Hydrological Significance of Input Sequence Lengths in LSTM-Based Streamflow Prediction

F. Hosseini [a], C. Prieto [a], G. Nearing [b], C. Álvarez [a], M. Gauch [b]

[a] IHCantabria—Instituto de Hidráulica Ambiental de la Universidad de Cantabria, Santander, Spain

[b] Google Research

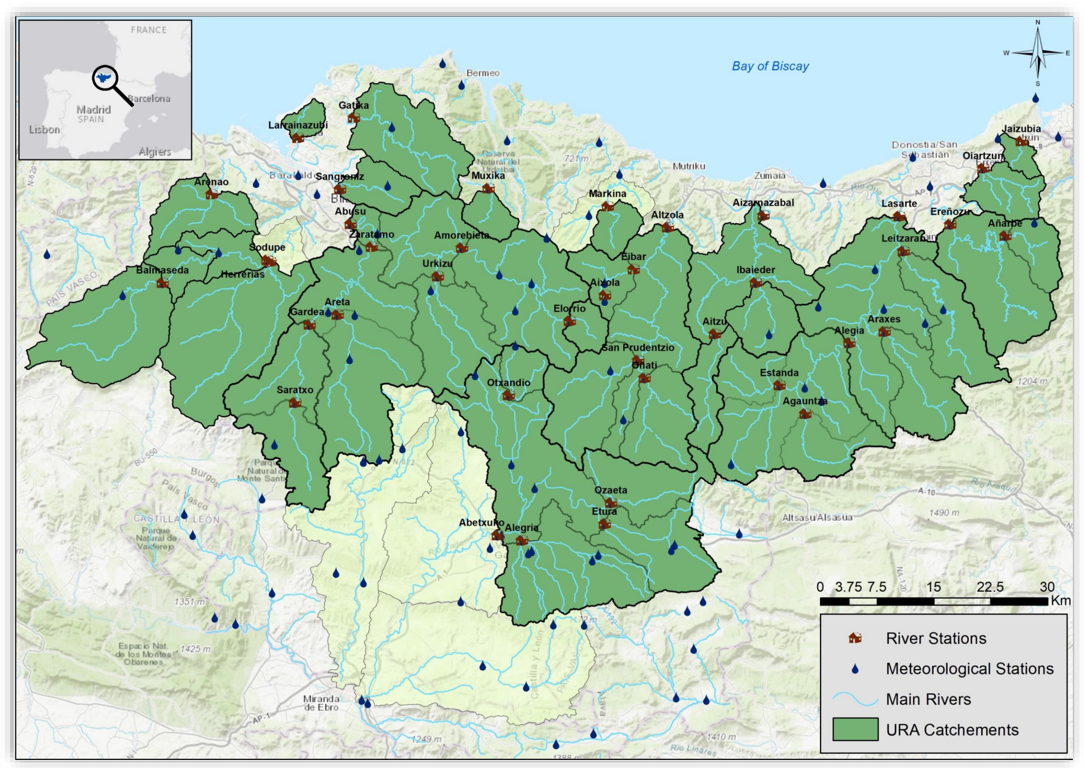
Introduction and Motivation

The hydrological modeling of flashy catchments, which are susceptible to floods, poses a significant practical challenge. Recent applications of deep learning, specifically Long Short-Term Memory networks (LSTMs), have demonstrated notable capabilities in delivering accurate hydrological predictions at daily and hourly intervals (Gauch et al., 2021; Kratzert et al., 2018).

The primary objective was to investigate the influence of various hyperparameter configurations (model settings) on the performance of regional streamflow LSTMs. We conducted a systematic hyper-tuning process by exploring among 1000 different configurations of 12 distinct hyperparameters. This approach led to the development of different final hyper-tuned LSTMs, which were then retrained and tested.

Method

We utilized a multi-timescale LSTM network (MTS-LSTM) (Gauch et al., 2021) to predict hydrographs in flashy catchments at hourly time scales. Our focus was on training regional hydrological MTS-LSTM networks to predict hourly streamflow and water level in the humid flashy catchments of Basque Country, located in north of Spain.



