A Comparative Study of the National Water Model Forecast to Observed Streamflow Data

CE394K GIS in Water Resources

Term Project Report

Fall 2018

Leah Huling

Introduction

As global temperatures increase, the volatility of localized weather patterns increases. Precipitation distribution becomes more severe, leading to storm and flooding events with higher frequency and greater intensity (Armel et al, 2018). In order to adapt to this change, there is a need for the scientific community to advance our methods of predicting events of high precipitation and stream flows. This would assist in preparation for large coastal storm events, such as hurricanes, as well as localized flash floods. Emergency response systems require both projected and current flooding information in order to protect the people at risk during these events.

In August 2016, the National Oceanic and Atmospheric Administration (NOAA) released a new modeling system named the National Water Model (NWM) that revolutionized the precision and scope in which precipitation and streamflow in the continental United States can be predicted. The NWM uses a vast array of inputs, such as stream gages, atmospheric modeling, soil moisture content, and Radar data, to forecast the flow of rivers and streams at 2.7 million stream locations nationwide. The forecasts are classified into four categories: analysis and assimilation, short-range, medium-range, and long-range forecasts. The analysis and assimilation step in the NWM operates as an hourly snapshot of the current hydrological conditions at the time of forecast. Observed data such as meteorological radar, 8,000 USGS gages, and water monitoring from 1,506 reservoirs is forced into the algorithms to provide the most accurate estimate of current conditions. This snapshot is then used as a baseline to initialize the three other types of forecast.

This study examines the accuracy of the National Water Model through a comparison to a network of stream gages outside the USGS gage network. There are currently 8,000 USGS gages and 2.7 million reach catchments within the NWM, leaving the vast majority of stream reaches without the real-time correction of stage data. Therefore, it is critical that the NWM provides acceptable estimates of stream flow, even when not forced with streamflow observations. This study focuses on the National Water Model reach within the city of Austin, Texas. The Texas Department of Transportation (TxDOT) maintains a network of 44 stream gages throughout Austin, 32 of which are located on NWM stream reaches. A sample size of 19 gages had continuous data for the target date range of January 2017 to September 2017 (Figure 1).

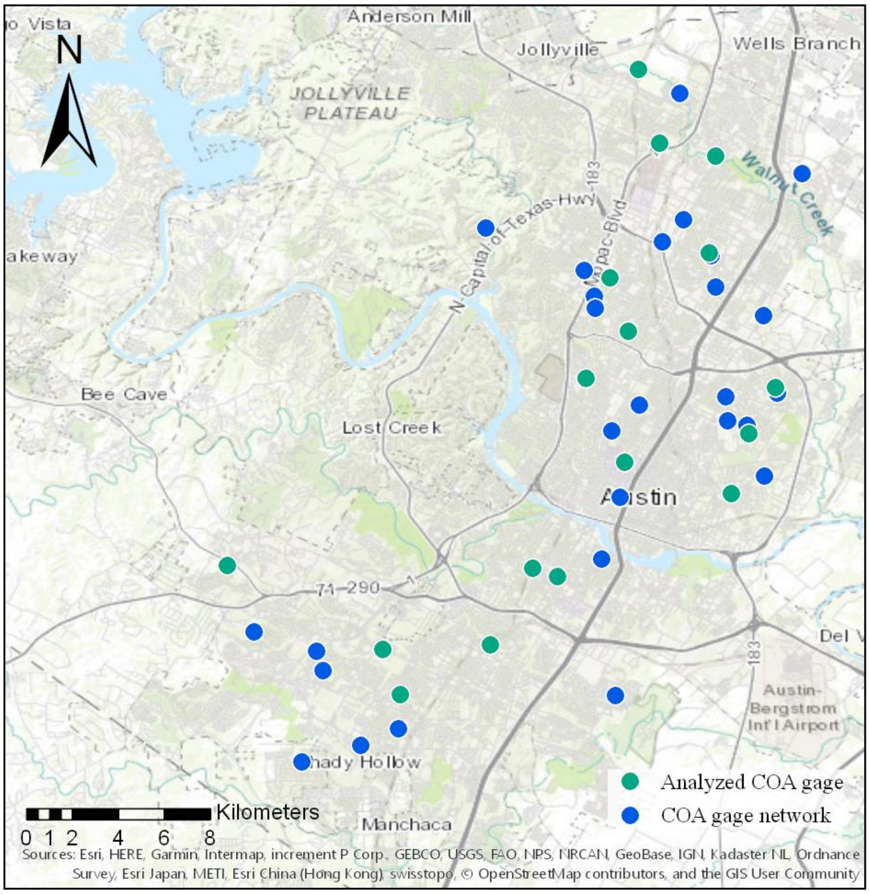


Figure 1. Network of total City of Austin gages. The gages analyzed in this study are illustrated in green.

The objective of this comparison will be to analyze how the NWM performs with respect to observed flow data in stream reaches with and without USGS gages. Ultimately, this study will qualitatively and quantitatively evaluate the National Water Model within a small study area and search for methods of improving the forecasting and forcing process for the NWM.

Methods

<u>Data Analysis</u>

City of Austin (COA) gage data is managed by the data information company KISTERS. Matt Ables of KISTERS created a tool that performs Height Above Nearest Drainage (HAND) calculations to transform the NWM flow rate outputs into equivalent stage heights at the locations of the City of Austin gages. HAND is a method of converting flow rate data to equivalent stage heights (Nobre et al, 2011).

In order to retrieve both the COA and NWM gage data, a python code was developed to perform an API call to the KISTERS database. The COA gages collect data every 15 minutes and the NWM every hour. Therefore, it was necessary to sort and select COA data from the same hourly timestamp as NWM outputs. Once the two data sources were temporally matched, the stage data could be directly compared (Figure 2-a) and compared over time (Figure 2-b).

