Streamflow Prediction Using LSTM Models Trained on Synthetic Gage Data

Quinn Y. Lee[1], Abel Andrés Ramírez Molina[2], James Halgren, Ph.D., PE[3], Jonathan M. Frame, Ph.D., PE[4]

[1]Department of Mathematics, The University of Alabama, [2]Department of Computer Science, The University of Alabama, [3]Alabama Water Institute, The University of Alabama, [4]Lynker Technologies, LLC

Abstract

Accurate streamflow predictions in ungaged basins are critical for water resource management, yet less than 10% of U.S. river segments are actively gaged. Addressing this, [1] proposed an LSTM model trained on synthetic basin networks, utilizing a "Leaky Bucket" model developed by [2] and signal separation techniques developed by [3] to generate and disaggregate streamflow data. This research aims to evaluate the robustness of this model by improving the synthetic hydrographs generated by the Leaky Bucket.

Bucket Model Structure

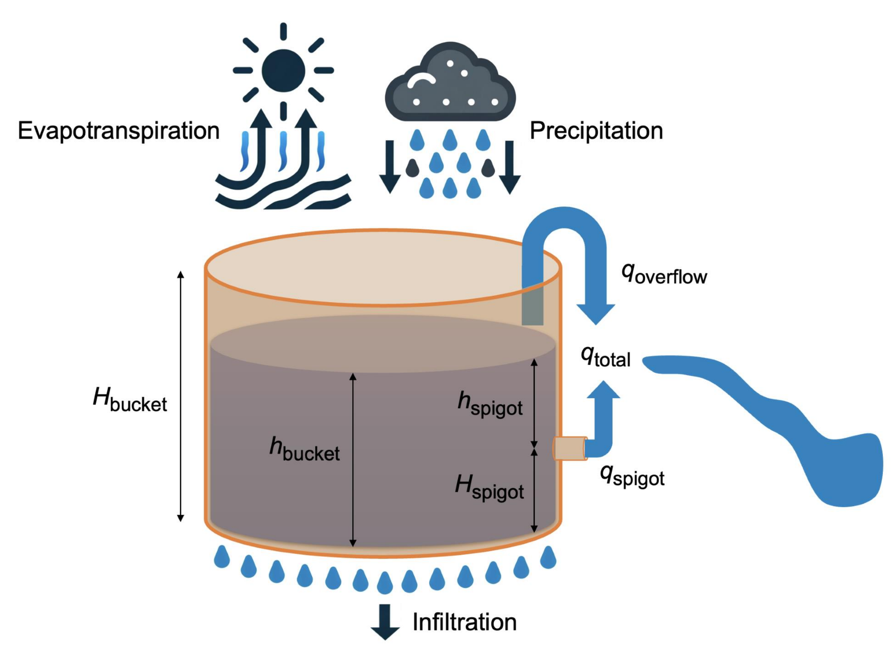

Figure 1. Bucket model structure. [1]