Results

We found that hyperparameters related to the length of the input sequence significantly impact the performance of regional models. Utilizing unsupervised machine learning models, we identified the optimal regional values for this hyperparameter in our case study, determining them to be 3 years for daily and 12 weeks for hourly data.

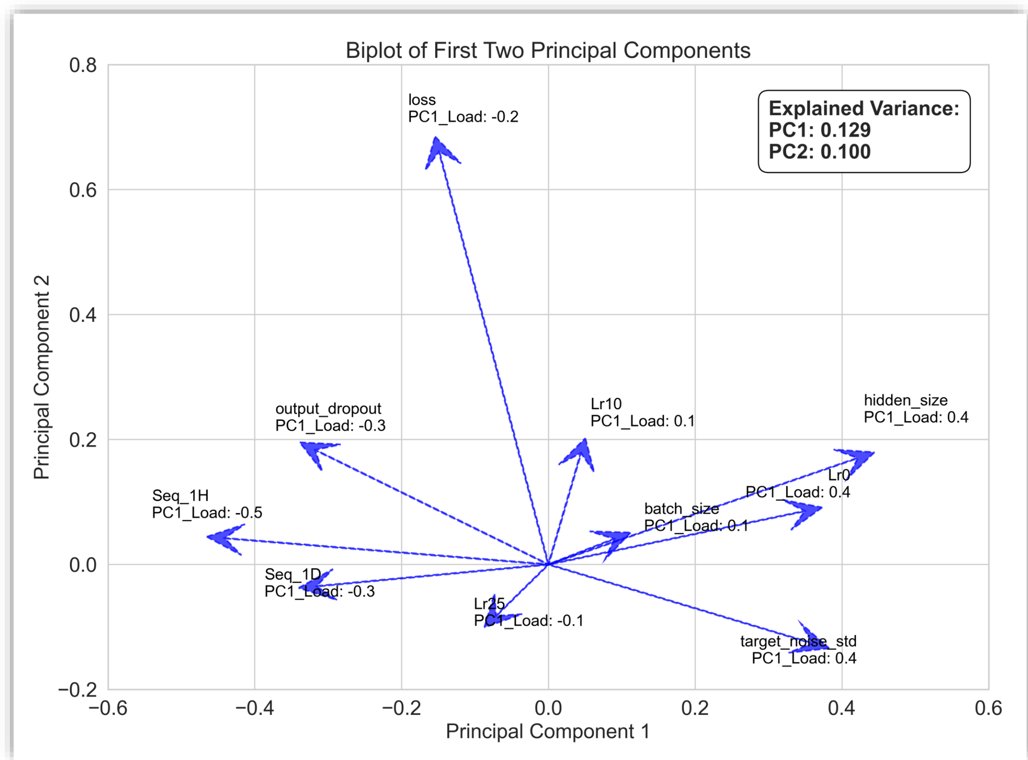


Fig.2: PCA Loadings for PCA1 vs. PCA2 demonstrates significance of Input Sequence Length Hyperparameters.

