# Hydrologic Database, Statistical Analysis, and Adjusted Historical Flow Development of Select Surface Water Stations on the Lower Santa Fe and Ichetucknee Rivers

**Prepared For:**

**Suwannee River Water Management District**

**Prepared By:**

INTERA Geosciences & Engineering



**April 23, 2012**

# Contents

## List of Figures

## List of Tables

## Executive Summary

The relative contributions of anthropogenic and climatic factors to recent decreases in streamflow in the Lower Santa Fe River Basin are not well understood because they cannot be directly measured. In order to investigate this relationship, trend analysis was conducted for long-term stream monitoring stations within the Santa Fe River Basin. Trend analysis of rainfall within the basin was also conducted in order to investigate the long-term relationship between rainfall and streamflow.

Prior to trend analysis, a comprehensive flow and stage database for all gauging stations in the Lower Santa Fe River basin was constructed. Gaps in the time series were filled using acceptable methods-of-practice, including interpolation, linear regression, multiple linear regression, and artificial neural networks. This database served as the source data for all unimpacted flow analysis and will be utilized by the Suwannee River Water Management District (the District) in support of MFL development.

Exploratory data analysis was completed on rainfall, long-term streamflow records and baseflow estimates throughout the Lower Santa Fe River Basin. The LOWESS analysis identified periods of monotonic trends in the data. Results of this analysis showed that many stations experienced changes in trends around 1970. Trend analysis of the stations revealed the presence of statistically significant decreasing trends in both total flow and baseflow, particularly for the post-1970 data.

In order to further explain the changes in flow over the period of record for each gauge, linear regression (LR) models were constructed for several stations in order to describe the relationship between baseflow and rainfall during the period of the record assumed to be minimally influenced by groundwater withdrawals (pre-1970). Overall, a LR model provided a good fit for estimating baseflow using rainfall as the independent variable(s). Application of the flows generated by the pre-1970 models using post-1970 rainfall data revealed that there is a systematic high bias in the pre-1970 model that increases with time when applied to the post-1970 data. This bias is due to the influence of factors other than rainfall on the dependent variable. These other factors include water use such as irrigation or municipal use, land use changes, and other anthropogenic factors. Using these LR models, adjustments were estimated by station. At the end of the period of record (March 30, 2011), adjustments were estimated to be 256 cubic feet per second (cfs) on the Santa Fe River at Ft. White and 35 cfs on the Ichetucknee River. The adjustment to the Ichetucknee Highway 27 record was approximately 1.1 cfs per year, and adjustments to Santa Fe at Ft. White and 441 were 6.2 cfs per year, and 3.6 cfs per year, respectively. Flow adjustment percentages are estimated to be between 10% and 19% of the historical flow, and were gradually applied over the post-1970 record.

## Introduction

The development of a comprehensive flow and stage database for the Lower Santa Fe River was an integral component of the modeling efforts in support of MFL development. Flow and stage data were available from several sources, including the USGS daily data, the District, and USGS field measurements. It was desirable to have a single data source for flow and stage to increase efficiency for both modeling efforts and statistical characterization of hydrologic data. In addition to data compilation, characterization of the trends in the data via statistical trend analysis is vital in order to understand the hydrologic trends in the time series. Since precipitation is typically a primary explanatory variable that affects the hydrologic response of a region, characterizing the trends in rainfall in conjunction with the trends in flow allows for the establishment of casual relationships and examination of changes in causal relationships. Of all the explanatory variables that contribute to streamflow, climatic variables such as rainfall offer the most complete historical data set. Historical anthropogenic changes are difficult to obtain with sufficient resolution to examine the impacts of a single variable on streamflow over time. If the relationship between rainfall and flow can be established during a time period with minimal anthropogenic effects, then this relationship can be projected into later time period. The difference between the projected flow and the measured flow is a measurement of the influences due to non-rainfall factors, such as anthropogenic factors. This methodology was employed for gauging stations in the Lower Santa Fe River basin in order to derive estimates of adjusted flow.

## Project Location

Stage and flow data for twelve (12) surface water stations and groundwater levels for one well were collected. As shown in Figure 1, the sites are located in the Suwannee River Water Management District, throughout Suwannee, Lafayette, Columbia, Gilchrist, Union, Alachua, and Bradford counties. The Luraville, Branford and Bell USGS surface water stations are located on the Suwannee River; Dampier's Landing near Hildreth and Highway 27 near Hildreth USGS stations are located on the Ichetucknee River; and the remaining six USGS surface water stations are located on the Santa Fe River. Table 1 shows only the USGS surface water stations. The Santa Fe River at O'Leno State Park by footbridge gauge is monitored by the Suwannee River Water Management District (the District) and it is not shown in Table 1. This surface water station data was also added to the hydrologic database. The single well utilized for this analysis, well -41705001, is maintained by the Florida Department of Transportation (FDOT) and is located in Lake City.