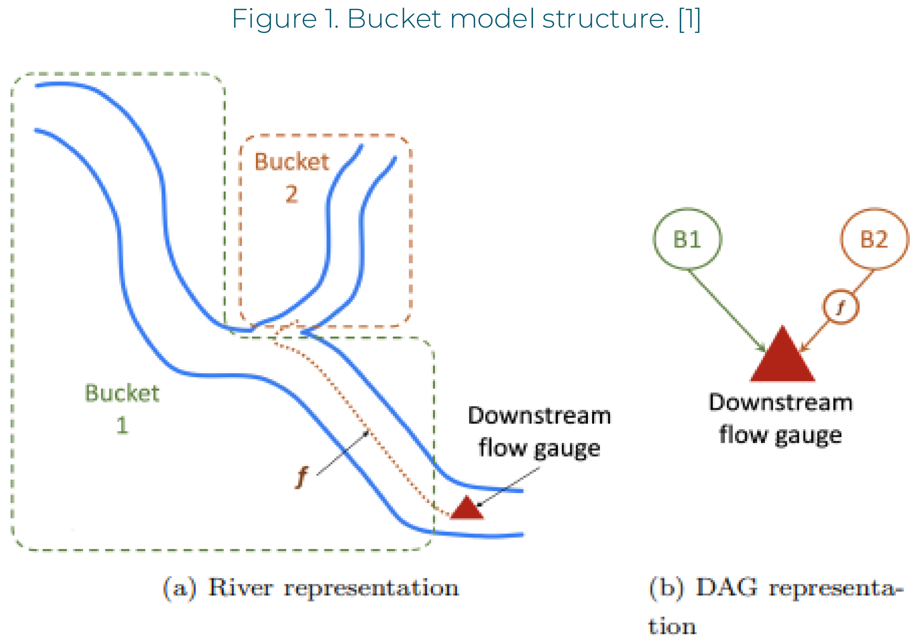

Figure 2. Bucket network structure. [1]

Issues Addressed in Bucket Model

- Unrealistic distributions
- Unit and timestep mismatches
- Didn't obey physics
- Temporal resolution too fine
- Unrealistic hydrograph shapes

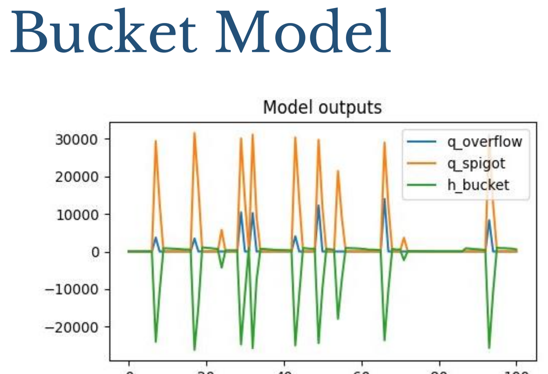

Figure 3. Strange bucket outputs.

Long Short-Term Memory (LSTM)

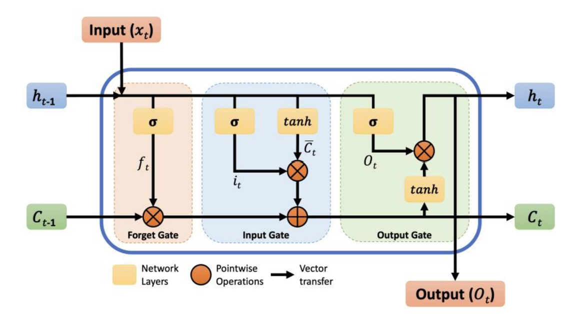

Figure 4. An LSTM cell. [1]

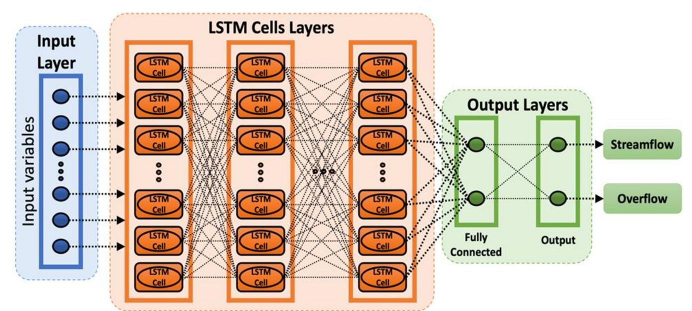

Figure 5. LSTM model structure. [1]

Generating Inputs

Random Value Generation

Process to generate probability distributions for inputs:

1. Identify potential distributions during literature review
2. Collect and transform data to desired units
3. Use Maximum Likelihood Estimation and Method of Moments to fit data to distributions
4. Use Kolmogorov-Smirnov test to compare goodness-of-fit