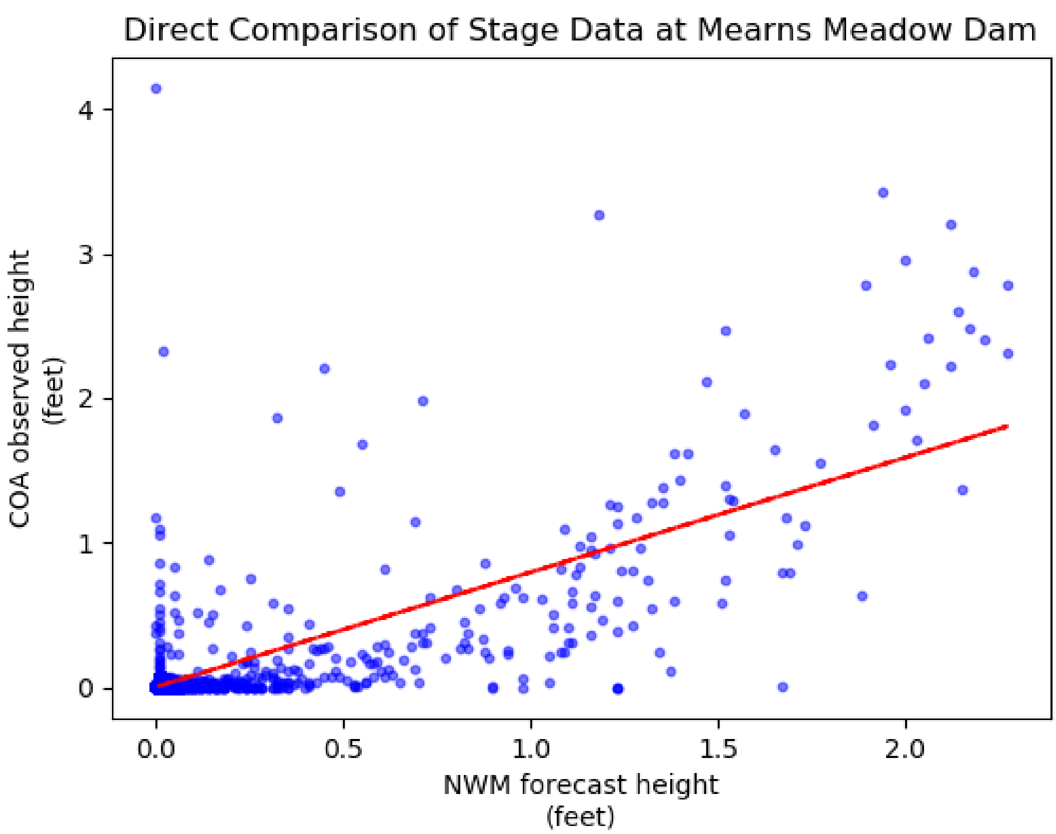


Figure 2-a. Direct comparison of observed stage height to the predicted National Water Model stage height. The trendline of the relationship indicates a positive correlation.

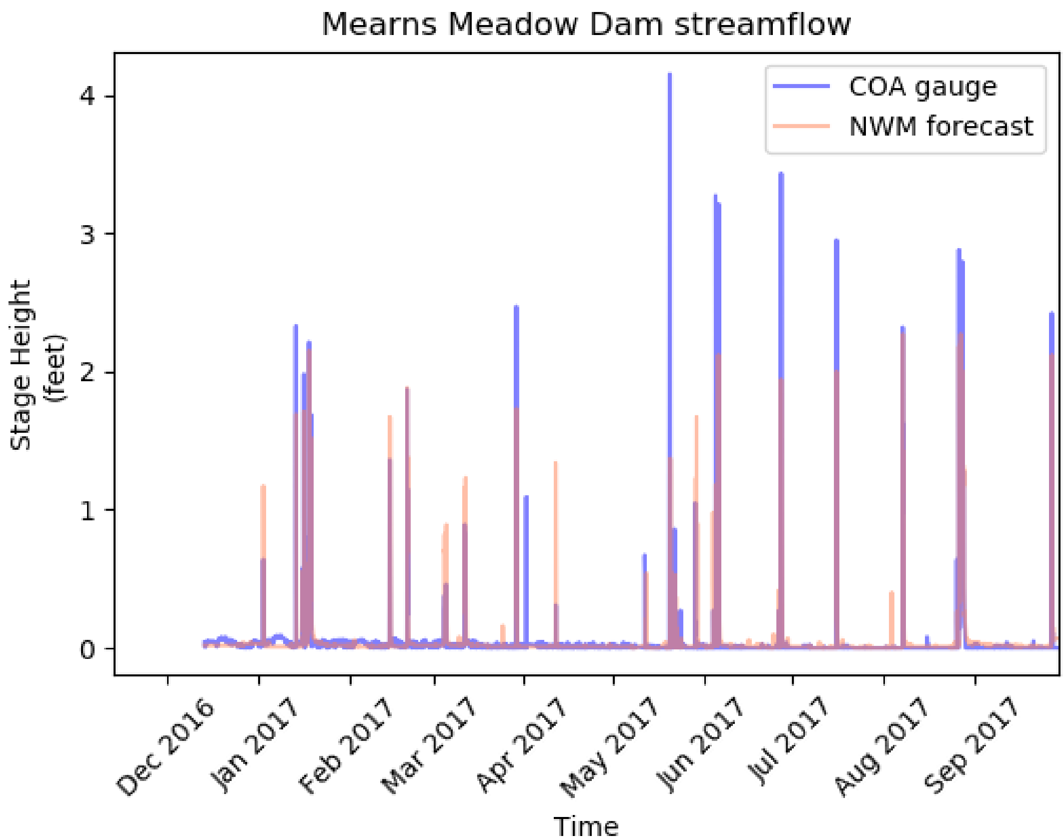


Figure 2-b. National Water Model and City of Austin stage heights over the time of study.

Method of Comparison

The Pearson correlation coefficient was used to measure the correlation between the National Water Model and the observed stage data. The formula is as follows:

$$r = \frac{\Sigma(X_i - \overline{X})\,(Y_i - \overline{Y})}{[\Sigma(X_i - \overline{X})^2\,\Sigma(Y_i - \overline{Y})^2]^{1/2}}.$$

This correlation coefficient centers and standardizes the raw data points by subtracting the mean of each data set and adjusts the scales of the datasets to have equal units (Rodgers and Nicewander, 1988). This was an essential qualification for the correlation analysis of the datasets because two of the COA gages have baseline measurements above zero. This was most likely because the gages are buried in debris or something similar. By definition of the Pearson correlation coefficient, 1 is a perfectly positive correlation, -1 is a totally negative correlation, and 0 indicates no correlation.

The downsides to the Pearson correlation coefficient are: 1) if neither dataset varies over the given timeframe, no coefficient is given; 2) if one dataset is zero for the given timeframe, no coefficient is given; and 3) the python application for this operation occasionally returns “inf” and “-inf” for non-ideal data. In this study, if any of these issues occurred, no correlation coefficient was applied to the return matrix.

Two storm events that registered at all gage locations were chosen for correlation analysis. The first event, “Storm A,” was simple with one streamflow surge over the span of approximately 48 hours (Figure 3). The second storm event, “Storm B” was more complicated, with three defined flow surges over 6 days (Figure 4).

