Deep Learning for Rainfall-Runoff Modeling

neuralhydrology.github.io

Grey S. Nearing¹,², Frederik Kratzert³, Alden K. Sampson⁴, Craig S. Pelissier⁵, Daniel Klotz³, Jonathan M. Frame², Cristina Prieto⁶, Hoshin V. Gupta⁷

¹Google Research, ²University of Alabama, ³Johannes Kepler University, LIT AI & Machine Learning Laboratory, ⁴Upstream Tech, Public Benefit Corporation, ⁵NASA Center for Climate Simulation, ⁶Instituto de Hidraulica Ambiental de la Universidad de Cantabria, ⁷University of Arizona
Long Short Term Memory (LSTM)

The LSTM is a recurrent neural network with an input-output-state relationship.
LSTMs are State-Space Models

State space model:
\[
S[t] = f(I[t], S[t-1]; \Theta_i) \\
O[t] = g(S[t]; \Theta_j)
\]

LSTM model:
\[
\{c[t], h[t]\} = f(x[t], c[t-1], h[t-1]; \theta_i) \\
\hat{y}[t] = g(h[t]; \theta_j)
\]
Embedding into Deep Learning Models

LSTM (time series model)

Legend:
- Fully coupled layer
- Convolutional layer
- Graph convolution
Experimental Setup

LSTM-based model

Catchment Attributes

Meteorological Forcings

531 Basins

9 Years Training Data

Regional LSTMs are better than catchment-specific hydro models.
Prediction in Ungauged Basins

LSTMs are better in ungauged basins than SAC-SMA is in gauged basins.

Learning a General Model

The LSTM is better when trained on multiple catchments than when trained on individual catchments.

Certain “hard” tasks are easy with DL

Multiple Forcings w/o Ensembles
Certain “hard” tasks are easy with DL
Certain “hard” tasks are easy with DL

Estimating Uncertainty

Quantiles of Predictive Distributions

Alden Sampson;
Upstream Tech, PBC
Physics Integration
Post-Processing

National Water Model Post-Processing

Adding NWM states and fluxes as inputs did not improve the LSTM

NWM States and Fluxes (Daily average)

Meteorological Forcing Data (NLDAS)

LSTM as a Dynamic Post-Processor

Improved Streamflow Response

Post-Processing

The LSTM “listens” to the NWM, but there isn’t any extra information.

**Figure 7.** Attributions to the LSTM post-processor predictions. The vertical axis shows the relative magnitude of attribution (importance) for each input, with precipitation (PRCP) as the top contributor and NWM-predicted runoff into channel reach (q.lateral) contributing the least.
Physics into Deep Learning Models
Physics into Deep Learning Models

Each neural connection in the DNN is now a mass (or energy) flux.
Physics into Deep Learning Models

A standard LSTM
Physics into Deep Learning Models

An “LSTM” with conservation laws

\[ \vec{x}_t = x_{t-1} + i_t - \vec{O}_t \]
Physics into Deep Learning Models

Table 1: Benchmarking Results. All values represent the median over the 447 basins.

<table>
<thead>
<tr>
<th>Model</th>
<th>MC(^a)</th>
<th>KGE(^b)</th>
<th>Bias(^c)</th>
<th>(\sigma_{vat})^d</th>
<th>(r^2)</th>
<th>FHV(^e)</th>
<th>FLV(^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deep Learning Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC-LSTM Ens.</td>
<td>yes</td>
<td>0.764*</td>
<td>-0.020*</td>
<td>0.842</td>
<td>0.873*</td>
<td>-14.689*</td>
<td>-24.651*</td>
</tr>
<tr>
<td>LSTM Ens.</td>
<td>no</td>
<td>0.762</td>
<td>-0.034</td>
<td>0.838</td>
<td>0.886</td>
<td>-15.740</td>
<td>36.267</td>
</tr>
<tr>
<td><strong>Conceptual Hydrology Models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAC-SMA</td>
<td>yes</td>
<td>0.632</td>
<td>-0.066</td>
<td>0.779</td>
<td>0.792</td>
<td>-20.356</td>
<td>37.415</td>
</tr>
<tr>
<td>VIC (basin)</td>
<td>yes</td>
<td>0.588</td>
<td>-0.018</td>
<td>0.725</td>
<td>0.760</td>
<td>-28.139</td>
<td>-74.769</td>
</tr>
<tr>
<td>VIC (regional)</td>
<td>yes</td>
<td>0.257</td>
<td>-0.074</td>
<td>0.457</td>
<td>0.651</td>
<td>-56.483</td>
<td>18.867</td>
</tr>
<tr>
<td>mHM (basin)</td>
<td>yes</td>
<td>0.691</td>
<td>-0.040</td>
<td>0.807</td>
<td>0.832</td>
<td>-18.640</td>
<td>11.433</td>
</tr>
<tr>
<td>mHM (regional)</td>
<td>yes</td>
<td>0.468</td>
<td>-0.039</td>
<td>0.589</td>
<td>0.793</td>
<td>-40.178</td>
<td>36.795</td>
</tr>
<tr>
<td>HBV (lower)</td>
<td>yes</td>
<td>0.391</td>
<td>-0.023</td>
<td>0.584</td>
<td>0.713</td>
<td>-41.859</td>
<td>23.883</td>
</tr>
<tr>
<td>HBV (upper)</td>
<td>yes</td>
<td>0.681</td>
<td>-0.012</td>
<td>0.788</td>
<td>0.833</td>
<td>-18.491</td>
<td>18.341</td>
</tr>
<tr>
<td>FUSE (900)</td>
<td>yes</td>
<td>0.668</td>
<td>-0.031</td>
<td>0.796</td>
<td>0.815</td>
<td>-18.935</td>
<td>-10.538</td>
</tr>
<tr>
<td>FUSE (902)</td>
<td>yes</td>
<td>0.690</td>
<td>-0.047</td>
<td>0.802</td>
<td>0.821</td>
<td>-19.360</td>
<td>-68.224</td>
</tr>
<tr>
<td>FUSE (904)</td>
<td>yes</td>
<td>0.644</td>
<td>-0.067</td>
<td>0.783</td>
<td>0.808</td>
<td>-21.407</td>
<td>-67.602</td>
</tr>
</tbody>
</table>

\(^a\)Mass conservation (MC).
\(^b\)Kling-Gupta Efficiency: \((-\infty, 1]\), values closer to one are desirable.
\(^c\)Bias: \((-\infty, \infty)\), values closer to zero are desirable.
\(^d\)Variance Ratio: \((-\infty, \infty)\), values closer to one are desirable.
\(^e\)Top 2\% high flow bias: \((-\infty, \infty)\), values closer to zero are desirable.
\(^f\)Bottom 30\% low flow bias: \((-\infty, \infty)\), values closer to zero are desirable.

Slight performance increase over LSTM, but currently the best peak-flow model we’ve tested.
Physics into Deep Learning Models

Snow represented as a sum over 4 states in the MC-LSTM.

Snow was not part of the training data set.
Contributors

Google Research
University of Alabama
Johannes Kepler University
Upstream Tech, PBC
University of Arizona
NASA

Frederik Kratzert
Daniel Klotz
Sepp Hochreiter
Craig Pelissier

Grey Nearing
Alden Sampson
Jonathan Frame
Hoshin Gupta