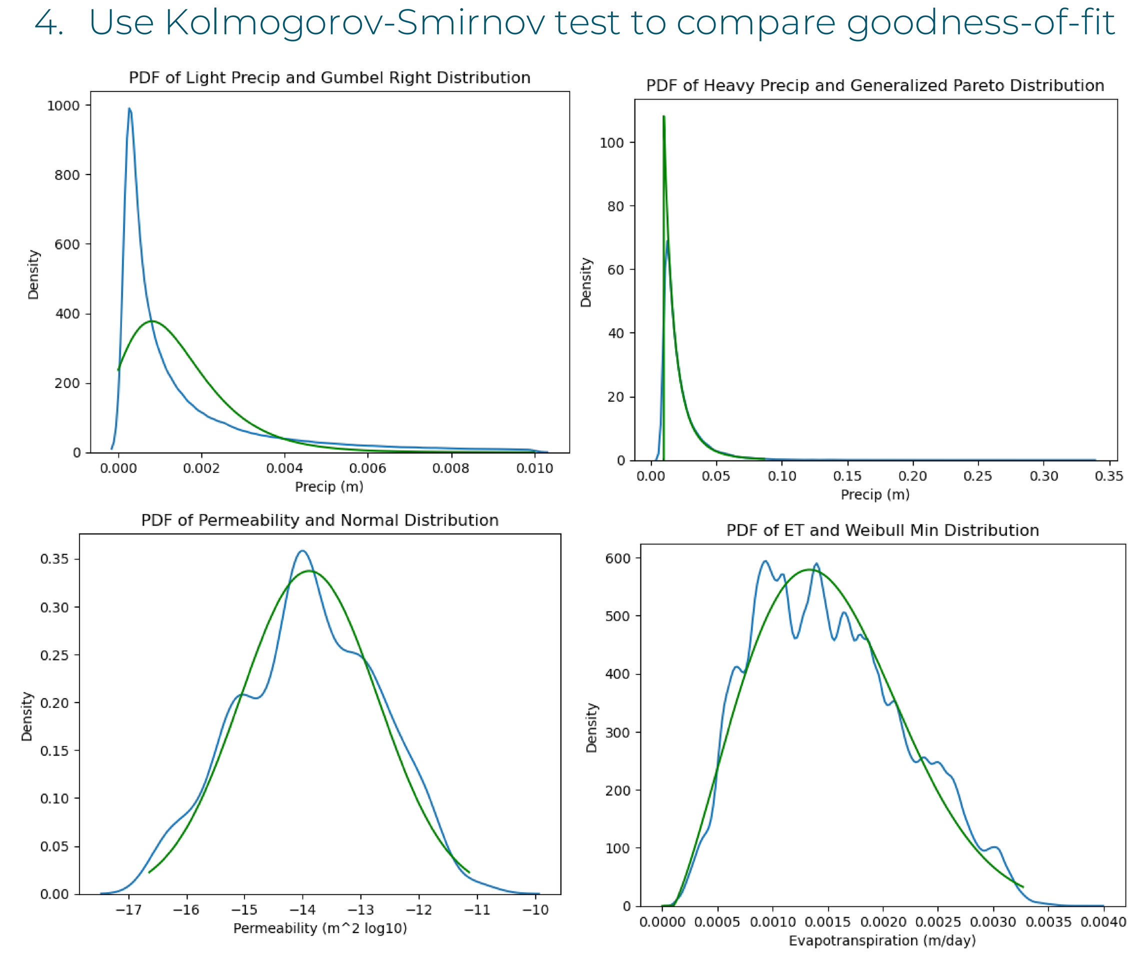

Figure 6. Probability distribution functions.

Other Calculations

Area and height of spigot are now functions of height of bucket instead of being independent

- Expanded ranges of possible dimensions
- Evapotranspiration (ET) transformed using a sine wave to represent diurnal fluctuations
- Soil depth added as a parameter to calculate groundwater infiltration using Darcy's Law

Generating Outputs

Refining Calculations

- Timestep changed from second to hour
- Prevented water from draining if it was below spigot
- Discharge values normalized by basin area to simplify model evaluation

