30min, 1hr, and 6hr Time - Series Prediction of White Oak Bayou, Houston Texas

Machine Learning Approach to modeling hydrographs

By - Roberto Vega

Date: 9/23/18
Flood Prediction Neural Network

Introduction

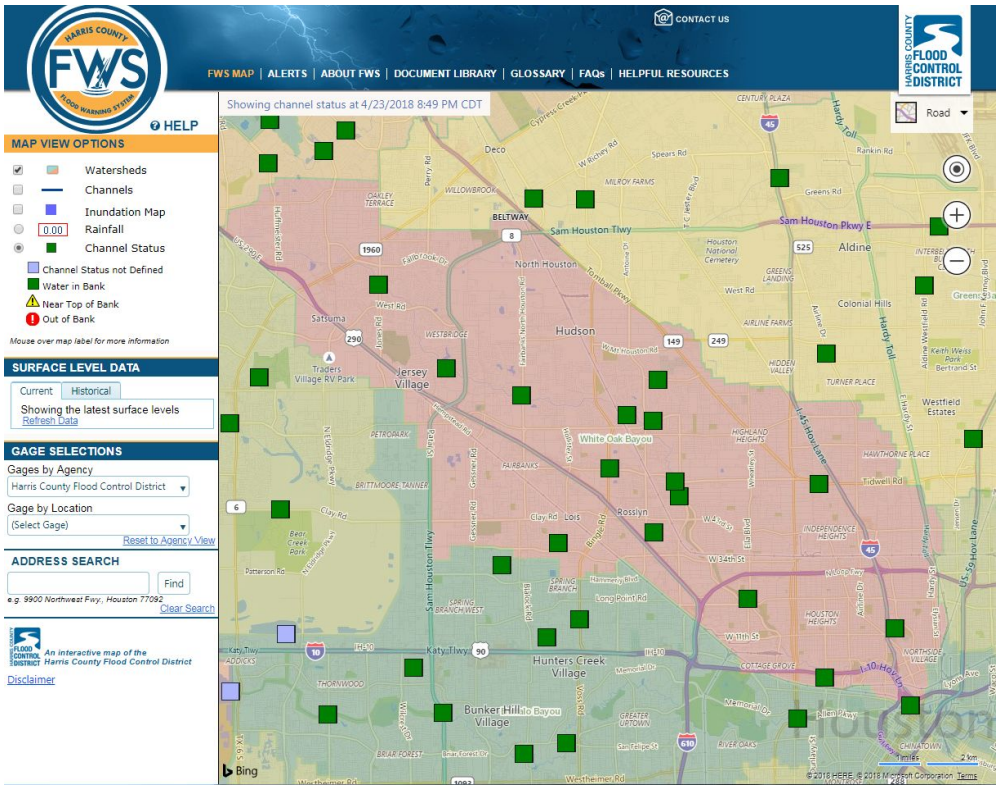
This lab tries to predict the future stream elevation, using one of two options, Ground Truth Validation data or user supplied rainfall in the White Oak Bayou watershed.

Because I will be working with sequences of data, I will use a Long Term Short Term Memory Network (LSTM). I will use gage data from the 2015 - 2018 time frame, thich encompasses Memorial Day, Tax Day, and Harvey storm events. The raw-data is 38.9 mbs. The Data is populated from 26 recorded rainfall and stream elevation events; resulting in 154,868 data points.

I trained two networks one network phase shifted by 30 minutes to see how accurate the network is on a short temporal shift. I then trained the network on a 6 hour phase shift to take past rainfall data and output the future 6hrs. I accomplish the phase shift by adjusting the output ground truth by 6 hours. Additionally the Flood Prediction Dashboard allows the user to add rainfall to the test batch as can be seen on the final prediction depicted below.

Location

Using Harris County Flood Control District stream gage data from White Oak Bayou Watershed in Houston, Texas.



