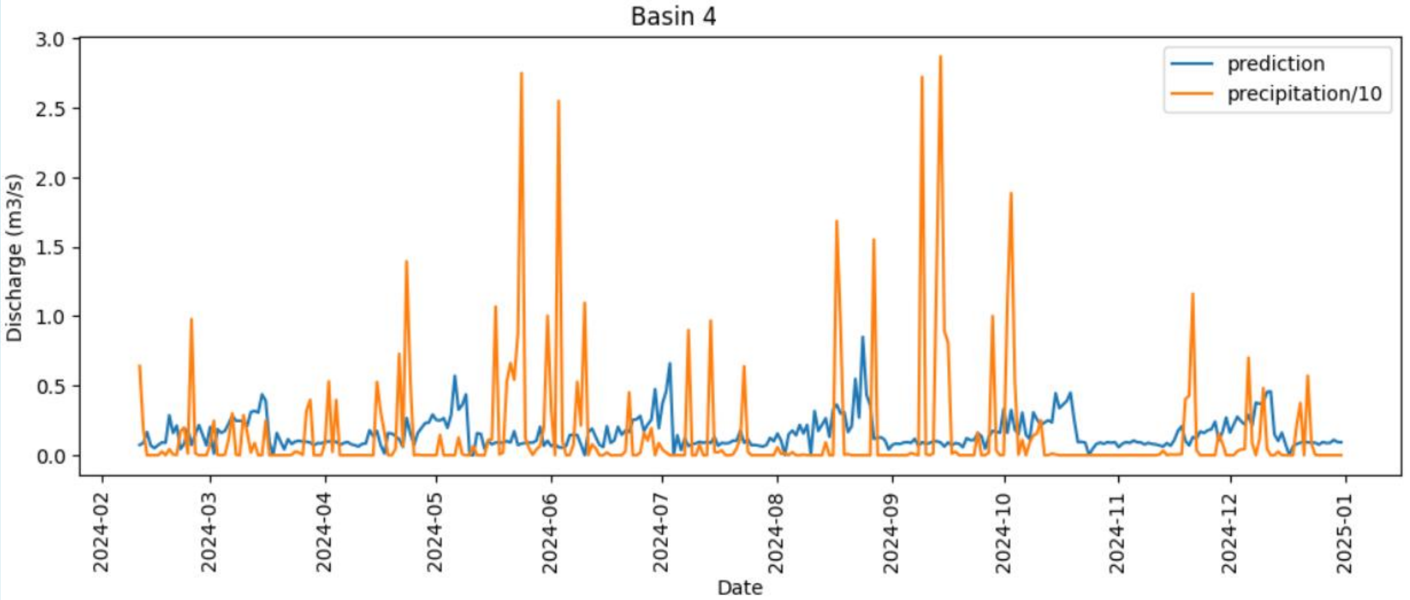
High complexity -> multiphase optimization strategy

Basin 1, NSE = 0.65 | Basin 3, NSE = 0.66

Basin 2, NSE = 0.76 | Basin 6, NSE = 0.68

Good performance on all gauged sub-basins (NSE = 0.1-0.3 on test period)!

NSWC calibration helped predictions on ungauged subbasins become much more realistic!



Final semi-distributed optimization sequence planned:

- first apply separately the full catchment model to top routing level gauged subbasins as well (1, 3 here)
- then switch these into evaluation mode and optimize the next routing level gauged subbasin (2 here), where a learned discharge transfer rate must be dependent on recent meteorology: $\sigma(r_{12}(PET))$
- and continue this way, keeping transfer learning, calibration, ungauged handling principles …

Baseflow contributions and recharge/discharge type

Special „dry test set": (last 3weeks) periods from the training/evaluation interval („well-known by the trained model") with (<1mm/4weeks) almost no precipitation/runoff/interflow

INTERPRETATION No. 3

Average of measured/predicted volumes gives the approximate **baseflow contribution share of those subbasins** (%)

Basin1	Basin2	Basin3	Basin4	Basin5	Basin6
6.4	48.8	16.0	4.5	8.0	16.3

Rem.: Basin2 is a karst area

INTERPRETATION No. 4

Coefficient of variance of volumes on this same „dry test set" will show **recharge/discharge characteristic**

Basin1	Basin2	Basin3	Basin4	Basin5	Basin6
0.24	0.18	0.35	0.81	0.88	0.40

Basins 4 and 5 are high mountain areas

Further steps and lessons learned

The trained full-basin LSTM model emulates well hydrological behavior and can be used for mid-term predictions or testing what-if scenarios

High 0.94 correlation, reached between cell state output and the ERA5 Near Surface Water Content estimation is a good base for calibration

Transfer learning and cross-hydrograph based discharge routing support well the semi-distributed Hydrological Response Unit (HRU) model

Further refinement/smaller HRUs and better handling of ungauged subbasins need geostatistical data preparation methods for GRACE satellite data downscaling by elevation, geology and Sentinel-2 index covariance + multi-phase optimization

Still, important explainable results are gained from the present model for baseflow contributions & HRU recharge/discharge characteristics