

Streamflow prediction in ungauged basins: benchmarking the efficiency of deep learning

Georgy Ayzel*, Liubov Kurochkina, Eduard Kazakov, and Sergei Zhuravlev

State Hydrological Institute, 199004 Saint Petersburg, Russia

Abstract. Streamflow prediction is a vital public service that helps to establish flash-flood early warning systems or assess the impact of projected climate change on water management. However, the availability of streamflow observations limits the utilization of the state-of-the-art streamflow prediction techniques to the basins where hydrometric gauging stations exist. Since the most river basins in the world are ungauged, the development of the specialized techniques for the reliable streamflow prediction in ungauged basins (PUB) is of crucial importance. In recent years, the emerging field of deep learning provides a myriad of new models that can breathe new life into the stagnating PUB methods. In the presented study, we benchmark the streamflow prediction efficiency of Long Short-Term Memory (LSTM) networks against the standard technique of GR4J hydrological model parameters regionalization (HMREG) at 200 basins in Northwest Russia. Results show that the LSTM-based regional hydrological model significantly outperforms the HMREG scheme in terms of median Nash-Sutcliffe efficiency (NSE), which is 0.73 and 0.61 for LSTM and HMREG, respectively. Moreover, LSTM demonstrates the comparable median NSE with that for basin-scale calibration of GR4J (0.75). Therefore, this study underlines the high utilization potential of deep learning for the PUB by demonstrating the new state-of-the-art performance in this field.

1 Introduction

Providing reliable streamflow predictions in ungauged basins (PUB) has crucial importance for understanding hydrological cycle processes where there are none or episodic hydrometric records [1]. While the most river basins in the world are ungauged, there is also a recent tendency for shrinking the number of gauging stations worldwide [2]. Therefore, PUB continues to be a pressing topic in hydrological modeling and requires particular attention from the research community [3].

In recent years, the emerging field of deep learning has advanced many scientific disciplines, including hydrology [4]. Deep learning models, such as convolutional and recurrent neural networks, proved their ability to finding complex relationships in natural phenomena by utilizing the power of big data that is recently available in open domain [5]. However, the use of modern deep learning techniques for PUB is lacking.

* Corresponding author: ayzelgv@gmail.com

The presented study aims to benchmark the prediction performance of deep learning for streamflow simulation in ungauged basins in comparison to standard and well-established technique of hydrological model parameters regionalization for 200 basins in Northwest Russia.

2 Data and Methods

2.1 Data

In the presented study, we use observed river streamflow data for 200 relatively undisturbed small-to-medium scale ($10 < \text{area} < 10\,000 \text{ km}^2$) river basins in Northwest Russia (within domain of $25\text{--}57^\circ \text{ E}$ and $55\text{--}70^\circ \text{ N}$). This data was digitized from the annual digest archives (hydrological yearbooks) in the framework of the R5 project [6]. The corresponding basin boundaries were semi-automatically digitized using a digital elevation map.

The source of meteorological data is the WFDEI global meteorological reanalysis dataset [7]. WFDEI is based on the ERA-Interim atmospheric reanalysis and has a 0.5° spatial and daily temporal resolution. We extract gridded WFDEI precipitation (P) and air temperature (T) data at a basin-scale using a spatial averaging. Potential evaporation (PE) was calculated based on Oudin's formulation [8] using air temperature data.

2.2 Models

In the presented study, we use two models to simulate streamflow: (1) conceptual lumped hydrological model GR4J [9] that incorporates the CemaNeige snow routine [10] (Fig. 1), and (2) deep learning model, namely Long Short-Term Memory Network (LSTM, Fig. 2). Both models proved their efficiency for PUB in recent studies [11, 12].