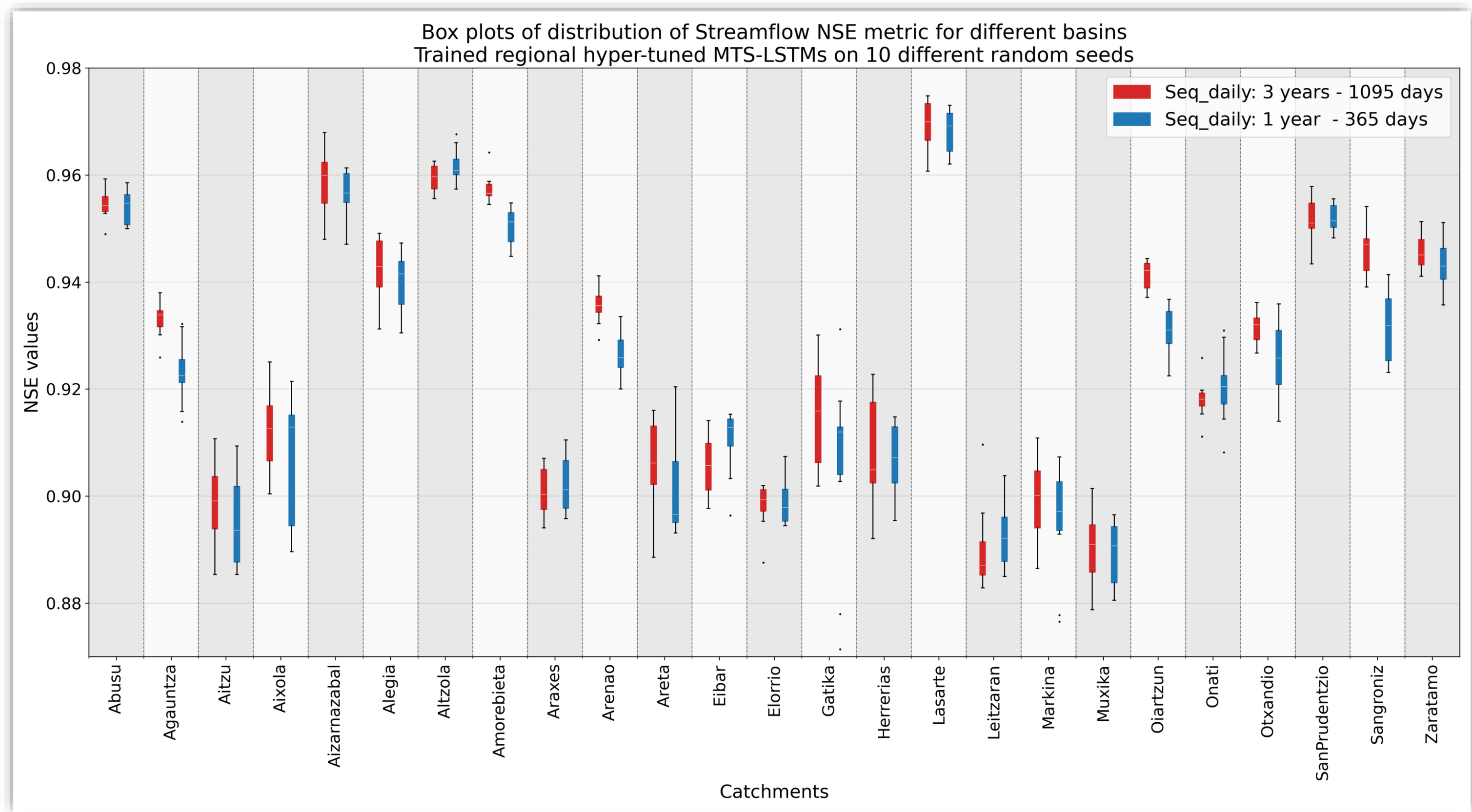
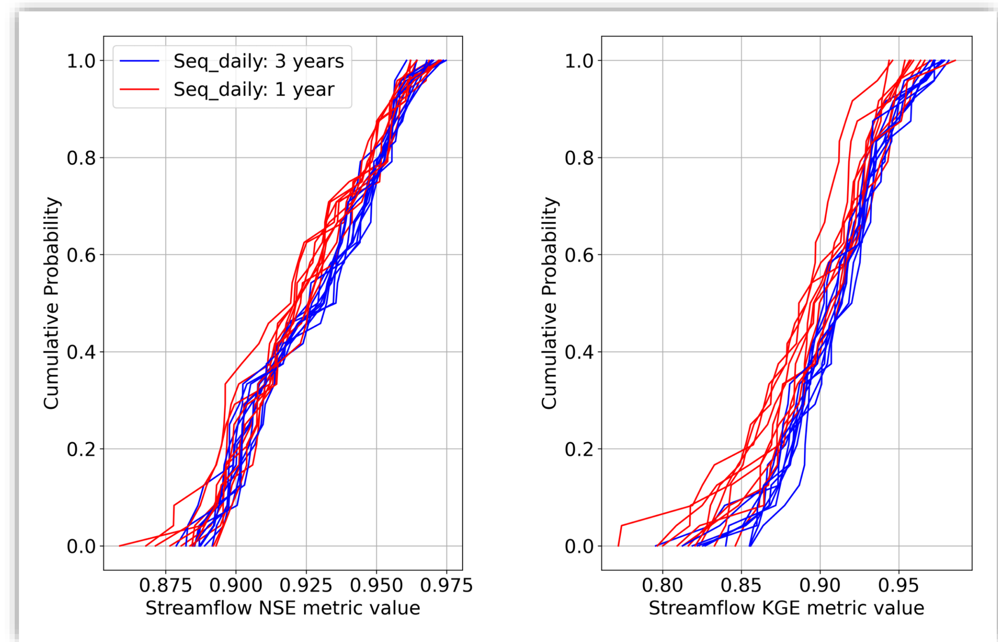