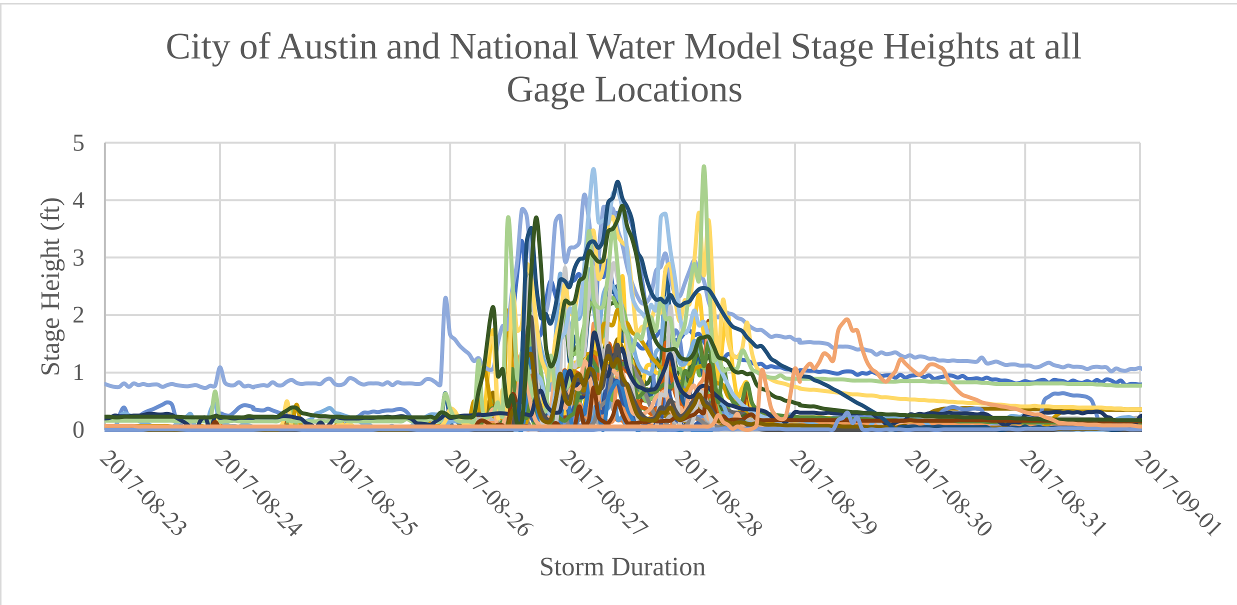


Figure 3. Visual representation of Storm A duration. Data from all stage locations and both resource forms is included to characterize the total storm behavior.

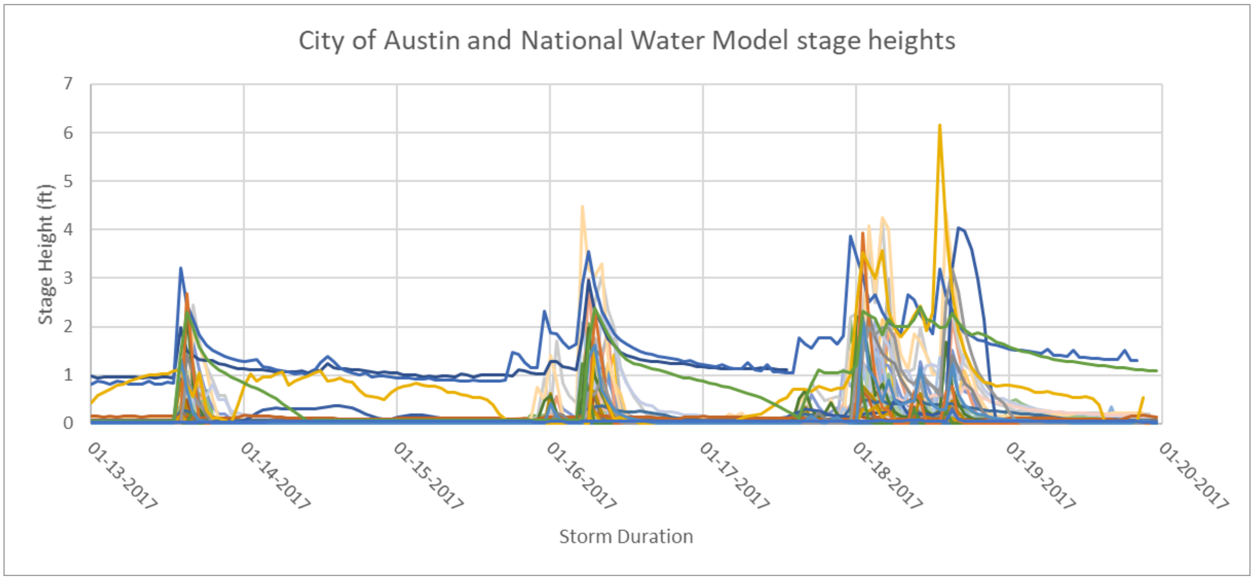
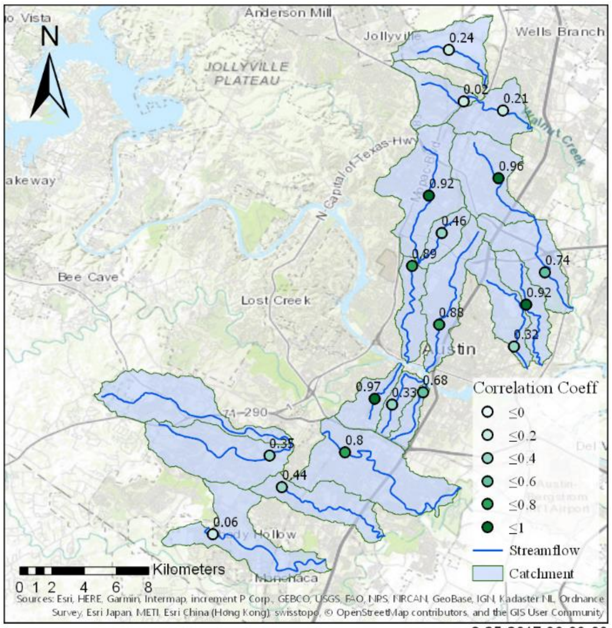
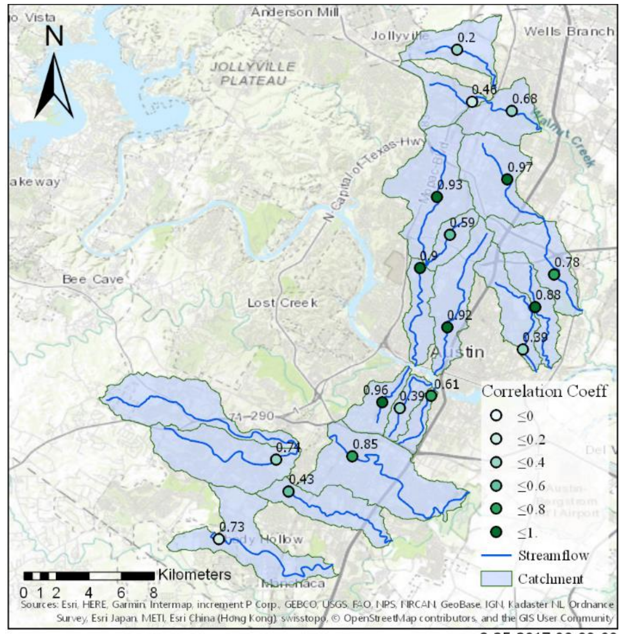


Figure 4. Visual representation of Storm B duration. Data from all stage locations and both resource forms is included to characterize the total storm behavior.