Gage Information

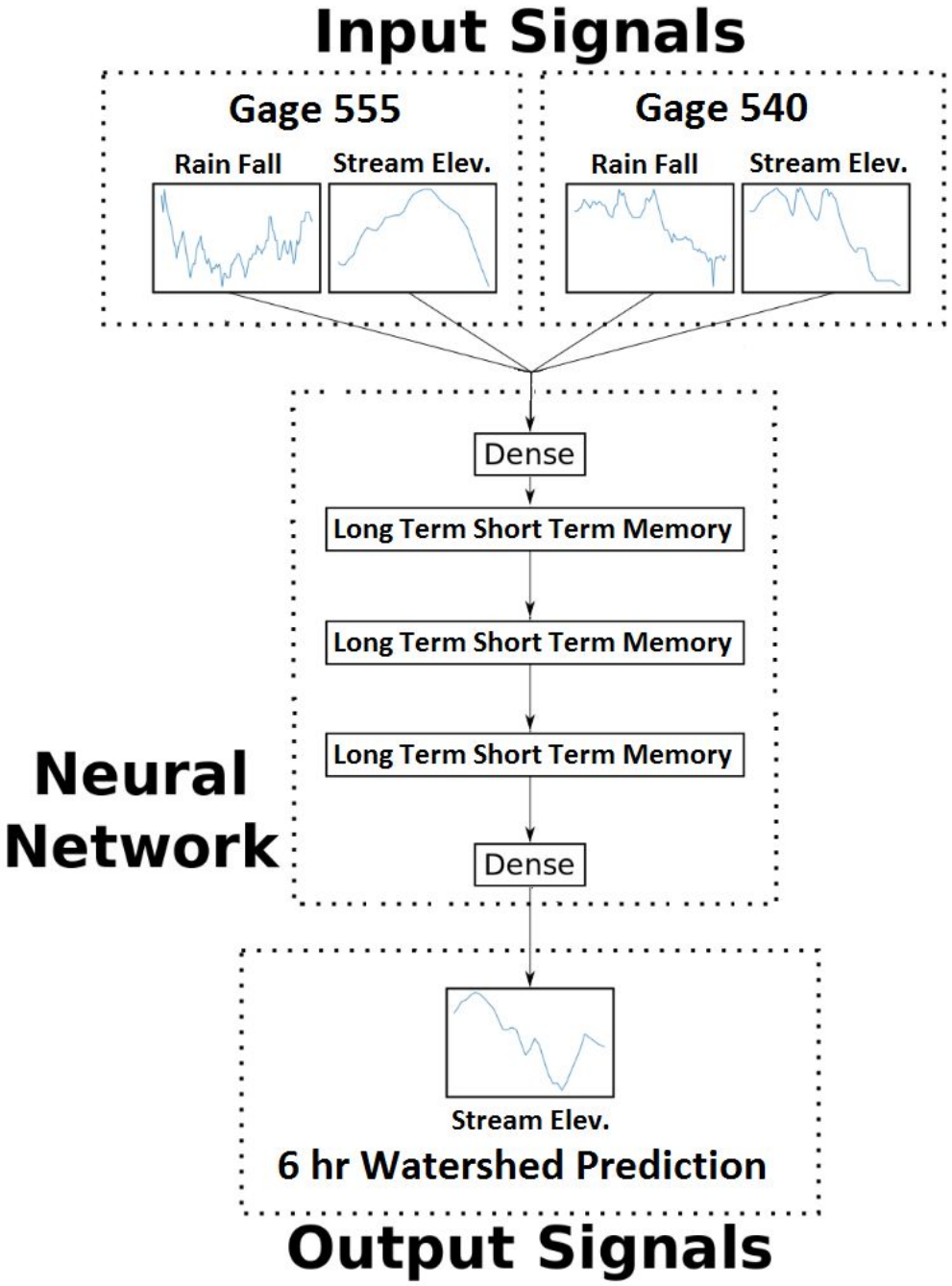
Site Id	Desc	Bottom of Channel
555	WOB @ Jones Road	99.12
550	WOB @ Lakeview	83.31
545	WOB @ Fairbanks North Houston Road	70.13
595	Vogel Creek @ Bank Road	65.55
590	Cole Creek @ Deihl Road	60.26
585	Vogel Creek @ Victory Drive	57.18
540	WOB @ Alabonson Road	57.81
570	Little WOB @ Tidwell Road	51.98
575	WOB @ Tidwell Road	44.53
580	Brickhouse Gully @ Costa Rica Road	47.24
535	WOB @ Pinemont Drive	40.52
560	Little WOB @ Trimble Street	18.63
520	WOB @ Heights Boulevard	6.59
2220	Buffalo Bayou @ Milam Street	-6.71