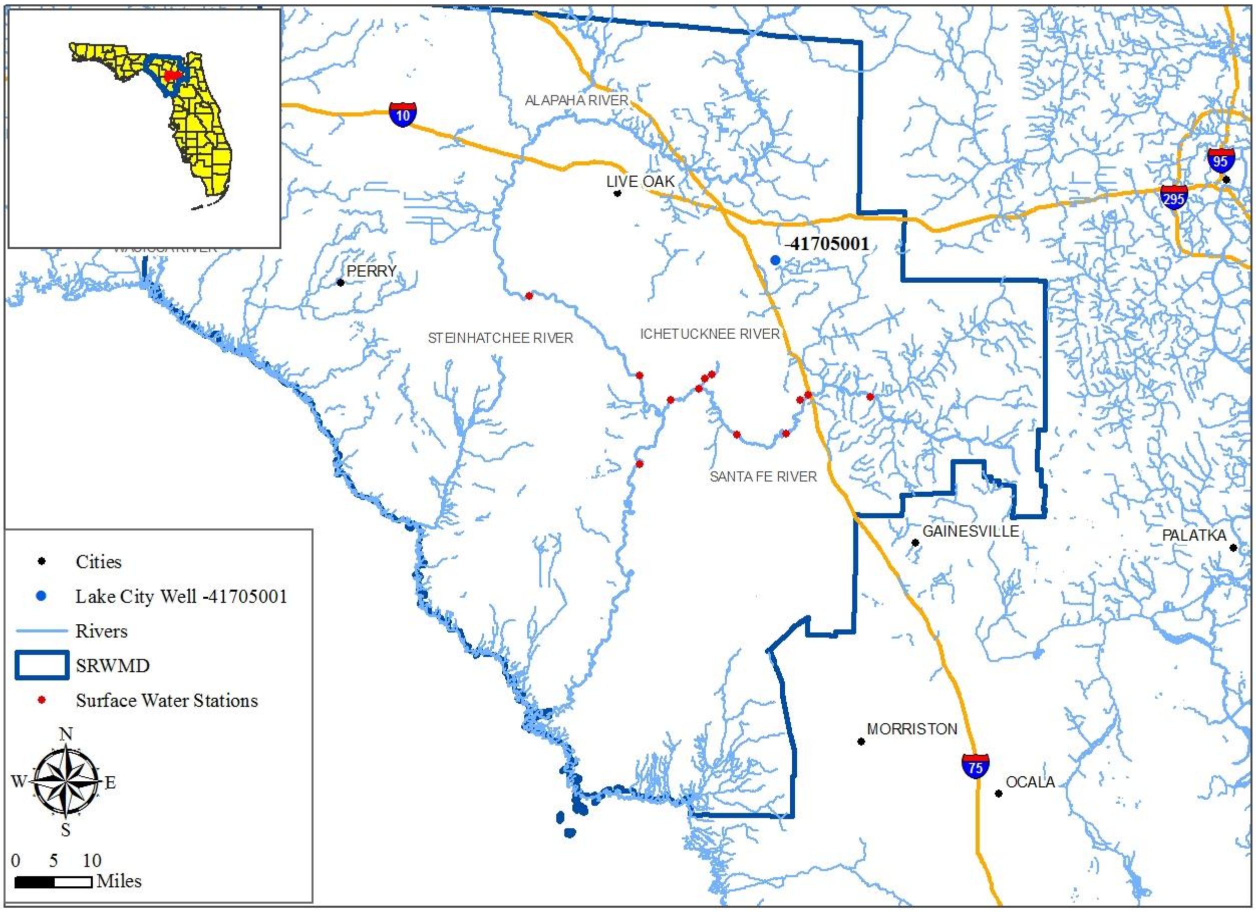



**Figure 1. Location Map: Surface Water Stations and Groundwater Well Site**

**Table 1. USGS Surface Water Stations**

| Station ID | USGS Name | River |
|---|---|---|
| 2320000 | SUWANNEE RIVER AT LURAVILLE, FLA. | SUWANNEE |
| 2320500 | SUWANNEE RIVER AT BRANFORD, FLA. | SUWANNEE |
| 2321500 | SANTA FE RIVER AT WORTHINGTON SPRINGS, FLA. | SANTA FE |
| 2321898 | SANTA FE RIVER AT O'LENO STATE PARK FLA | SANTA FE |
| 2321975 | SANTA FE RIVER AT US HWY 441 NEAR HIGH SPRINGS,FL. | SANTA FE |
| 2322500 | SANTA FE RIVER NEAR FORT WHITE, FLA. | SANTA FE |
| 2322698 | ICHETUCKNEE RIVER AT DAMPIER'S LANDING NR HILDRETH | ICHETUCKNEE |
| 2322700 | ICHETUCKNEE RIVER AT HWY27 NR HILDRETH, FL | ICHETUCKNEE |
| 2322703 | SANTA FE RIVER AB ICHETUCKNEE RIVER NR HILDRETH,FL | SANTA FE |
| 2322800 | SANTA FE RIVER NR HILDRETH FLA | SANTA FE |
| 2323000 | SUWANNEE RIVER NEAR BELL, FLORIDA | SUWANNEE |

## Data Collection and Gap Filling

The majority of flow and stage data of the surface water stations of interest was imported from the USGS National Water Information System (NWIS) database. Table 2 and Table 3 summarize the USGS surface water daily time series and percent of available data when compared to a complete daily record from 10/1/1927 until 3/30/2011. This period of record was originally chosen as an ultimate period of record (to fill) since several surface water stations had a period of record that went back to 10/1/1927. However, the starting date of the period of record desired for gap filling was revised as it was learned that the Ichetucknee at Highway 27 station could only be filled from 6/4/1948. The final statistical model for flow for this station utilized data from a local well as an explanatory variable. The majority of flow and stage data was retrieved from the USGS on 3/31/2011. The USGS stage data was retrieved again on 11/30/2011 to extend the stage records until 9/30/2011. Some of the flow and stage data was provisional at the time of retrieval from the USGS database. USGS water-data reports were consulted to account for datum shifts and to make adjustments in the time series.