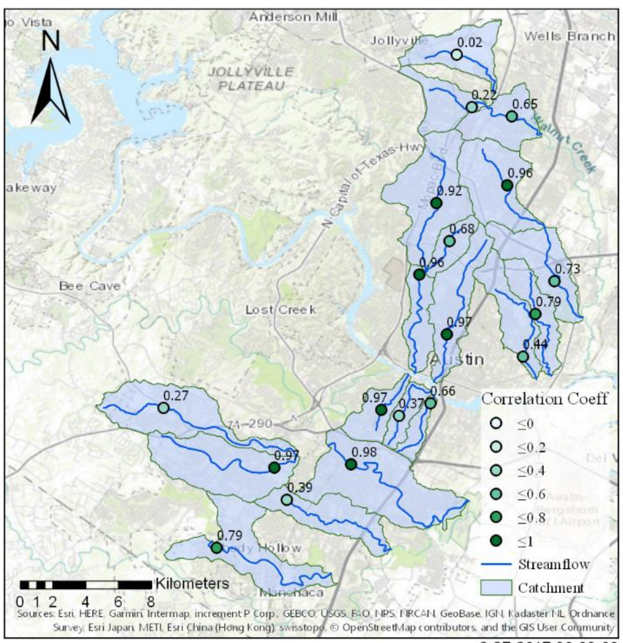
The correlation coefficients were calculated with a 48-hour rolling window for every hour within a week of the storm events. For example, at the timestamp 1-13-2017 00:00, a correlation coefficient was calculated from that point through the next 48 hours. Then the next correlation would be calculated at 1-13-2017 01:00 through 1-15-2017 01:00, and so on. To analyze how the National Water Model forecasts behaved both spatially and throughout the storm events, correlation coefficients were chosen at all 19 gage locations every 24 hours at the 00:00 timestamp for 5 days. For Storm A, the correlation coefficients were collected from 8-25-2017 to 8-29-17. For Storm B, coefficients are collected from 1-15-2017 to 1-19-2017. The correlation coefficients spatial distribution is illustrated in Figures 5 and 6.



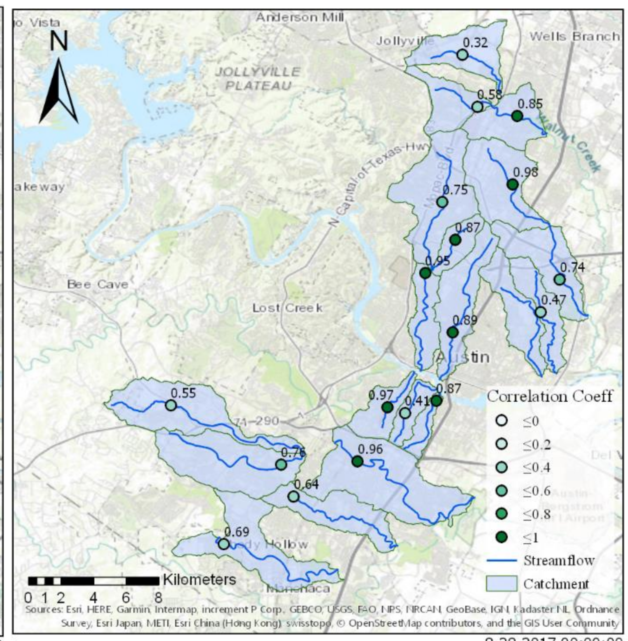
8-25-2017 00:00:00



8-25-2017 00:00:00



8-27-2017 00:00:00



8-28-2017 00:00:00

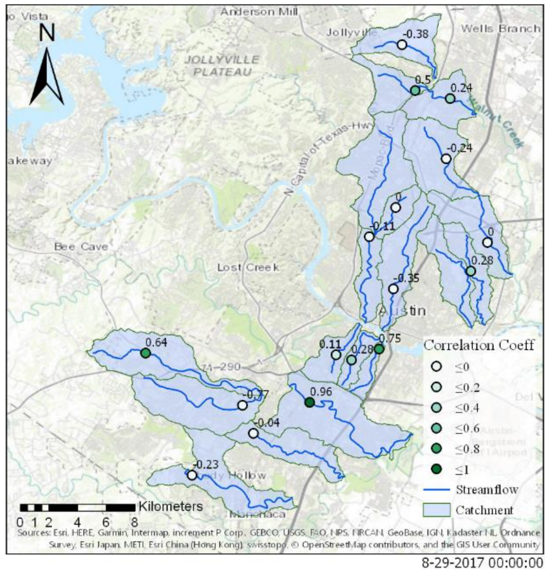
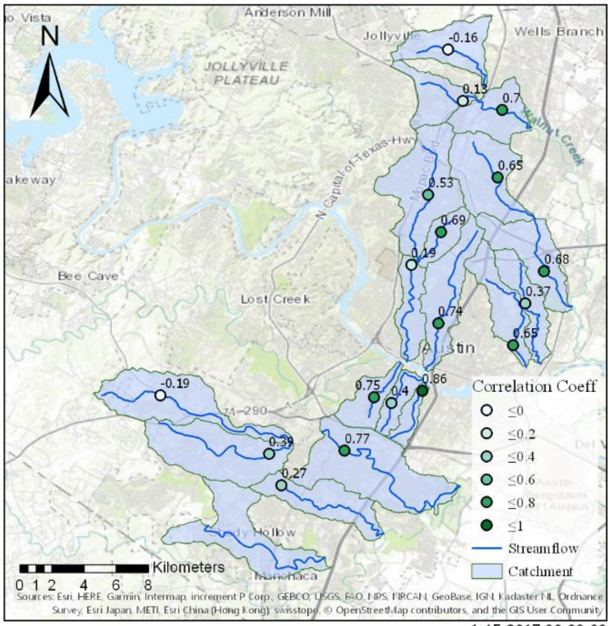
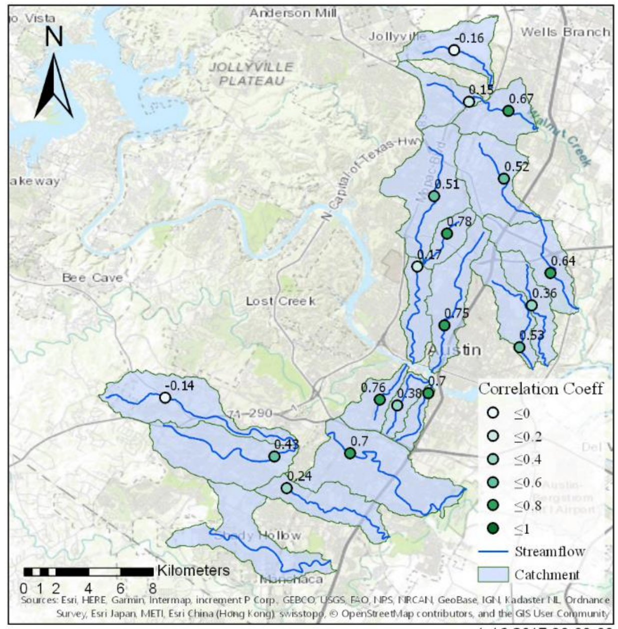