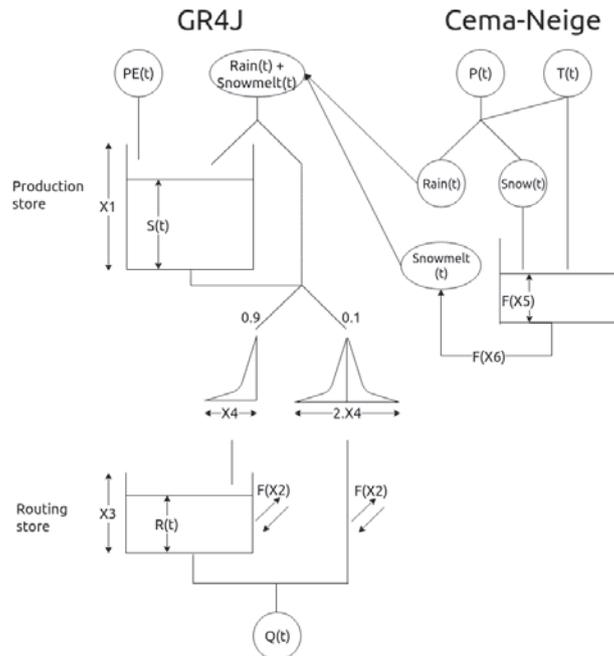


Fig. 1. Conceptual schemes of GR4J hydrological model and Cema-Neige snow routine.

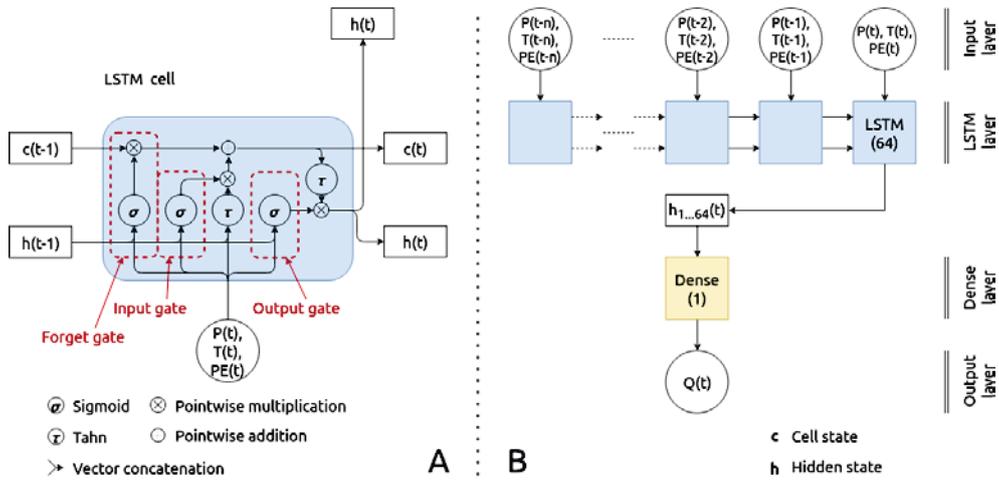


Fig. 2. The standard LSTM cell (A) and LSTM model for streamflow prediction (B).

2.3 Experimental design

Fig. 3 shows the flowchart of the established benchmark experiment.

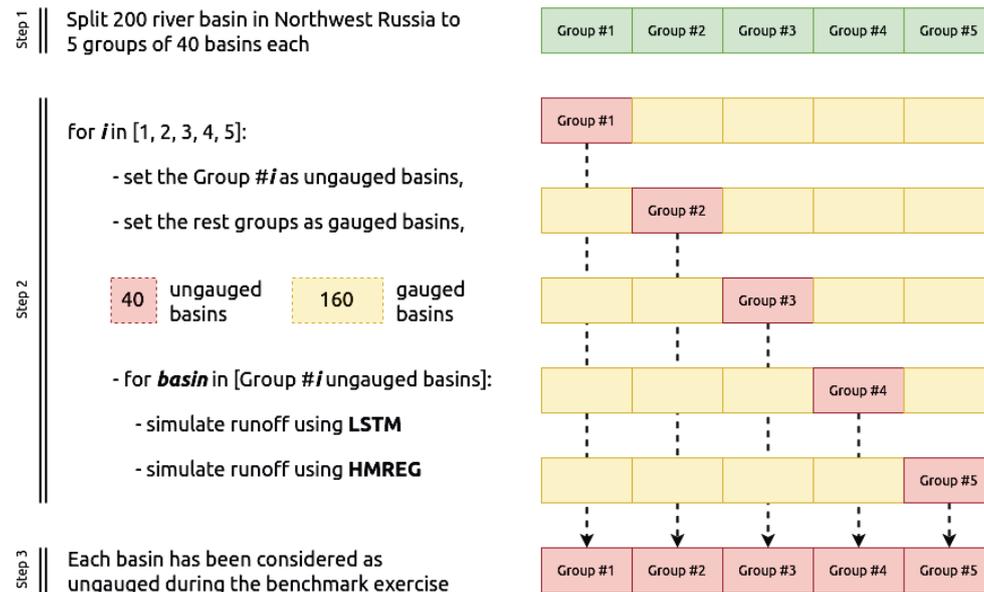


Fig. 3. Flowchart of the benchmark experiment.