Flowchart

In this lab I am trying to predict the stream elevation at 14 gage locations for the White Oak Bayou Watershed, 6 hrs into the future, given the current and past stream-data (The Flowchart below only shows two locations).

I use a Long Term Short Term Memory Network (LSTM) to work on the sequences of arbitrary length. During training I will use a sub sequence of 12 time steps (6 hours; 30 min intervals) from the training-set, with each data-point or observation having 28 input signals for the rainfall, and Stream Elevation for each of the 14 gages. I then train the network to output the stream elevation for the 14 gage locations.

Additionally I can create predictions on the watershed for stream elevations using arbitrary rainfall data and stream elevation prediction to predict theoretical flood water elevation rises.



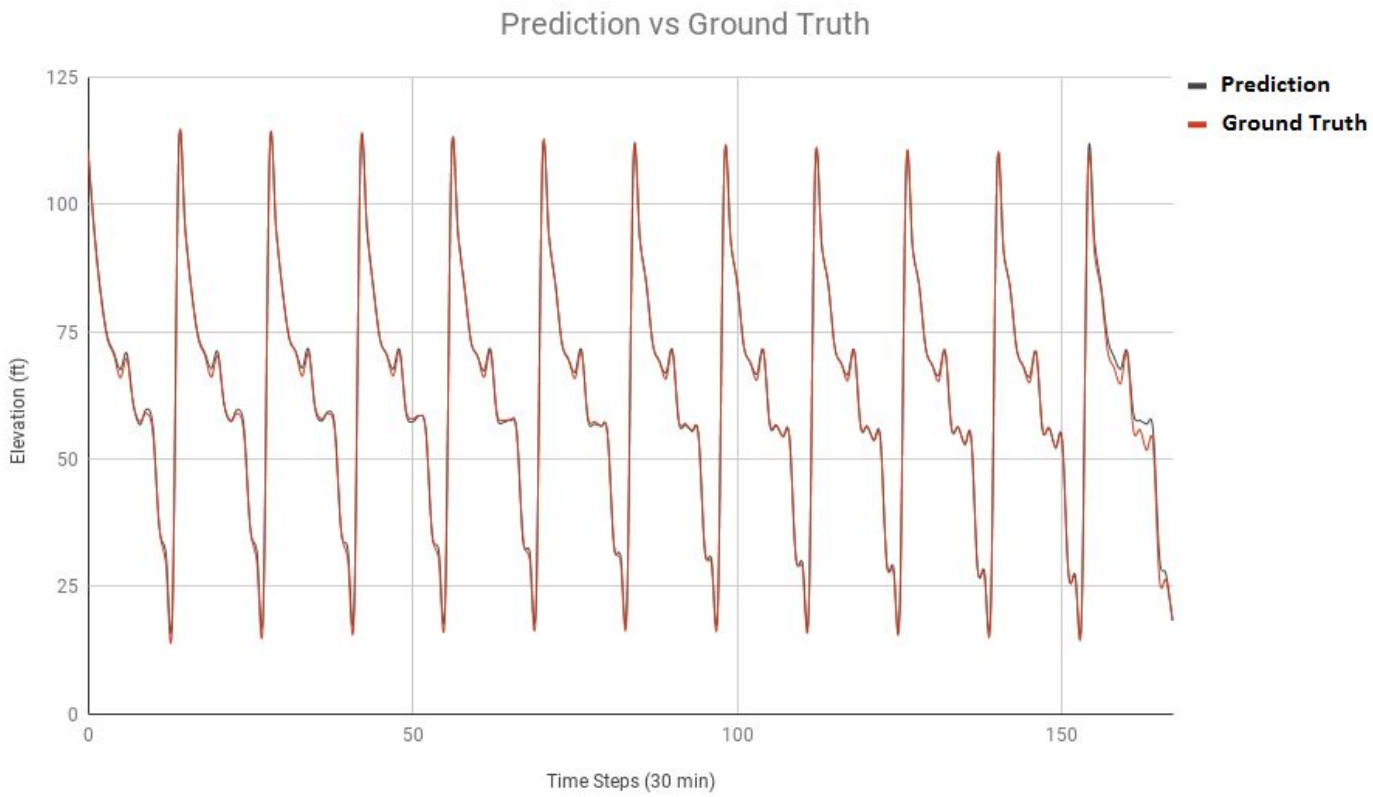
Output Predictions

Using 154,868 Training Data points the Flood Prediction Network will produce near real world accuracy.

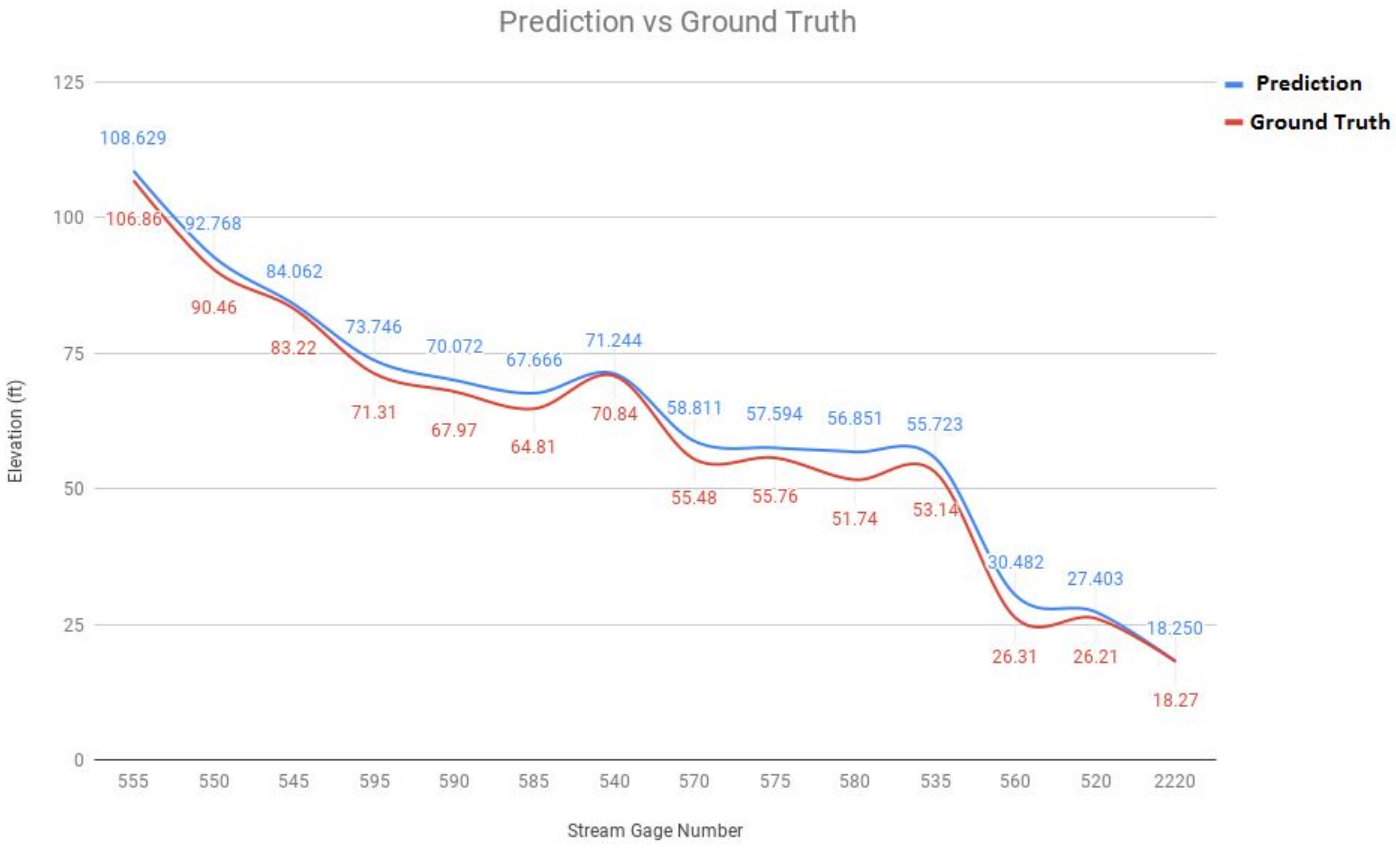
Giving the Flood Prediction Network 12 sample time steps (6hrs) of real world data the network will produce an output 6 hour flood elevation prediction. In the 6hr prediction graph (below) I add arbitrary rainfall data to the last 30min prediction, this results in the final prediction being performed on theoretical rainfall.