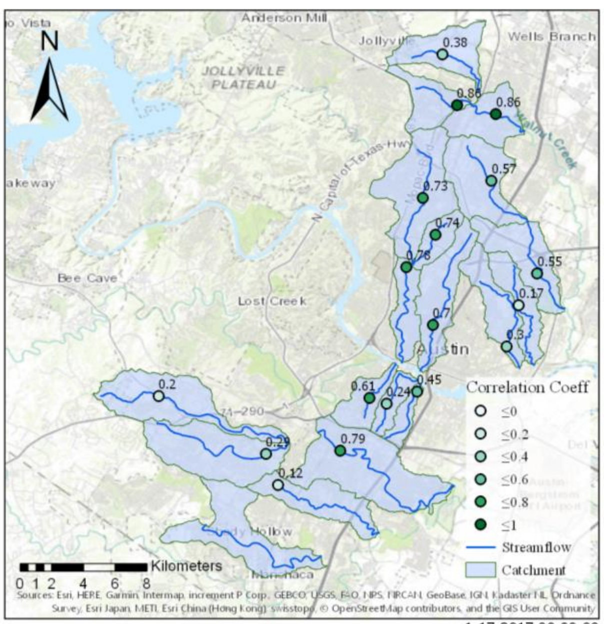
Figure 5. Example of correlation coefficient spatial distribution of Storm A through the 5 sampled timestamps.



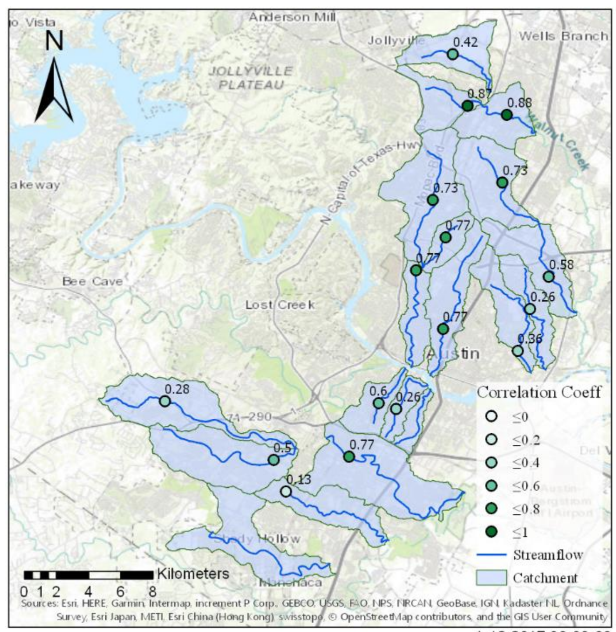
1-15-2017 00:00:00



1-16-2017 00:00:00



1-17-2017 00:00:00



1-18-2017 00:00:00

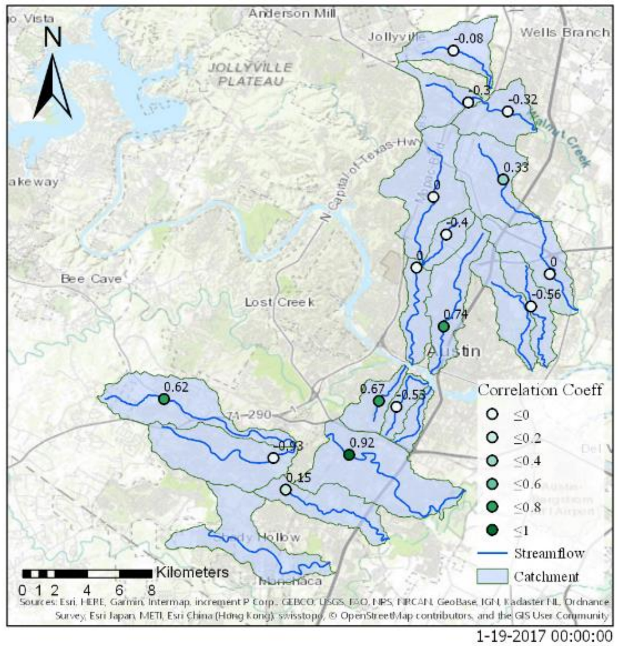


Figure 6. Series of correlation coefficient spatial distribution of Storm B through the 5 sampled timestamps.

Programming Process

In order to achieve these figures, the following process was implemented:

1. *Python script*: A storm time frame was isolated at each gage location and the correlation coefficients were calculated. An Excel spreadsheet was then generated with the stage heights and correlation coefficients.
2. *Excel to Table:* Excel spreadsheets containing the geographical coordinates and correlation coefficients of each gage location were imported into the GIS geodatabase.
3. *XY Table to Point:* Point features were created out of the coordinates of the at the gage locations.
4. *Symbology:* The visual representation in the GIS maps was manipulated by classifying the color and labels by the correlation coefficients for each timestamp.