**Table 2. USGS Discharge Daily Time Series and Percent Availability**

| Station ID | USGS NAME | Q Min Date | Q Max Date | % Q Data Available |
|---|---|---|---|---|
| 2320000 | SUWANNEE RIVER AT LURAVILLE, FLA. | 10/1/1927 | 4/4/2011 | 29.67% |
| 2320500 | SUWANNEE RIVER AT BRANFORD, FLA. | 7/1/1931 | 3/29/2011 | 95.50% |
| 2321500 | SANTA FE RIVER AT WORTHINGTON SPRINGS, FLA. | 10/1/1931 | 3/29/2011 | 95.20% |
| 2321975 | SANTA FE RIVER AT US HWY 441 NEAR HIGH SPRINGS,FL. | 10/1/1992 | 9/30/2002 | 10.78% |
| 2322500 | SANTA FE RIVER NEAR FORT WHITE, FLA. | 10/1/1927 | 3/30/2011 | 97.16% |
| 2322698 | ICHETUCKNEE RIVER AT DAMPIER'S LANDING NR HILDRETH, FLA | 2/15/2002 | 4/3/2011 | 10.51% |
| 2322700 | ICHETUCKNEE RIVER AT HWY27 NR HILDRETH, FL | 2/5/2002 | 4/4/2011 | 10.97% |
| 2322703 | SANTA FE RIVER AB ICHETUCKNEE RIVER NR HILDRETH,FL | | | |
| 2322800 | SANTA FE RIVER NR HILDRETH FLA | 11/1/2000 | 3/30/2011 | 10.03% |
| 2323000 | SUWANNEE RIVER NEAR BELL, FLORIDA | 6/1/1932 | 3/29/2011 | 42.20% |

**Table 3 . USGS Gauge Height Daily Time Series and Percent Availability**

| Station ID | USGS NAME | Gauge Min Date | Gauge Max Date | % Gauge Height Data Available |
|---|---|---|---|---|
| 2320000 | SUWANNEE RIVER AT LURAVILLE, FLA. | 10/1/1927 | 4/4/2011 | 29.00% |
| 2320500 | SUWANNEE RIVER AT BRANFORD, FLA. | 7/9/1931 | 3/29/2011 | 95.00% |
| 2321500 | SANTA FE RIVER AT WORTHINGTON SPRINGS, FLA. | 11/18/1931 | 3/29/2011 | 89.00% |
| 2321898 | SANTA FE RIVER AT O'LENO STATE PARK FLA | 6/8/2010 | 4/5/2011 | 1.00% |
| 2322500 | SANTA FE RIVER NEAR FORT WHITE, FLA. | 1/21/1994 | 3/30/2011 | 14.00% |
| 2322698 | ICHETUCKNEE RIVER AT DAMPIER'S LANDING NR HILDRETH, FLA | 2/15/2002 | 4/3/2011 | 10.00% |
| 2322700 | ICHETUCKNEE RIVER AT HWY27 NR HILDRETH, FL | 2/6/2002 | 4/4/2011 | 11.00% |
| 2322703 | SANTA FE RIVER AB ICHETUCKNEE RIVER NR HILDRETH,FL | 12/23/2009 | 3/29/2011 | 1.00% |
| 2322800 | SANTA FE RIVER NR HILDRETH FLA | 4/28/1947 | 3/30/2011 | 58.00% |
| 2323000 | SUWANNEE RIVER NEAR BELL, FLORIDA | 6/1/1932 | 3/29/2011 | 43.00% |

In addition, USGS field measurements were imported for the Ichetucknee River at Highway 27 near Hildreth surface water station (2322700). USGS field measurements consist of manual readings of streamflow and gauge height and are generally used to supplement the USGS daily time series (http://waterdata.usgs.gov/usa/nwis/sw). Rainfall and well data as well as the additional surface water flow and stage data was obtained from the District. The District flow and stage data was provided to INTERA as Microsoft Office Excel worksheets (SW_DVQ.xlsx, SW_POINTQ.xlsx, SW_DVSTAGE.xlsx, and SW_POINTSTAGE.xlsx).