Unit Hydrograph Transformation

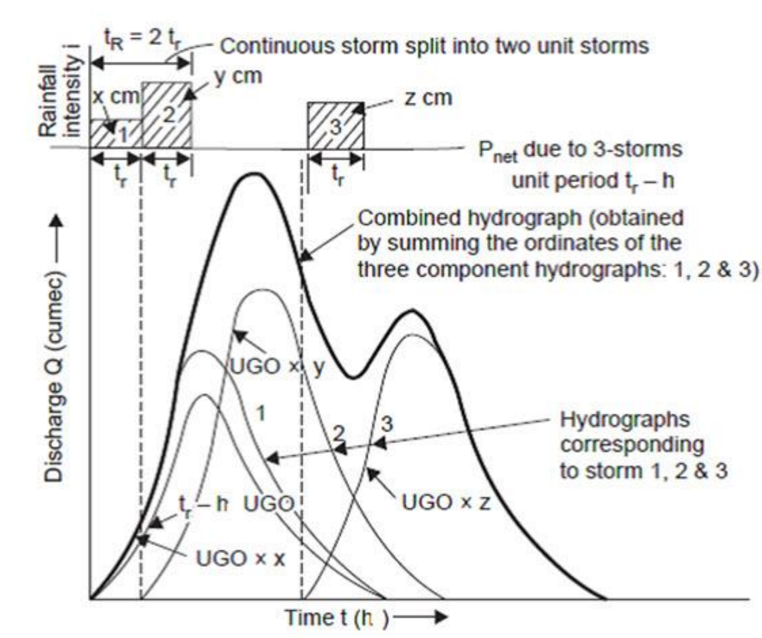

Figure 7. Elements of a unit hydrograph. [4]

Transforming q_spigot + q_overflow = q_total_pre into q_total

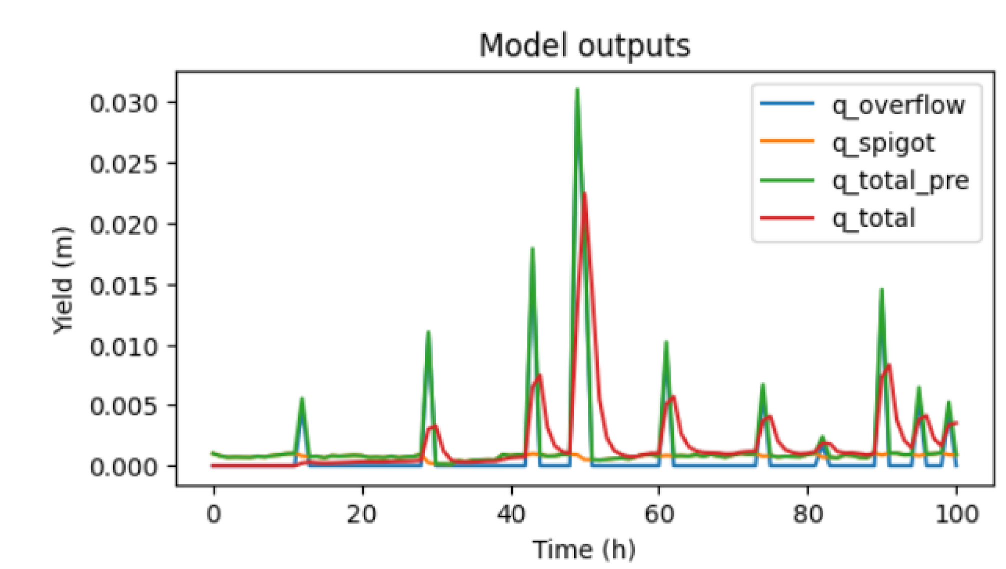

Figure 8. Results of unit hydrograph transformation.