Results

General Pattern

The correlation coefficients for both storms displayed a similar overall trend. The coefficients leading up to and during the storm events largely indicated a positive correlation. However, during the end of and immediately following a significant storm event, most correlation coefficients took a considerable dip in otherwise steady correlation coefficient values. In Storm A, this occurred in the 8-29-19 timestamp and in the 1-19-2017 timestamp for Storm B. The decrease was significant, with 37% of the correlation coefficients across both storms switching to negative values. The average difference between the other 4 correlation coefficients and the one decreased value was 0.57 and 0.47 for storm A and storm B respectively. Figure 7 illustrates the discrepancies of the correlation coefficients for the three flow surges of Storm B.

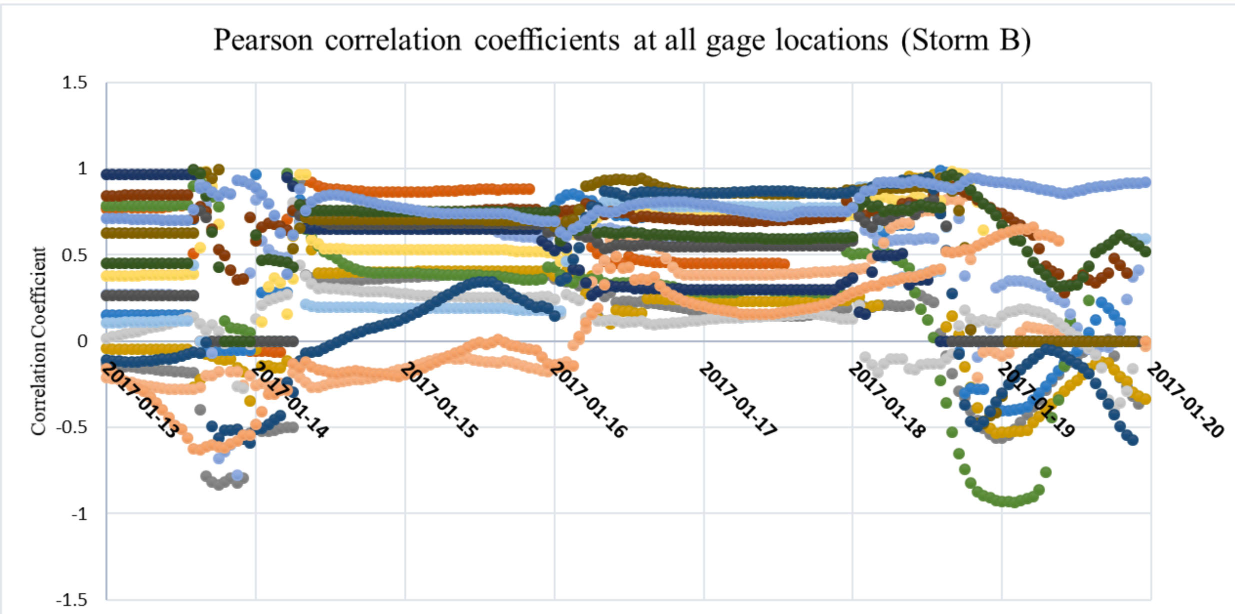


Figure 7. Correlation Coefficients for all City of Austin gage locations for Storm B. This clearly demonstrates the atypical behavior of the coefficients at the end of storm events.

Because of this anomalous behavior, the two timestamps above were excluded from further data analysis and will be discussed in the conclusions.

<u>Storm A</u>

The data for storm A was strongly positively correlated. For this simple isolated storm event 79% of all locations had an average correlation coefficient above 0.5, with a lowest average of 0.2 and greatest of 0.97. Figure 8 illustrates the distribution of coefficients for each timestamp as well as the average.

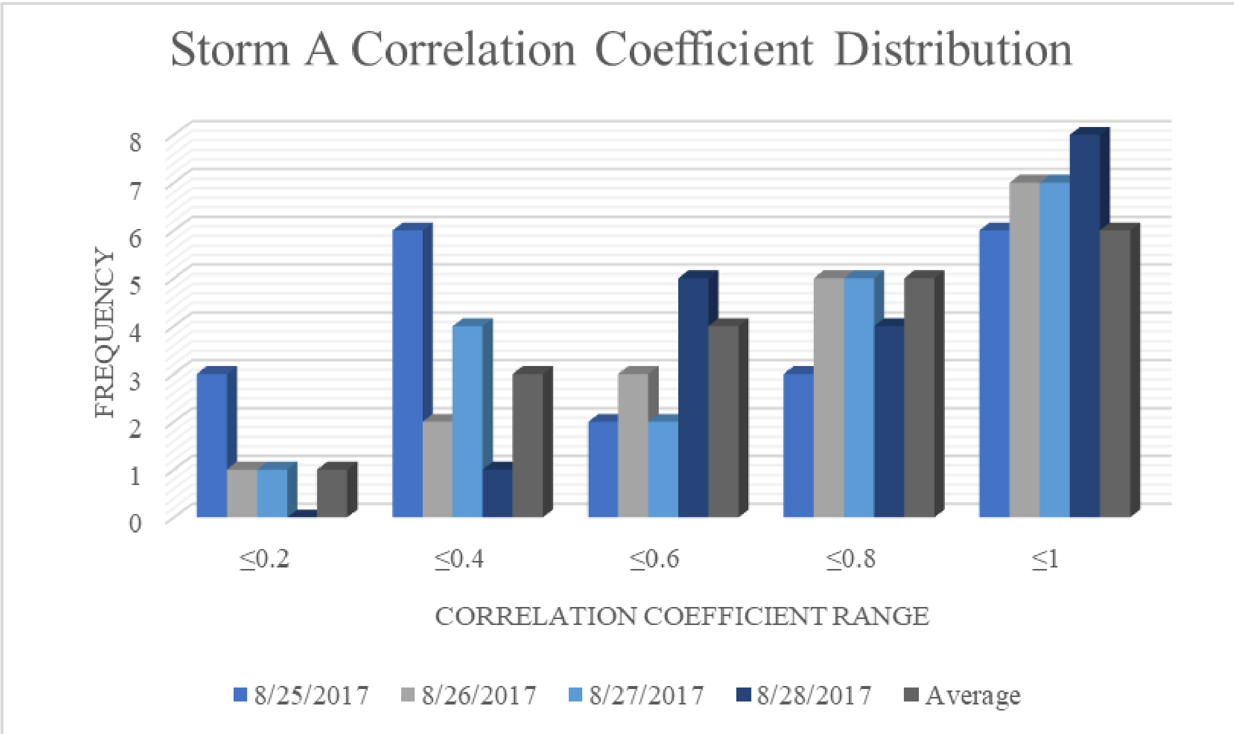


Figure 8. Distribution of Storm A coefficients. The results indicate a majority of the data points fall within regions of significant correlation.