Gauge data gaps were filled using various types of statistical methods. A summary of all gauges and the models developed for filling is shown in Table 4. The models used for filling include simple linear regressions (SLRs), rating curves, multiple linear regressions (MLRs), and artificial neural networks (ANNs). For all cases, a general hierarchy was followed for statistical model development. First, a stage-discharge relationship was developed for the dependent variable of interest, if possible. If there was adequate data to develop this relationship and the fit of the relationship was strong, the stage-discharge relationship was utilized for filling. If it was not possible to develop a stage-discharge relationship, a nearby gauge was examined in order to develop a simple linear regression for flow using nearby flow or stage using stage. If a simple linear regression did not adequately predict the response variable, a multiple linear regression or artificial neural network with multiple inputs was utilized to predict the response variable. For small gaps (ideally several days or less) linear interpolation was also utilized. Because the other Suwannee River stations did not have records extending as far back as Luraville, Luraville was utilized to fill stage during the early record for both the Branford and Bell stations. For both Branford and Bell, stage-discharge relationships were developed to fill flow records after filling the stage record. The Santa Fe River stations were filled using various methods, including stage-discharge relationships, simple linear regression, and multiple linear regression. The Ichetucknee stations were filled using ANNs with -41705001 well data and the Santa Fe River near Hildreth (2322800) data as inputs. All data gaps were filled from 1948 until 2011. For the Ichetucknee stations, the limiting factor for filling was the available -41705001 well data.

**Table 4. Statistical Model Summary**

| Station Number | Station Name | Data Type | Model Type | Explanatory Variable(s) |
|---|---|---|---|---|
| 2320500 | Suwannee River at Branford | Gauge height | SLR | Luraville gauge height |
| | | Flow | Rating curve | Branford gauge height |
| 2323000 | Suwannee River at Bell | Gauge height | SLR | Branford gauge height or Luraville gauge height |
| | | Flow | Rating curve | Bell gauge height |
| 2321500 | Santa Fe River at Worthington | Gauge height | Rating curve | Worthington flow(4 flow regimes) |
| 2321975 | Santa Fe River at 441 | Gauge height | MLR | Ft. White gauge height, flow, Worthington gauge height, flow |
| | | Flow | MLR | Ft. White flow, Worthington flow |
| 2322500 | Santa Fe River near Ft. White | Gauge height | MLR | Bell gauge height, Worthington flow |
| 2322703 | Santa Fe River above Ichetucknee near Hildreth (3 Rivers) | Gauge height | SLR | Santa Fe Hildreth gauge height |
| 2322800 | Santa Fe River near Hildreth | Gauge height | MLR | Ft. White gauge height, Branford gauge height |
| | | Flow | MLR | Hildreth gauge height Ft. White flow |
| 2322700 | Ichetucknee River Highway 27 near Hildreth | Gauge height | ANN | FDOT well -41705001, Santa Fe Hildreth gauge height |
| | | Flow | MLR | Santa Fe Hildreth flow, Santa Fe Hildreth 7-day lagged flow, FDOT well -41705001, |

For all models, the root mean squared error (RMSE), average residual, and coefficient of determination ($R^2$) are shown in conjunction with the graphical model fit to assess the goodness-of-fit of the statistical model. These regression diagnostics are defined in Table 5.

**Table 5. Regression Diagnostics**

| Regression Diagnostic | Indicator |
|---|---|
| RMSE | Root mean squared error: measures the 'typical' error in the model (without regard to sign) |
| Average Residual | The general trend of the model to over-estimate or underestimate output |
| $R^2$ | Coefficient of Determination: the proportion of variability in the data set that the statistical model accounts for |

## Suwannee River Surface Water Stations

There are three USGS surface water stations located on the Suwannee River: Luraville, Branford, and Bell. Figure 2 shows the surface water stations located on the Suwannee River. The stage and flow data of the Suwannee River stations, Branford (2320500) and Bell (2323000) are of primary interest on the Suwannee as they bracket the Suwannee-Santa Fe confluence, and provide useful information on tailwater conditions in the Lower Santa Fe River. The Luraville (2320000) station was successfully used to fill the Branford and Bell sites as needed. Linear regressions generally produced $R^2$ values of 0.90 or greater; hence, it was appropriate to use linear regressions to fill data gaps. Small intermittent gaps were filled using linear interpolation. Since the Luraville data is not utilized in modeling efforts and was used solely to fill gaps at other Suwannee River surface water stations, the Luraville data gaps were not filled.