In order to test the neural network I separated Tax Day Flood Event to utilize as a validation and Testing Set. This data the network has never seen. I take a 6hour sample of rain and creek data and plot the prediction from the network.

Depicted below is a 30 min phase shift.

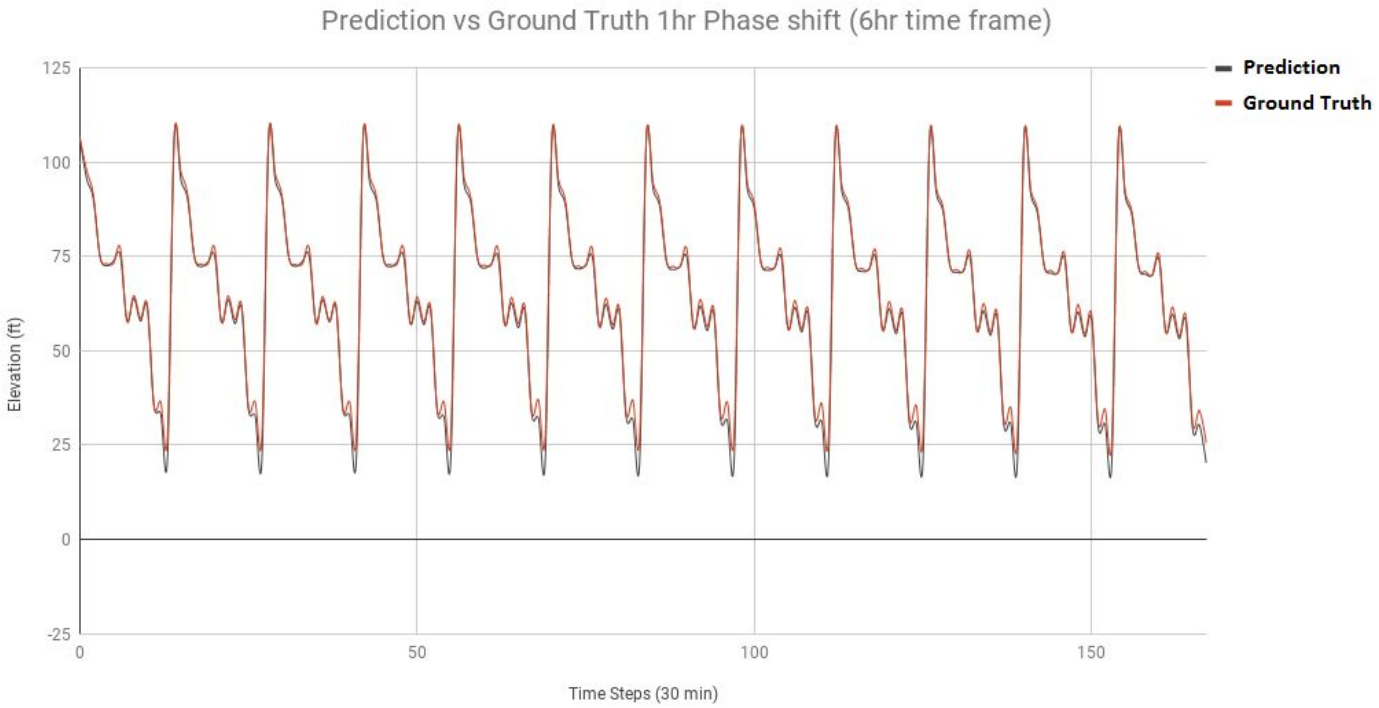


As close up of the final 30min prediction (below)