Storm B

The data for Storm B was positively correlated, but not as strongly as in Storm A. 59% of the gage locations had an average correlation coefficient above 0.5, with the lowest average of 0.038 and the greatest average of 0.78. Four individual correlation coefficients were less than 0. One gage location (Slaughter Creek) registered no observable increase of flow but the NWM predicted a small storm surge (Figure 9). This resulted in no correlation coefficients for the entire storm event at this location.

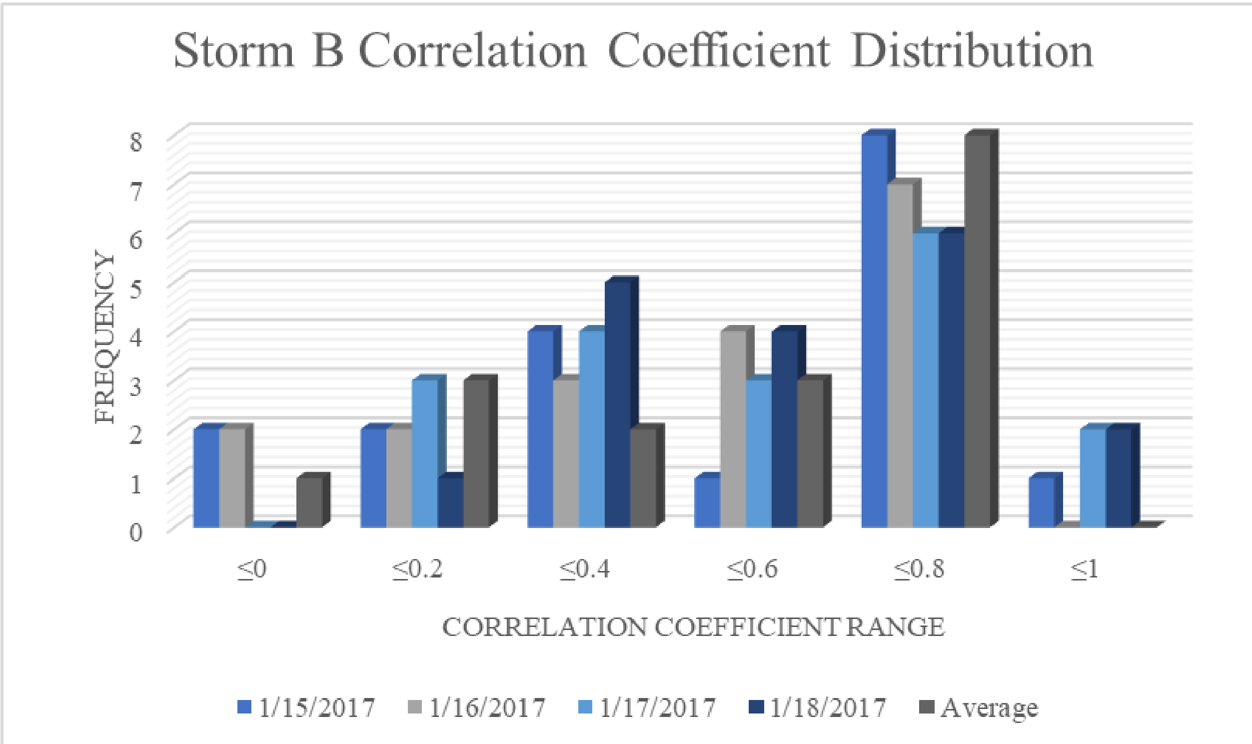


Figure 9. Frequency distribution of correlation coefficients across the 5 timestamps and overall average coefficient values.

USGS stream gages

There are 20 USGS stream gages within the catchments of interest for this study area. Thirteen City of Austin gages are on stream reaches that also have a USGS gage, and 5 of those have the USGS gages upstream of the COA gage.

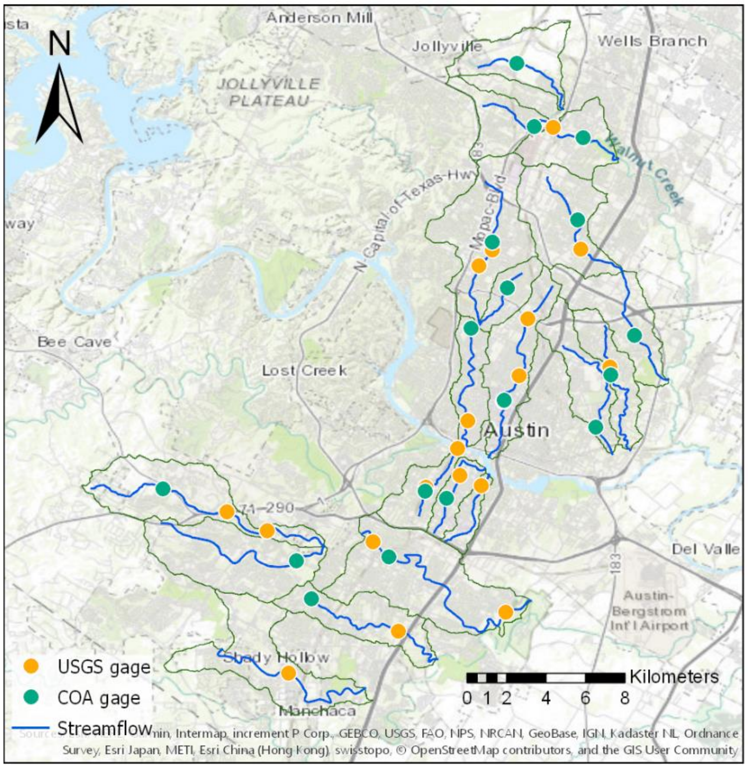


Figure 10. Spatial relationship of City of Austin gages to USGS stream gages within study area.

The data indicate a positive relationship between the correlation coefficients and the presence of USGS gages. The average Pearson correlation coefficient for a COA gage with a USGS gage upstream is 0.69, with a USGS gage downstream is 0.58, and with no USGS gage on the reach is 0.47.

Table 1. Correlation coefficients with respect to the presence of USGS stream gages.

Correlation Coefficients of locations relative to USGS gages			
	USGS gage upstream	USGS game downstream	No USGS gage
Storm A	0.81	0.68	0.46
Storm B	0.57	0.47	0.49
Average	0.69	0.58	0.47

Within the upstream and downstream datasets, there was no significant correlation between the distance to a USGS gage and the coefficient value.

Catchment size

The catchments in the study area varied from 6.82 km^2 to 26.9 km^2. To examine the relationship to the correlation coefficients within the catchments, the catchments and correlation coefficients were distributed evenly across six intervals so that an even number of observations fell within each interval (Figure 11). There was no meaningful relationship between the two variables, with the lowest average correlation coefficients frequently within the largest catchments and vice versa.