Network Structure – Two Buckets

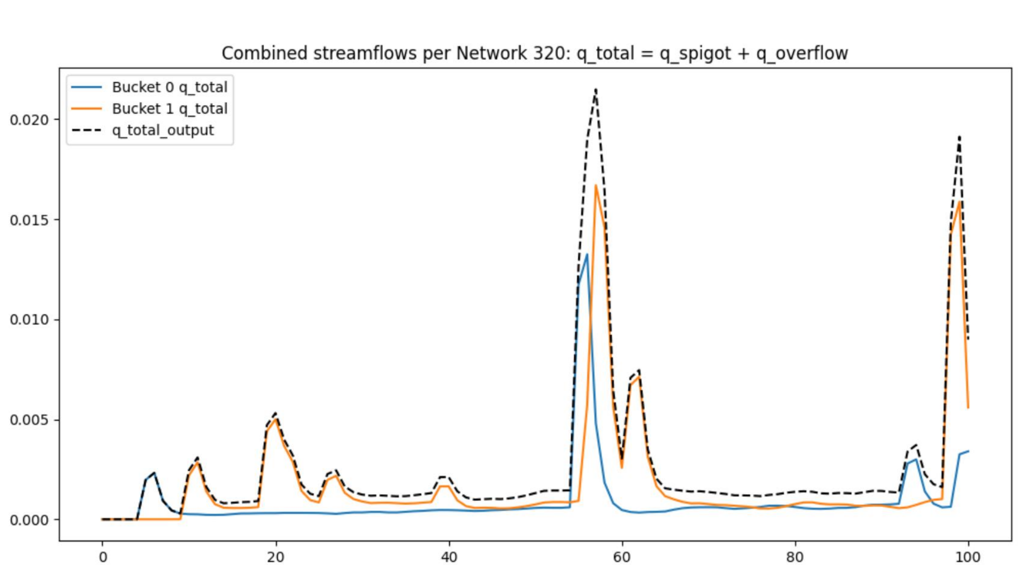

Figure 9. New bucket network structure.

Models and Results

Model type	Inputs	Outputs	Nash-Sutcliffe Efficiency (NSE)
Individual bucket with combined gage data (1)	Precipitation, spigot area, spigot height, bucket height + network combined discharge (NCD)	Total discharge	0.69
Combined buckets with combined gage data (2)	Model 1 inputs for each bucket in network + NCD	Total discharge for each bucket	0.78

Table 1. Model types and results.

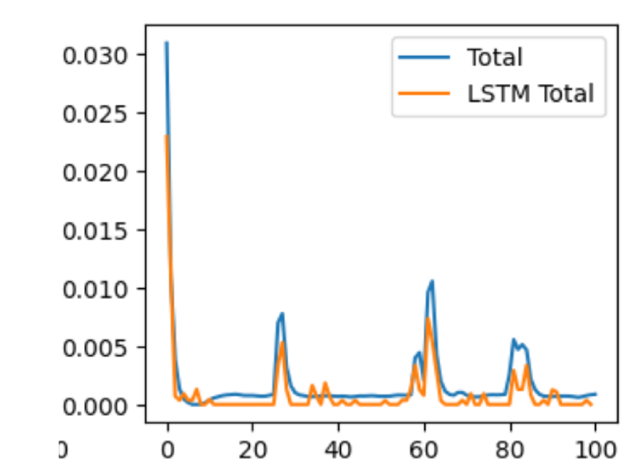

Figure 10. Model types and results.

An LSTM model trained on downstream gage data and basin parameters does predict upstream streamflow well, even when synthetic basin simulations are more realistic and parameter ranges are wider.

References and Acknowledgements

[1] A. A. Ramírez Molina, J. M. Frame, J. Halgren, and J. Gong, "Synthetic stream gauges: An lstm-based approach to enhance river streamflow predictions in unmonitored segments," 2024, submitted to JGR: Machine Learning and Computation.
[2] J. M. Frame, L. Hernandez Rodriguez, and M. Bassiouni, "Deepbucketlab – a playground for understanding deep learning for hydrologic process representations." To be published. [Online]. Available: doi:10.5072/zenodo.7349
[3] A. Ephrat, I. Mosseri, O. Lang, T. Dekel, K. Wilson, A. Hassidim, W. T. Freeman, and M. Rubinstein, "Looking to listen at the cocktail party: A speaker-independent audio-visual model for speech separation," *CoRR*, vol. abs/1804.03619, 2018. [Online]. Available: http://arxiv.org/abs/1804.03619.
[4] H. M. Raghunath, *Hydrology: principles, analysis, and design*. New Delhi: New Age International, 2006.

GitHub Repository:
https://github.com/quinnylee/synthetic_stream_gages

Funding for this project was provided by the National Oceanic and Atmospheric Administration (NOAA), awarded to the Cooperative Institute for Research on Hydrology (CIROH) through the NOAA Cooperative Agreement with The University of Alabama, NA22NWS4320003.