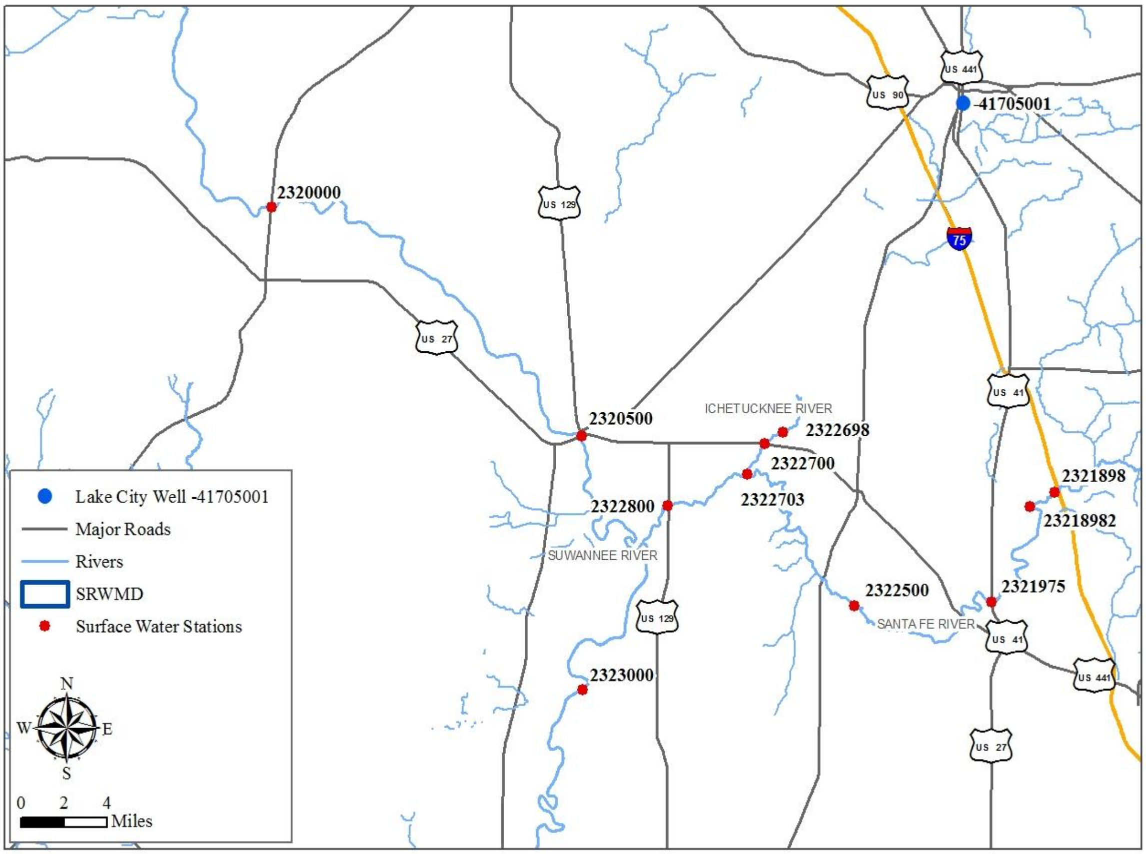



**Figure 2. Surface Water Stations on the Suwannee River**

### Suwannee River at Luraville (2320000)

The Suwannee River at Luraville stage and flow data were obtained from the USGS. The Luraville stage and flow data gaps were not filled. The Luraville stage and flow data are shown in Figures 3 and 4. As shown, data at this station begins in 1927, while data collection at Bell and Branford begin in 1932 and 1931, respectively. For this reason, the Luraville station represents the best available data to utilize for gap filling the early period of record.

