Predicting Flood Responses from Spatial Rainfall Variability and Basin Morphology through Machine Learning

Flood Response Characterization

- Characterize flood-generating flow responses with respect to:
  - Causating rainfall event:
    - Lag Time
  - Pre-defined discharge thresholds:
    - Flood Stage Threshold Exceedance (NWS)
- Information for Flash Flood Forecasting
- Flashiness relates closely to basin morphology
Precipitation Spatial Variability

- Rainfall spatial variability influences basin response [1,5,6,7,8].
- Spatial rainfall moments:
  - Dimensionless indices
  - Based on precipitation location.
  - Based on flow distance.
  - Describe the behavior of the storm event, relative to the basin.
- Statistical Moments:
  - Mean, Variance, Skewness...
  - Single and cross-moments.
**Data Set**

- **Event Dataset**
  - Morphological
  - Climatological
  - Precipitation Moments*
  - Streamflow

- **Morphological**
  - 21,143 Events

- **Climatological**

- **Precipitation Moments***
  - 21,143 Events

- **Streamflow**
  - Event Dataset
  - 133 variables

- **Kirstetter & Saharia**
  - 21,143 Events

- **Working Dataset**
  - 31 variables
  - 2 Responses

**POST-PROCESSING**

- **Data Transformation:**
  - Yeo-Johnson Transform
    - Similar to Box-Cox
    - Accepts negatives
    - Accepts zeros
  - Log + Yeo-Johnson

- **Variable Selection:** *Elastic Net* [9]
  - Penalized Maximum Likelihood Regression
  - Selection such that Error is within 1SE of optimum (minimum)
Some Variable Examples

Climatological

- Flashiness
- Mean Diurnal Temperature Range
- Mean Temp. of Warmest Quarter
- Isothermality (Mean Diurnal Range / Annual Range)
- Precip. Of Warmest Quarter
- Annual Mean Temperature

Morphological

- Outlet Slope
- Curve Number
- River Length
- Basin Imperviousness
- Slope Index

Precip. Moments

- Precipitation Mean
- Precip. Std. Deviation
- Precip. Skewness
- Flow Distance Mean
- Flow Dist. Std. Deviation
- Flow Dist. Skewness
- Precipitation x Flow Dist. Mean
- Precip. x Flow Dist. Std. Deviation
- Precip. x Flow Dist. Skewness
- Precip. x Flow Dist. Kurtosis
Response Modeling

- Dataset divided randomly into 75% - 25% training and validation subsets:
  - Both subsets showed similar distributions for the data.
  - Preserved representativeness in both subsets.
- Three different approaches:
  - MARS: Multidimensional piecewise linear fits.
  - Support Vector Machines: Kernel-based spatial transformations (radial basis function).
  - Random Forest: Randomized, bagged tree ensemble.
- All training subject **10 x 10-fold cross-validation**.
- Validation with unseen data **v.s.** cross-validated training.
- **Variable importance analysis** performed on models.

Images taken from [2, 3, 4]
Results - Training Times

<table>
<thead>
<tr>
<th></th>
<th>MARS</th>
<th>Random Forest</th>
<th>Support Vector Machines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag Time</td>
<td>1.76 hours</td>
<td>8.20 hours</td>
<td>344.13 hours</td>
</tr>
<tr>
<td>Flood Stage Threshold Exceedance</td>
<td>9.5 hours</td>
<td>3.0 hours</td>
<td>~150 hours</td>
</tr>
</tbody>
</table>

○ What are the relationships and trade-offs between ML technique, training time and performance?
MARS Results - Validation

Lag Time

Flood Stage Threshold Exceedance

<table>
<thead>
<tr>
<th>Prediction</th>
<th>No Action</th>
<th>Action</th>
<th>Minor</th>
<th>Moderate</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Action</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Action</td>
<td>13</td>
<td>123</td>
<td>45</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>Minor</td>
<td>1</td>
<td>5</td>
<td>21</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Moderate</td>
<td>2</td>
<td>30</td>
<td>86</td>
<td>164</td>
<td>86</td>
</tr>
<tr>
<td>Major</td>
<td>0</td>
<td>14</td>
<td>55</td>
<td>238</td>
<td>3258</td>
</tr>
</tbody>
</table>

- Performance remains consistent between training and validation, for both responses.
- Class imbalance - favors **Major** events.
- Sensitive to **Action** and **Major** events

VALIDATION "CC 0.64" V.S. TRAINING "CC 0.66"

VALIDATION "ACCURACY 0.84" "KAPPA 0.58" V.S. TRAINING "ACCURACY 0.85" "KAPPA 0.5288"
RF Results - Validation

Lag Time

Validation Lag Time (transformed)

Validated Lag Time (transformed)

Flood Stage Threshold Exceedance

<table>
<thead>
<tr>
<th>Prediction</th>
<th>No Action</th>
<th>Action</th>
<th>Minor</th>
<th>Moderate</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Action</td>
<td>9</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Action</td>
<td>12</td>
<td>132</td>
<td>63</td>
<td>33</td>
<td>28</td>
</tr>
<tr>
<td>Minor</td>
<td>0</td>
<td>16</td>
<td>56</td>
<td>30</td>
<td>13</td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>16</td>
<td>65</td>
<td>210</td>
<td>109</td>
</tr>
<tr>
<td>Major</td>
<td>0</td>
<td>6</td>
<td>23</td>
<td>183</td>
<td>3218</td>
</tr>
</tbody>
</table>

- Performance lost between training and validation, for both responses.
- Even using 10x10-fold CV!