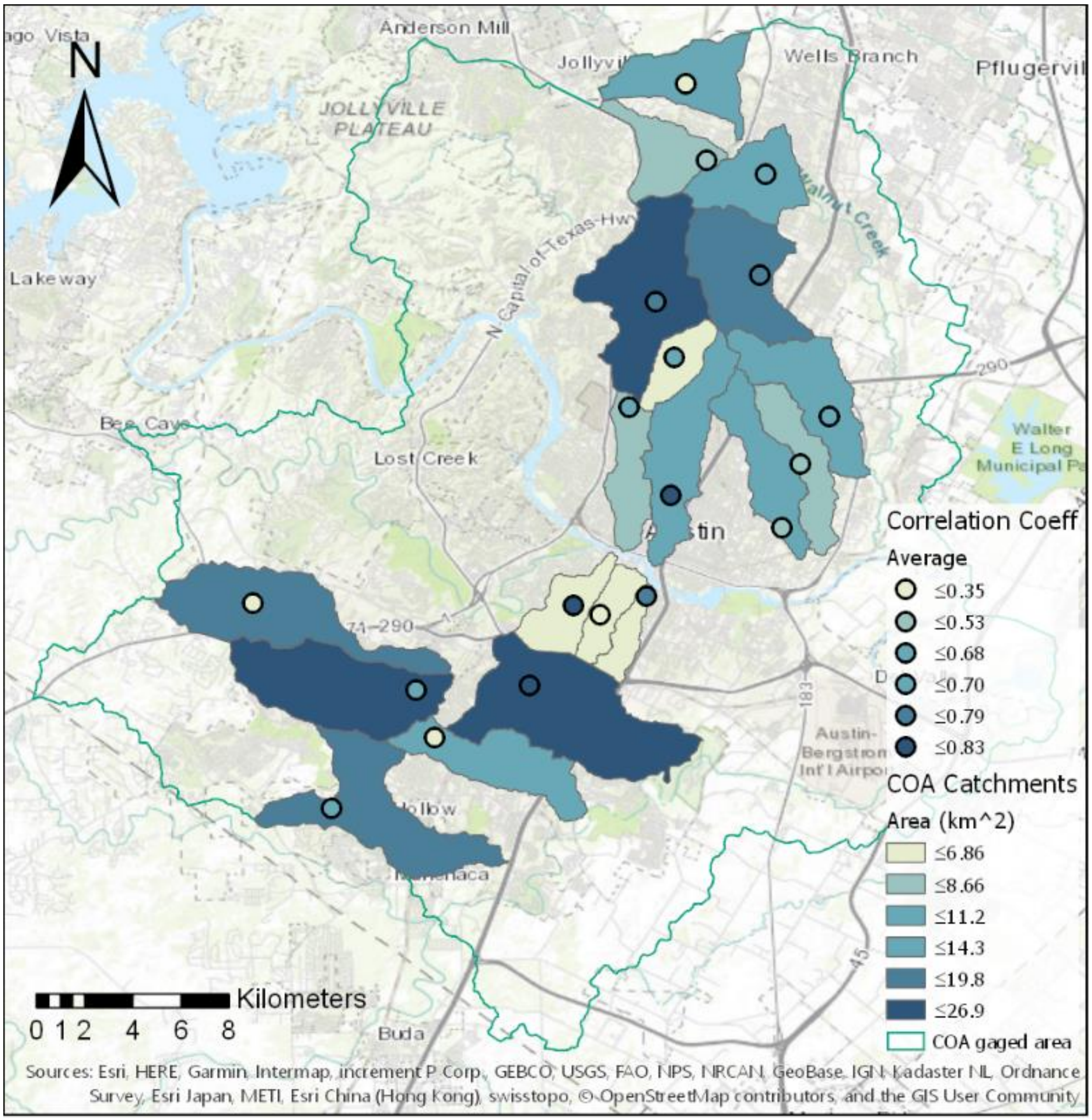


Figure 11. Correlation between catchment area and Pearson Correlation Coefficient. The ranges for both variables are evenly split into six classes.

Conclusions

The National Water Model analysis and assimilation outputs were positively correlated with observed stream flow data. The 48-hour range Pearson correlation coefficients averaged between 0.44 to 0.82 leading up to and during significant storm events. On the downhill side of the increased flows, the correlation coefficients behaved erratically. This was most likely because the trailing end of a storm is difficult to predict, and the Pearson correlation coefficient performs poorly in times of low to no flow. An example of a typical data point at no flow is a NWM forecast of 0 ft and an observational reading of 0.1 or 0.02. Although the difference is physically insignificant, if the reading is repeated for a large portion of the 48 hours, the correlation coefficient formula registers a poor correlation. Further investigation into the data is needed to confirm this theory.

The National Water Model more accurately estimated stream flows for a simple single storm event like Storm A. Two gage locations, Mearns Meadow Dam and West Bouldin Creek at Oltorf, registered an average correlation of 0.97. The series of flow surges in Storm B registered lower correlation values. This may present an issue for prolonged storm events of multiple rain spells.

Spatially, the positive correlation was greater in the central cluster of gage locations. The gages located on the peripheral edges of the study area chiefly represented the lower half of the correlation values. This could be because USGS gages are more densely located in the central area of the city. This would further vindicate the USGS relation data from Table 1 that shows COA gages with USGS gages located upstream within the same reach have an average correlation coefficient 0.22 greater than those with no USGS gages within the reach.

In evaluation of the National Water Model, this study concludes that, with the exception of a few outliers, the analysis and assimilation outputs are satisfactorily correlated to physical observed data. The accuracy of the prediction suffers when no USGS stages are present to actively correct the NWM algorithms. In order to progress the National Water Model, it is proposed that in the analysis and assimilation step is forced with stream gage networks outside the USGS systems. Several major metropolitan areas maintain valuable gage network resources that are currently unutilized by the NWM. An implementation of this policy would likely increase the accuracy of all National Water Model predictions and further the overall goal of improving analysis and saving lives in emergency flood response.

References

Armal, S., N. Devineni, and R. Khanbilvardi. 2018. Trends in extreme rainfall frequency in the contiguous United States: Attribution to climate change and climate variability modes. Journal of Climate 31(1):369– 385. https://doi.org/10.1175/JCLI-D-17-0106.1

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., ... Saleska, S. (2011). Height Above the Nearest Drainage - a hydrologically relevant new terrain model. Journal of Hydrology, 404(1-2), 13-29.

Rodgers, J; Nicewander, W. (1988). Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, 42(1) 59-66.