Fig.1: Distinct regional hyper-tuned MTS-LSTM networks with different configuration settings; one has input sequence length of 3 years and the other 1 year. Although, both versions demonstrate high accuracy, regionally and locally; the one having 3 years of input sequence length outperformed the other, regionally and in several places.

Fig.3: Cumulative distribution function plots of KGE and NSE metrics for ten simulations of two distinct regional hyper-tuned MTS-LSTM networks with different configuration settings. Plots clearly demonstrate outperformance of the model having daily input sequence length of 3 years.



Our methodology, which involves simultaneous hyper-tuning of LSTMs, demonstrates high accuracy for regional predictions both regionally and on a catchment-scale basis (see Fig. 1), with hourly and daily NSE values reaching up to 0.969 and 0.941, respectively. Moreover, Principal Component Analysis (PCA) on the performance of different configurations reveals that the first principal component explains 13% of the variance among the 12 hyperparameters. Within this set of hyperparameters, the input sequence length for hourly data exhibits the highest loading in PC1, with a value of -0.5, while the loading of the input sequence length for daily data is also very high (-0.3) (see Fig. 2). This suggests that properly tuning the hyperparameter strongly contributes to the network's performance.

Precisely tuned input sequence length for a catchment may encapsulate hydrological information pertaining to water transit over short and long-term periods. Notably, the regional daily sequence length aligns with the highest local daily sequence values across all catchments.

Conclusion

Catchment-scale analyses reveal distinctive input sequence lengths for individual basins. This underscores the necessity of customizing the length of the input sequence in LSTMs, based on the "uniqueness of the place" paradigm (Beven, 2020).

This study emphasizes the hydrologically critical role of the input sequence length hyperparameter that needs to be tuned in LSTM networks for accurate streamflow and water level predictions.

We suggest that each catchment may require specific hydrologically meaningful daily and hourly input sequences. These values can be tuned through systematic hyper-tuning of hyperparameters to encode hydrologically meaningful information in this hyperparameter, which subsequently assists tuned networks generating accurate predictions.

References

Beven, K. (2020). Deep learning, hydrological processes and the uniqueness of place. Hydrological Processes, 34(16), 3608–3613. doi:10.1002/hyp.13805

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., and Hochreiter, S. (2021). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network, Hydrol. Earth Syst. Sci., 25, 2045–2062, DOI:10.5194/hess-25-2045-2021.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall--runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrology and Earth System Sciences, 22(11), 6005–6022. DOI:10.5194/hess-22-6005-2018.

EGU General Assembly 2024

Deep learning in hydrology; A.56|EGU24-571|HS3.4|ECS | Posters on site: Thu, 18 Apr

Sharing is encouraged