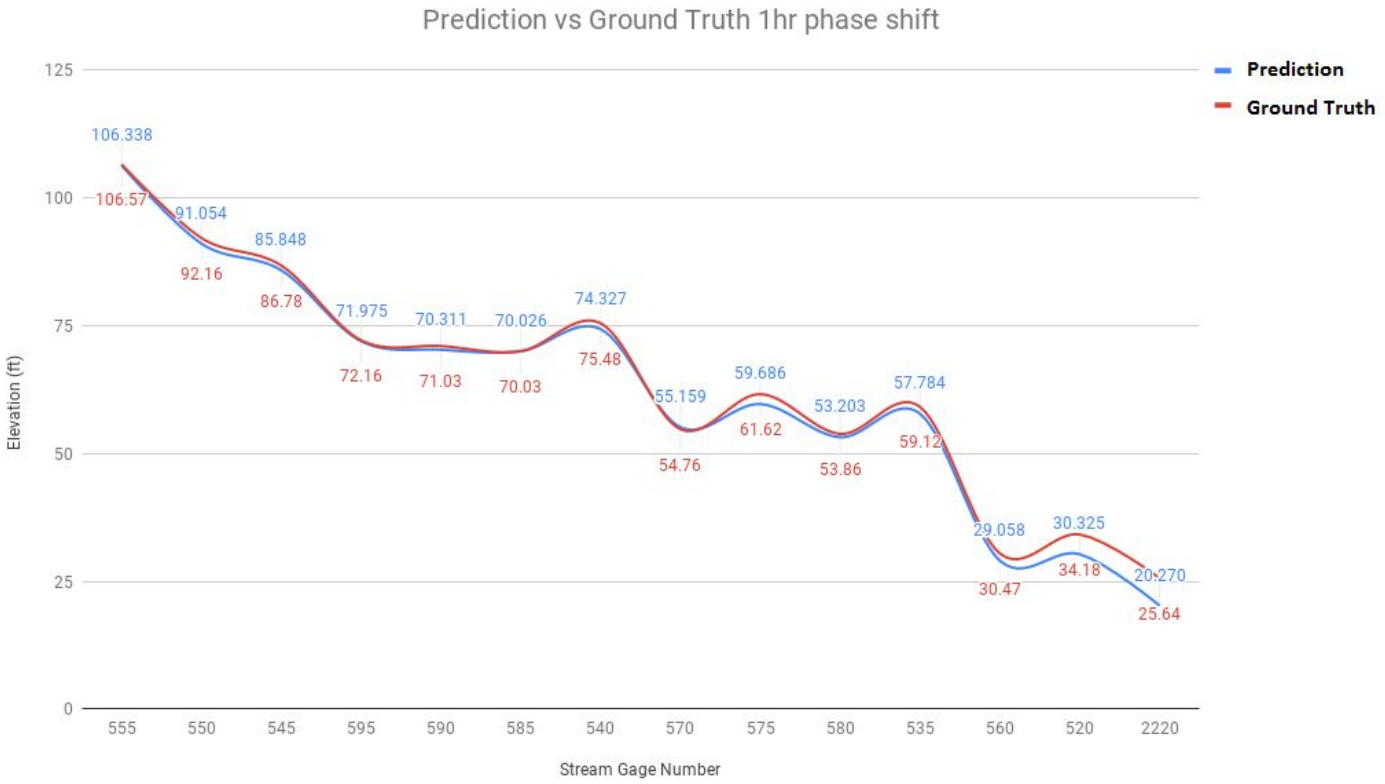


The Final Stream Elevation Prediction Data shows the hydrograph for White Oak bayou with the user specified rain vs how the data specified.