We use a 5-fold cross-validation technique to test the out-of-sample performance of two PUB techniques: (1) hydrological model parameters regionalization (HMREG) and (2) regional LSTM model. In the first step, we randomly split 200 available basins into five groups. Then, in the second step, we consider basins from one of the five groups as ungauged (i.e., we use them only for the evaluation) and the rest as of gauged (which we use to support the techniques under consideration). We repeat this exercise iteratively until every group, as well as every basin, has been considered as ungauged.

HMREG workflow for PUB is based on GR4J hydrological model and well-established nearest-neighbor regionalization technique [13], and can be summarized as follows:

1. For each ungauged basin in the corresponding group, find ten nearest donor basins from the set of the rest 160 gauged basins,
2. For each donor basin, calibrate the GR4J model against streamflow observations using the global optimization method of differential evolution (for details see [14]),
3. For each ungauged basin, simulate streamflow using the GR4J model and ten different optimal parameter sets obtained for respective donor basins. Then, calculate the ensemble mean of streamflow simulations.

LSTM-based workflow for PUB utilizes the entirely different idea that can be expressed in using all the data from the gauged group of basins for regional LSTM model calibration. To this end, LSTM-based workflow can be summarized as follows:

1. Calibrate parameters of ten regional LSTM models against streamflow observations for 160 gauged basins. These ten LSTM models differ from each other by randomly assigned initial conditions at the beginning of the calibration procedure (for details see [15]).
2. For each ungauged basin in the corresponding group, simulate streamflow using ten calibrated regional LSTM models. Then, calculate the ensemble mean of streamflow simulations.

We use the Nash-Sutcliffe efficiency (NSE) metric to evaluate the daily streamflow prediction efficiency of HMREG and LSTM-based workflows for PUB.

3 Results and Discussion

The results of the benchmark experiment are summarized in Fig.4. It is a common practice in PUB studies that the basin-scale calibration results provide a so-called “superlative estimate,” i.e., show the best performance estimate that can be reachable in case of using the hydrological model for streamflow prediction. In the presented study, GR4J showed a median NSE of 0.75 with interquartile range (IQR; Q75-Q25) of 0.11, and the lowest NSE of 0.51. Thus, the prediction efficiency of GR4J for 200 river basins in Northwest Russia can be considered as satisfactory [16].

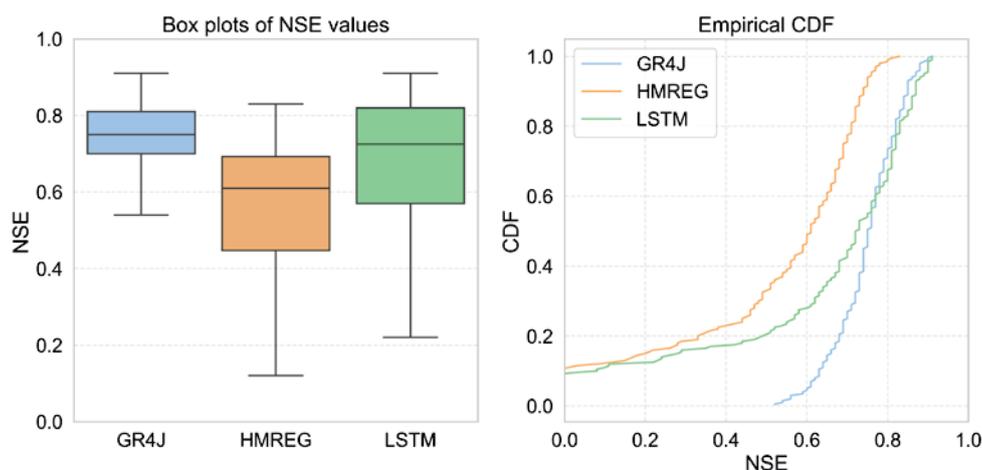


Fig. 4. Boxplots (left panel) and cumulative distribution functions (CDF) of NSE values obtained for 200 basins in Northwest Russia.