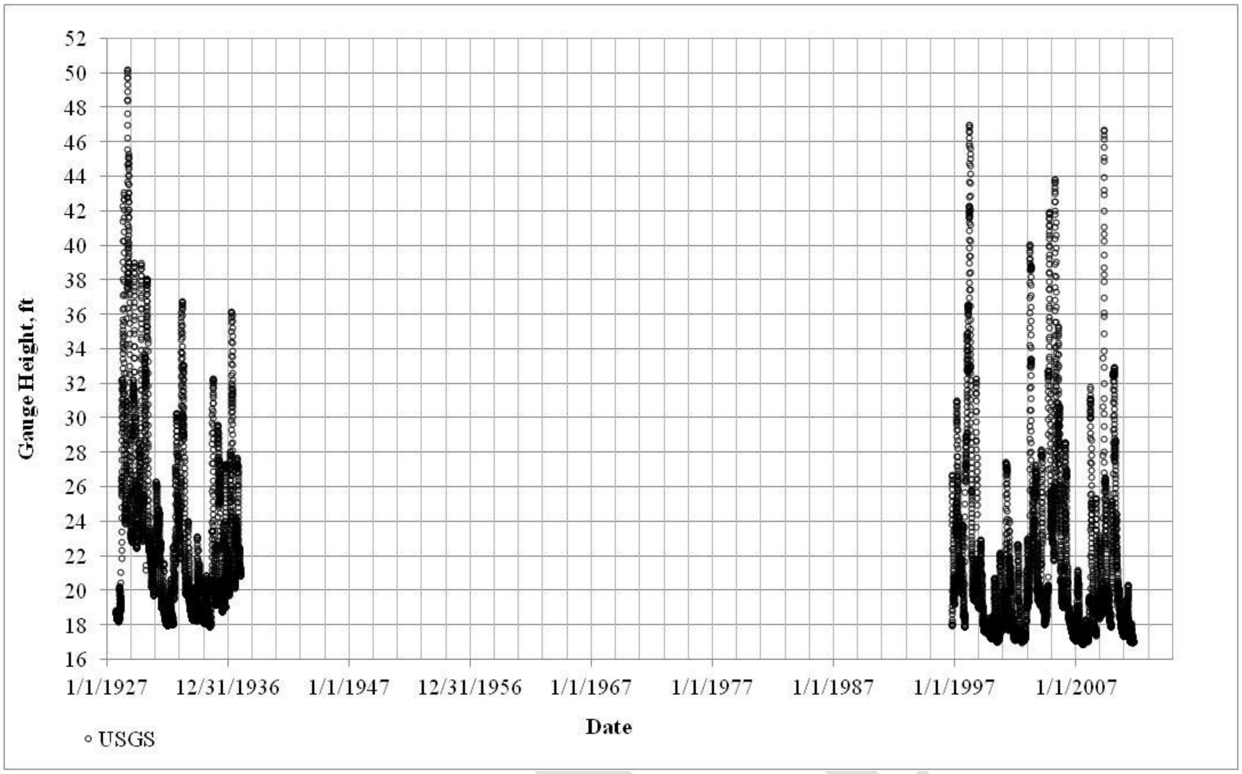



**Figure 3. Luraville Gauge Height (ft, NGVD29)**

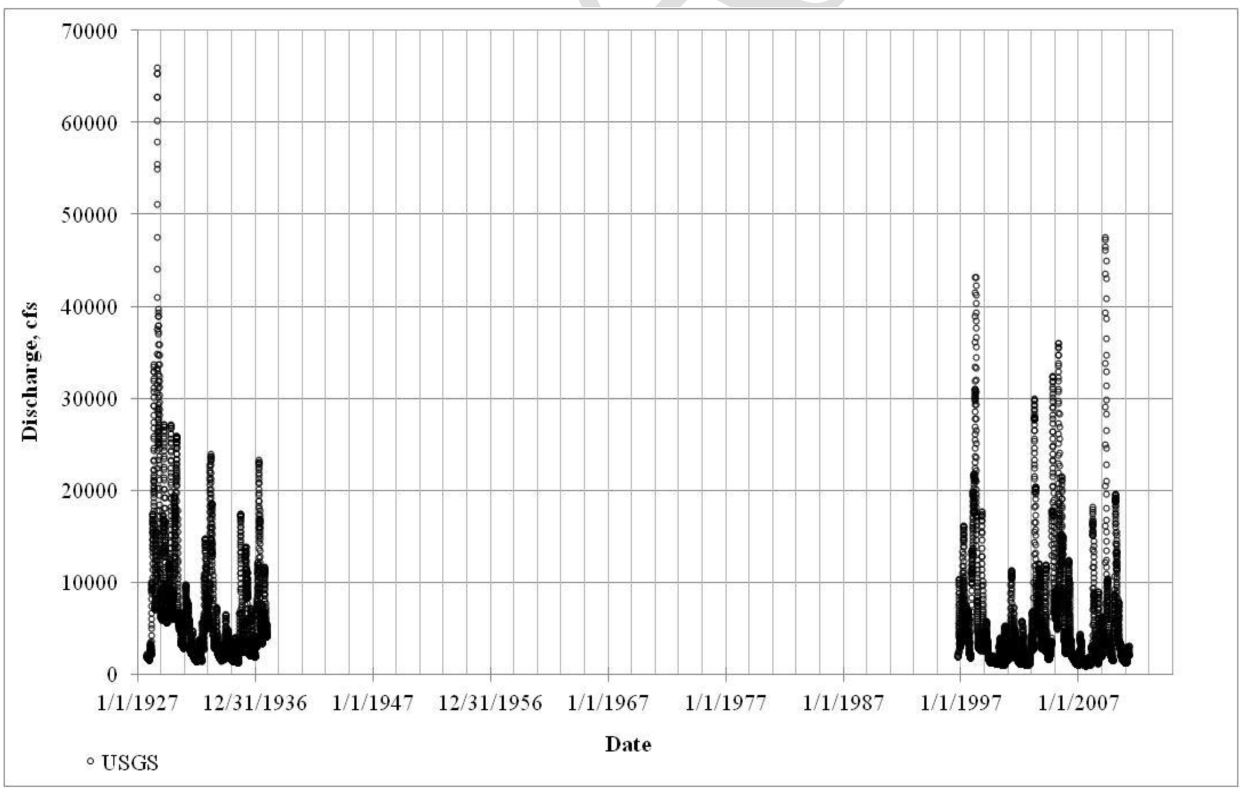



**Figure 4. Luraville Discharge (cubic feet per second, cfs)**

### Suwannee River at Branford (2320500)

The Branford gauge height record is one of the longest available USGS records among the surface water stations of interest. The stage data gaps were filled using the USGS Luraville gauge height record which was adjusted for a datum shift. Two SLR models were developed for high and low stages, as shown in Figures 5 and 6. A summary of the performance of the SLRs is shown in Table 6. The high and low stage models were divided based on the natural inflection point of the Luraville stage data and the Branford stage data, at a Luraville stage of 30 ft. The complete (filled) Branford gauge height record is shown in Figure 7.