Depicted below is a 1hr phase shift

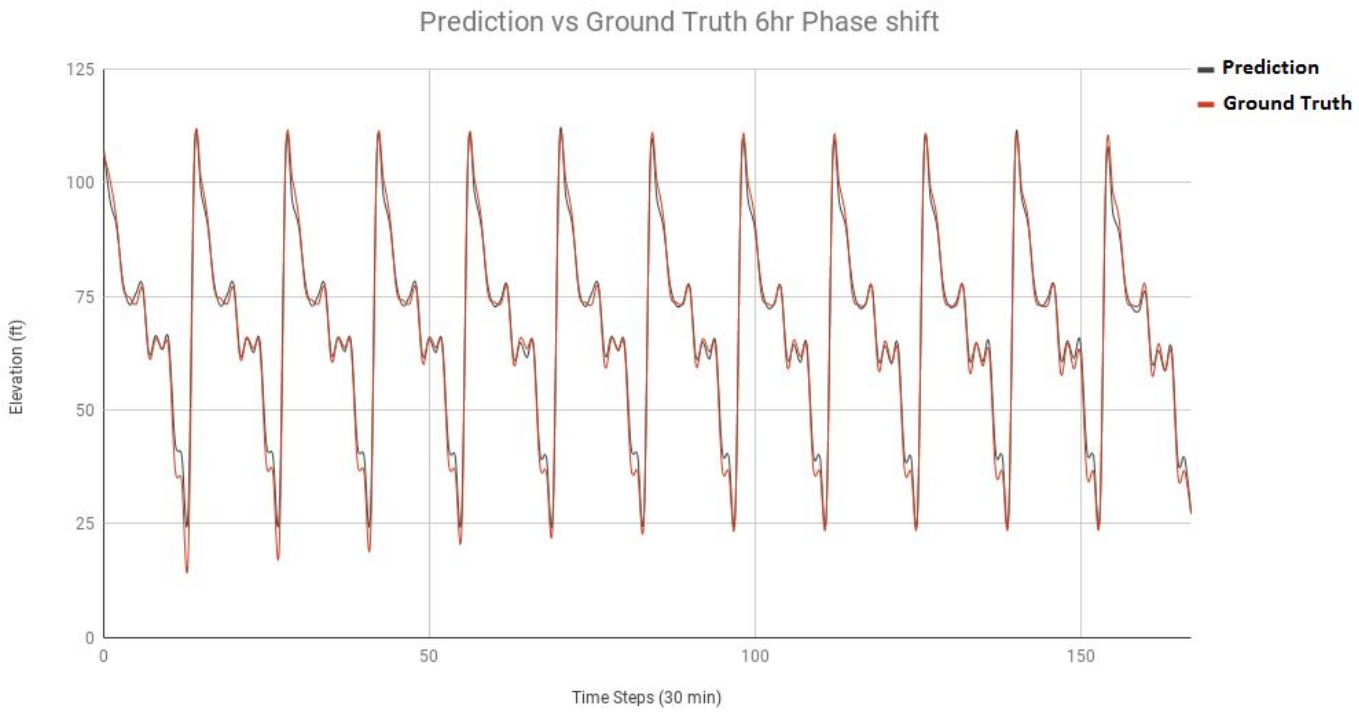


A close up of the final 30min prediction

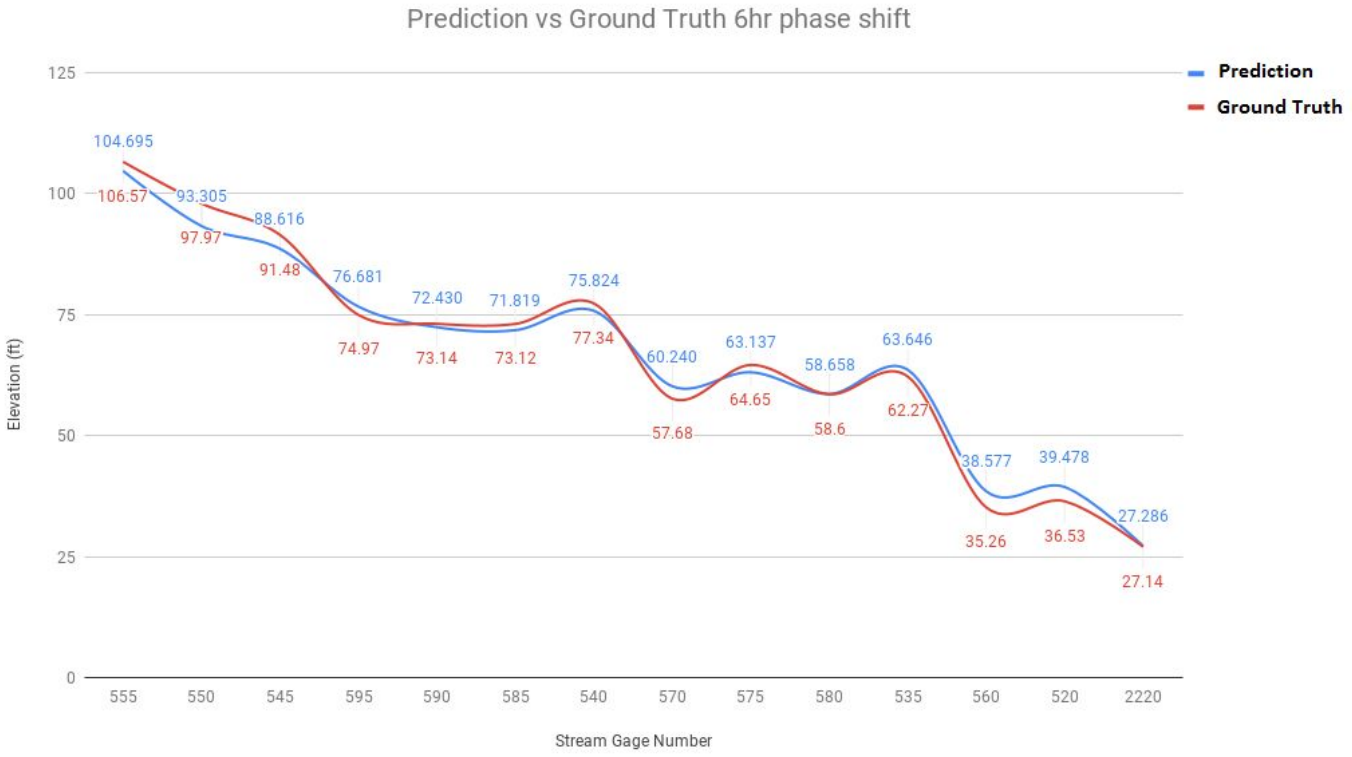


The Final Stream Elevation Prediction Data shows the hydrograph for White Oak bayou with the user specified rain vs how the data reacts. You can see that the model is starting to predict lower than the actual data as a result of the 0 inches of rainfall supplied by the user.

Depicted below is a 6hr phase shift



Below is the final 30min prediction



Usage

Using the network I can generate a 6hr prediction of stream elevations using user specified rainfall intensity. The rainfall events will cause stream elevation reactions.

I can extend the functionality of the Network by incorporating detention data to determine what volume of detention would affect the hydrograph negatively and positively.

To increase the accuracy It would be necessary to collect the distance between gages, and slope. Adding slope information would allow the network to refine the time of concentration and clean up variances in gages that are immediately following an abrupt slope change. Detention could also be added to refine the model further and add the ability to analyze basins in specific locals and how they affect the hydrograph.

Lab Equipment

Custom Modeling Dashboard depicted below Private License
Tensorflow.Js v0.13.0 Copyright 2018 Apache V2.0 January 2004
Chrome 69.33443 Open License
NodeJS v 8.6.4 Open License
Express.JS v4.0 Open License