Class imbalance - favors Major events.
Sensitive to Action and Major events
Better “understanding” of Moderate events than MARS

EGU General Assembly 2020 - ITS4.6/NH6.7-D2391, Wed. May 6

Jorge Duarte – jduarte@ou.edu
SVM Results - Validation

Lag Time

![Graph showing SVM results for validation and training datasets.]

<table>
<thead>
<tr>
<th>Prediction</th>
<th>No Action</th>
<th>Action</th>
<th>Minor</th>
<th>Moderate</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Action</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Action</td>
<td>5</td>
<td>61</td>
<td>19</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Minor</td>
<td>0</td>
<td>8</td>
<td>27</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Moderate</td>
<td>0</td>
<td>11</td>
<td>35</td>
<td>97</td>
<td>43</td>
</tr>
<tr>
<td>Major</td>
<td>12</td>
<td>93</td>
<td>126</td>
<td>330</td>
<td>3310</td>
</tr>
</tbody>
</table>

**VALIDATION**

"CC 0.64"

**TRAINING**

"CC 0.82"

**Flood Stage Threshold Exceedance**

- Performance lost between training and validation, for both responses.
- Even using 10 x 10-fold CV!

- Class imbalance - favors **Major** events.
- Sensitive ONLY to **Major** events.

EGU General Assembly 2020 - ITS4.6/NH6.7-D2391, Wed. May 6

Jorge Duarte – jduarte@ou.edu
## Results - Summary

<table>
<thead>
<tr>
<th></th>
<th>MARS</th>
<th>Random Forest</th>
<th>Support Vector Machines</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lag Time</strong></td>
<td>1.76 hours</td>
<td>8.20 hours</td>
<td>344.13 hours</td>
</tr>
<tr>
<td></td>
<td>CC - 0.64</td>
<td>CC - 0.66</td>
<td>CC - 0.64</td>
</tr>
<tr>
<td><strong>Flood Stage</strong></td>
<td>9.5 hours</td>
<td>3.0 hours</td>
<td>~150 hours</td>
</tr>
<tr>
<td><strong>Threshold</strong></td>
<td>ACC. - 0.84 K - 0.50</td>
<td>ACC. - 0.85 K - 0.57</td>
<td>ACC. - 0.82 K - 0.34</td>
</tr>
<tr>
<td><strong>Exceedance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Class imbalance makes it difficult to provide accurate predictions for intermediate flood stage thresholds.
  - Prediction skill for Major events.
- All models exhibit comparable skill levels.
  - MARS performs consistently across all data, and requires less training time for Lag Time.
  - RF overfits training data, but performs comparably.
  - SVM requires extensive parameter tuning and training times, no drastic performance gains.
### Lag Time - Variable Importance Ranking

<table>
<thead>
<tr>
<th>Lag Time</th>
<th>MARS</th>
<th>Random Forest</th>
<th>Support Vector Machines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Flashiness</td>
</tr>
</tbody>
</table>

- Precipitation moments as well as flashiness appear to play a decisive role across all three algorithms.
  - Flashiness as an abstracted measure of morphological and climatological information.
- Models consistently rank Precip. x Flow Dist. Mean, Precipitation Mean and Flashiness as significant variables.
- Differences in variable ranking order imply that the choice of ML approach potentially impacts the underlying understanding of the processes being modeled.
### Flood Stages - Variable Ranking

<table>
<thead>
<tr>
<th>MARS</th>
<th>Random Forest</th>
<th>Support Vector Machines</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Flashiness</td>
<td>1. Curve Number</td>
<td>1. River Length</td>
</tr>
<tr>
<td>2. Precip. X Flow Dist. Mean</td>
<td>2. Flow Distance Mean</td>
<td>2. Flow Distance Mean</td>
</tr>
</tbody>
</table>

- Both precipitation moments and morphological variables appear to hold relevance in performing classification.  
  - Thresholds are directly related to morphology and may be influenced by climatology.  
- Climatological variables appear as well, most notably measures of temperature and precipitation.  
- Differences in variable importance and ranking order imply that the choice of ML approach impacts the underlying understanding of modeled processes.
Conclusions

- Characterization of floods was achieved by training machine learning models.
  - Data-driven approach for variable selection using Elastic Net.
- Catchment-scale precipitation moments were effectively used to model Lag Time and Flood Stage Threshold Exceedance.
- Variable importance showed relevant factors that contribute to characterizing both responses:
  - Rainfall moments and flashiness lead the ranking for Lag Time.
  - Aside from moments, there is an intrinsic dependence on morphology and climatology for flood stage thresholds.
- MARS was the most consistent performer of all three approaches, and the best at predicting Lag Time.
- RF is more efficient at classification, but overfitting affects consistency.
Thank you!

jduarte@ou.edu
Image Sources and References


2. https://www.researchgate.net/profile/Bernd_Freimut/publication/2455400/figure/fig1/AS:341659566002178@1458469396649/l.png


