Streamflow Forecasting using Deep Learning

Zhongrun Xiang
PhD Candidate
Deep Learning models have been applied to hydrology and water resources fields since 2017.
The recurrent neural networks (RNNs) including LSTM and GRU can work on multiple tasks.

Inner algorithm of each green GRU cell:

\[ y_t = f_{\text{GRU}}(y_{t-1}, x_t) \]

\( f_{\text{GRU}} \) is the same in all timesteps.
Study Area and Data

All USGS gauged catchments in Iowa State
Green represents the upstream
Red represents the downstream

<table>
<thead>
<tr>
<th>Input Data</th>
<th>Data Type</th>
<th>Source</th>
<th>Spatial Resolution</th>
<th>Temporal Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>Stage IV multi-sensor measurement</td>
<td>NOAA</td>
<td>4km grid</td>
<td>hourly</td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>Empirical monthly data</td>
<td>Iowa Flood Center</td>
<td>State-based</td>
<td>monthly</td>
</tr>
<tr>
<td>Streamflow</td>
<td>USGS gauge measurement</td>
<td>USGS</td>
<td>Gauge-based</td>
<td>5-60 mins</td>
</tr>
</tbody>
</table>
Our study designed a Neural Runoff Model (NRM) for the runoff/streamflow forecast task.

- **Training dataset:** Water Year 2012-2017.
- **Evaluation dataset:** Water Year 2018.
- **Evaluation Metric:** Nash-Sutcliffe efficiency (NSE)

*Diagram showing the structure of the NRM model with layers and connections.*
Distributed structure adjusted the boundary
Reduced the error from unequal distribution of rainfall
Extra real-time data from upstream station
→ Better accuracies!!!
Results:
NSE of 120-hr ahead Predictions

NSE of 120-hr ahead Predictions

Persistence
Median NSE = -0.14

Ridge Regression
Median NSE = 0.63

Random Forest Regression
Median NSE = 0.56

Neural Runoff Model
Median NSE = 0.68

Neural Runoff Model - Distributed
Median NSE = 0.74
120-hr ahead streamflow forecast in time series

USGS 05389000 Yellow River near Ion

NSE values | Ridge | RF | NRM
---|---|---|---
Overall | 0.55 | 0.56 | 0.85
Peak (>Q90) | 0.44 | 0.21 | 0.74
Non-Peak | -1.31 | 0.31 | 0.64

Streamflow [m³/s]

Precipitation

Streamflow Observation

Ridge Regression

Random Forest Regression

Neural Runoff Model

Precipitation [m³/s]

2018/4/1 to 2018/10/1
120-hr ahead streamflow forecast in time series

<table>
<thead>
<tr>
<th></th>
<th>NSE values</th>
<th>Ridge</th>
<th>RF</th>
<th>NRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td>-0.76</td>
<td>-0.17</td>
<td>0.80</td>
</tr>
<tr>
<td>Peak (&gt;Q₉₀)</td>
<td></td>
<td>0.48</td>
<td>-0.49</td>
<td>0.75</td>
</tr>
<tr>
<td>Non-Peak</td>
<td></td>
<td>-28.63</td>
<td>-0.05</td>
<td>-0.49</td>
</tr>
</tbody>
</table>

USGS 05451700 Timber Creek near Marshalltown
Using the next 36 hours’ real-time rainfall forecast data from HRRR (High Resolution Rapid Refresh), we made our forecast model real-time.

In our preliminary study, deep learning models has the ability to forecast more accurate than National Weather Services AHPS (Advanced Hydrologic Prediction Service) and NWMv2.0 (National Water Model).