The standard baseline technique for PUB – HMREG – provides reliable results with a median NSE of 0.61 with IQR of 0.25. Additionally, for 135 out of 200 basins, HMREG implementation ensures satisfactory results ($NSE > 0.5$) in streamflow predictions in ungauged basins. However, there are 21 basins where NSE is negative, so HMREG has no skill at all. The obtained results confirm our previous study [17], where we showed that while model parameters regionalization technique based on spatial proximity provides reliable results for PUB, it cannot ensure uniform efficiency for a considerable portion of analyzed basins.

In the presented study, we propose to use the LSTM deep learning model as a regional hydrological model that can assimilate all the information from gauged basins during the calibration procedure (also training or learning). Then, the calibrated LSTM model can be used for streamflow simulations in ungauged basins utilizing meteorological reanalysis data as input forcing. Results show that the LSTM-based regional hydrological model significantly outperforms the HMREG scheme in terms of NSE. Thus, LSTM has a median NSE of 0.73 with IQR of 0.25. However, while HMREG provides comparable or better results than LSTM for 24 out of 200 basins, for 8 of 24 basins both schemes are not skilful ($NSE < 0$).

Additionally, LSTM demonstrates the comparable median NSE with that for basin-scale calibration of GR4J, which is 0.73 and 0.75, respectively (Fig. 4). Moreover, for 96 out of 200 basins, LSTM showed comparable or even better NSE compared to basin-scale GR4J calibration results. This finding revealed that the implementation of the regional LSTM model provides the new state-of-the-art prediction efficiency for PUB studies, and beyond.

4 Conclusions

In the presented study, we show the high utilization potential of deep learning for the PUB by demonstrating the new state-of-the-art performance in this field. Thus, the regional LSTM-based deep learning model for streamflow predictions in ungauged basins, on average, outperforms the standard hydrological model parameters regionalization technique (HMREG). Moreover, for almost half of analyzed basins, LSTM also outperforms the results of the hydrological model (GR4J) basin-scale calibration. Therefore, we underline the importance of adapting the new methods from the emerging field of deep learning for hydrological applications as they can demonstrate and set the state-of-the-art prediction performance in the field.

The reported study was funded by RFBR, project number 19-35-60005, and by RGS-RFBR grant under project number 17-05-41118.

References

1. M. Sivapalan et al., *Hydrol. Sci. J.*, **48**, 6, 857–880 (2003)
2. B. M. Fekete, U. Looser, A. Pietroniro, R. D. Robarts, *J. Hydrometeorol.*, **13**, 6, 1977–1986 (2012)
3. M. Hrachowitz et al., *Hydrol. Sci. J.*, **58**, 6, 1198–1255 (2013)
4. C. Shen et al., *Hydrol. Earth Syst. Sci.*, **22**, 11, 5639–5656 (2018)
5. M. Reichstein et al., *Nature*, **566**, 7743, 195–204 (2019)
6. G. Ayzel, L. Kurochkina, S. Zhuravlev, *Hydrol. Sci. J.* (in review)
7. G. P. Weedon, G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, P. Viterbo, *Water Resour. Res.*, **50**, 9, 7505–7514, (2014)

8. L. Oudin et al., *J. Hydrol.*, **303**, 1–4, 290–306 (2005)
9. C. Perrin, C. Michel, V. Andréassian, *J. Hydrol.*, **279**, 1–4, 275–289 (2003)
10. A. Valéry, V. Andréassian, C. Perrin, *J. Hydrol.*, **517**, 1176–1187 (2014)
11. R. Arsenault, M. Breton-Dufour, A. Poulin, G. Dallaire, R. Romero-Lopez, *Hydrol. Sci. J.*, **64**, 11, 1–15 (2019)
12. F. Kratzert, D. Klotz, M. Herrnegger, A. K. Sampson, S. Hochreiter, G. S. Nearing, *Water Resour. Res.*, 2019WR026065 (2019)
13. G. Ayzel and A. Izhitskiy, *Water*, **11**, 11, 2377 (2019)
14. G. Ayzel, N. Varentsova, O. Erina, D. Sokolov, L. Kurochkina, V. Moreydo, *Water*, **11**, 8, 1546 (2019)
15. G. Ayzel, Proceedings of V International Conference Information Technologies and High-Performance Computing (ITHPC-2019), available at: <http://ceur-ws.org/Vol-2426/paper9.pdf> (2019)
16. D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, T. L. Veith, *Trans. ASABE*, **50**, 3, 885–900 (2007)
17. G. V. Ayzel, E. M. Gusev, O. N. Nasonova, *Water Resour.*, **44**, 4, 547–558 (2017)