**Table 6. Branford Gauge SLR Performance Summary**

| Independent Variable | RMSE | Average Residual | $R^2$ |
|---|---|---|---|
| Luraville < 30 ft | 0.89 | -0.00072 | 0.94 |
| Luraville ≥ 30 ft | 0.82 | 0.00064 | 0.93 |

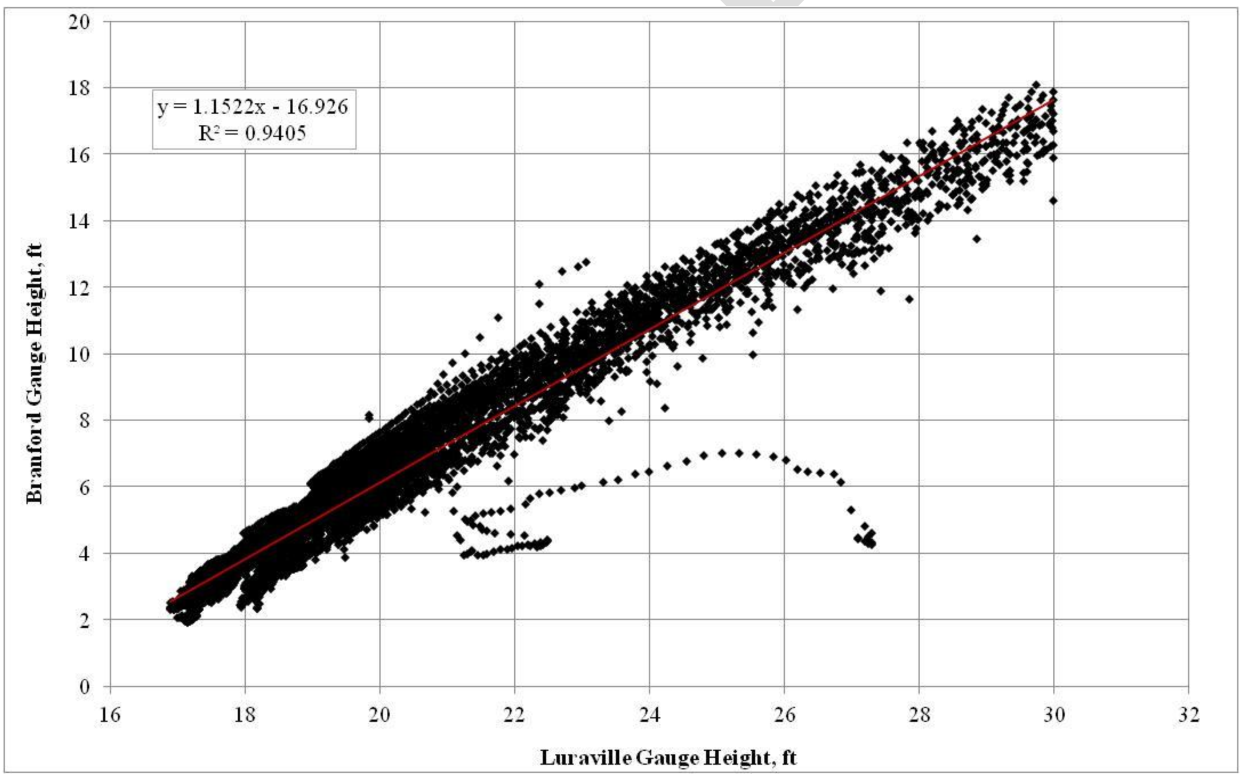



**Figure 5. SLR: Branford Gauge Height vs. Luraville Gauge Height (for Luraville Gauge less than 30 ft)**

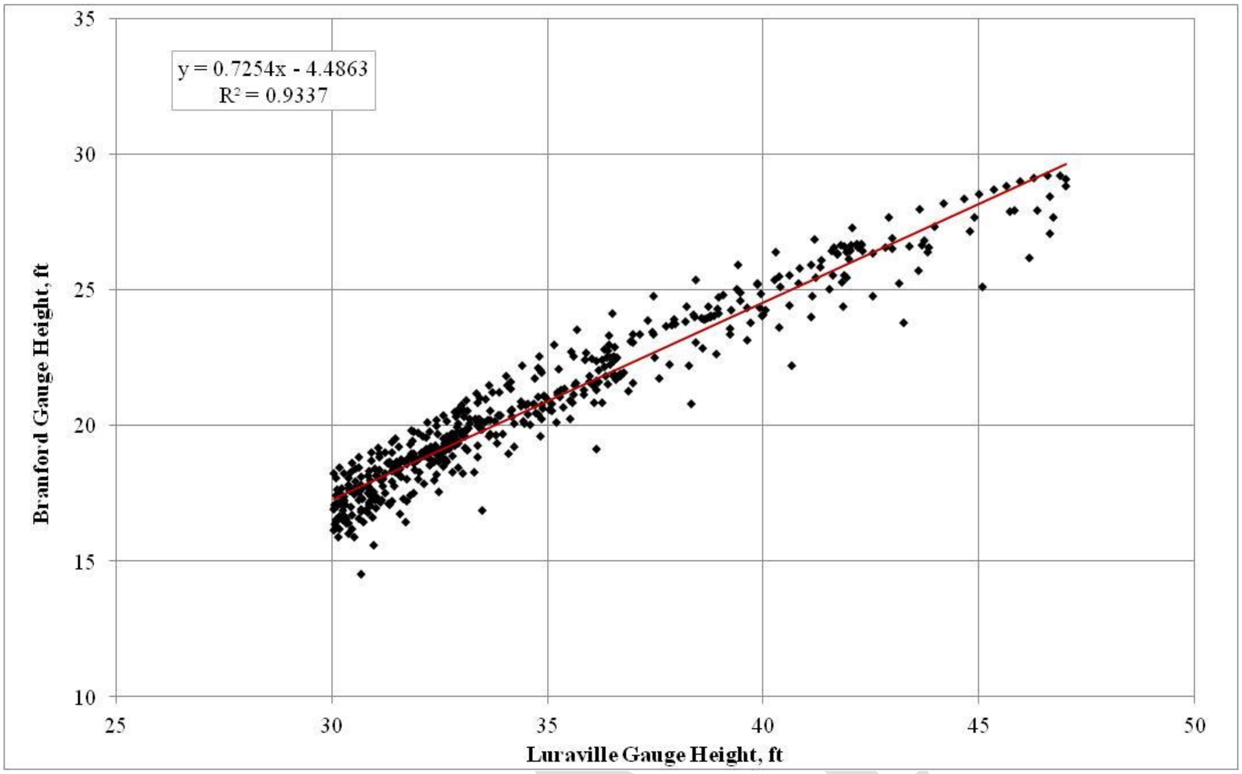



**Figure 6. SLR: Branford Gauge Height vs. Luraville Gauge Height (for Luraville Gauge greater than 30 ft)**

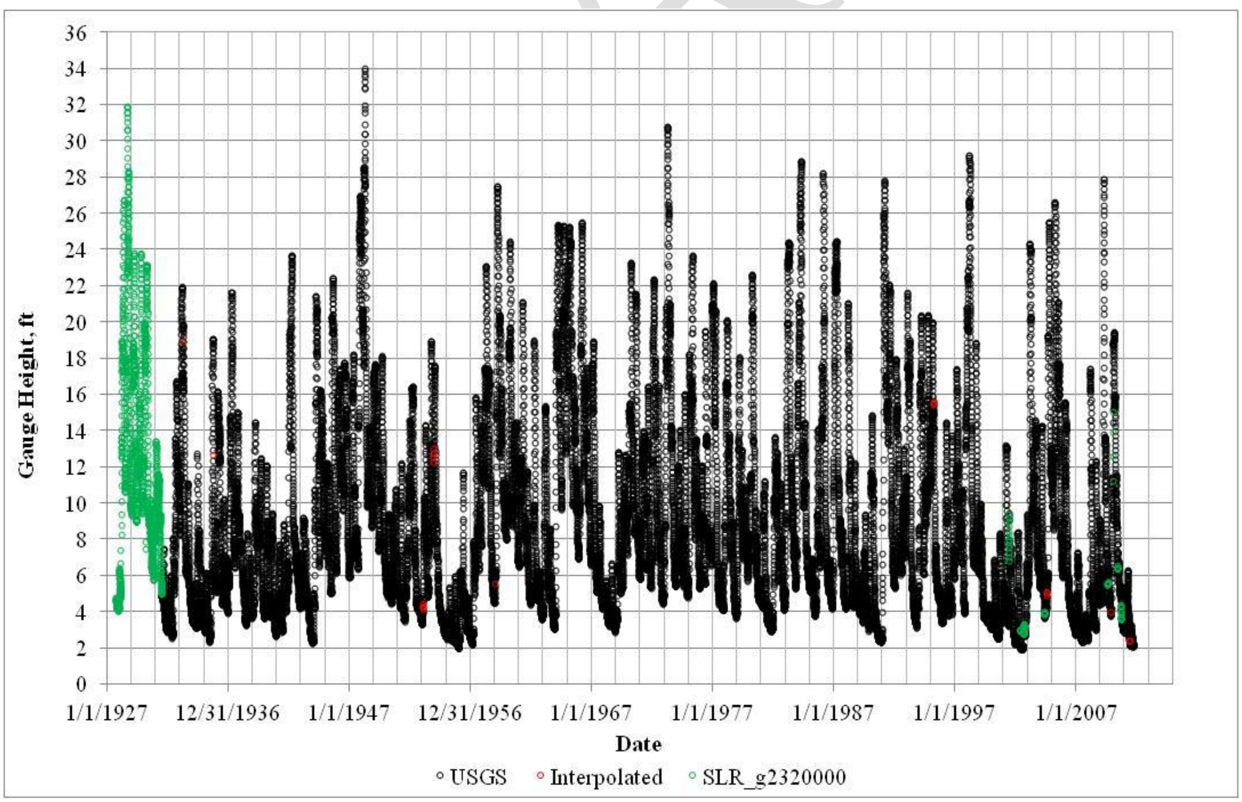



**Figure 7. Branford Gauge Height (4.81 ft above NGVD29)**

The gaps in the Branford discharge record were filled using stage-discharge relationship (Figure 8) at the Branford station. The filled discharge record is shown in Figure 9. Model statistics for the stage discharge relationship are shown in Table 7.

**Table 7.Branford Discharge Regression Performance Summary**

| Regression Performance Summary | |
|---|---|
| RMSE | 526.702 |
| Average Residual | -18.642 |
| $R^2$ | 0.9923 |

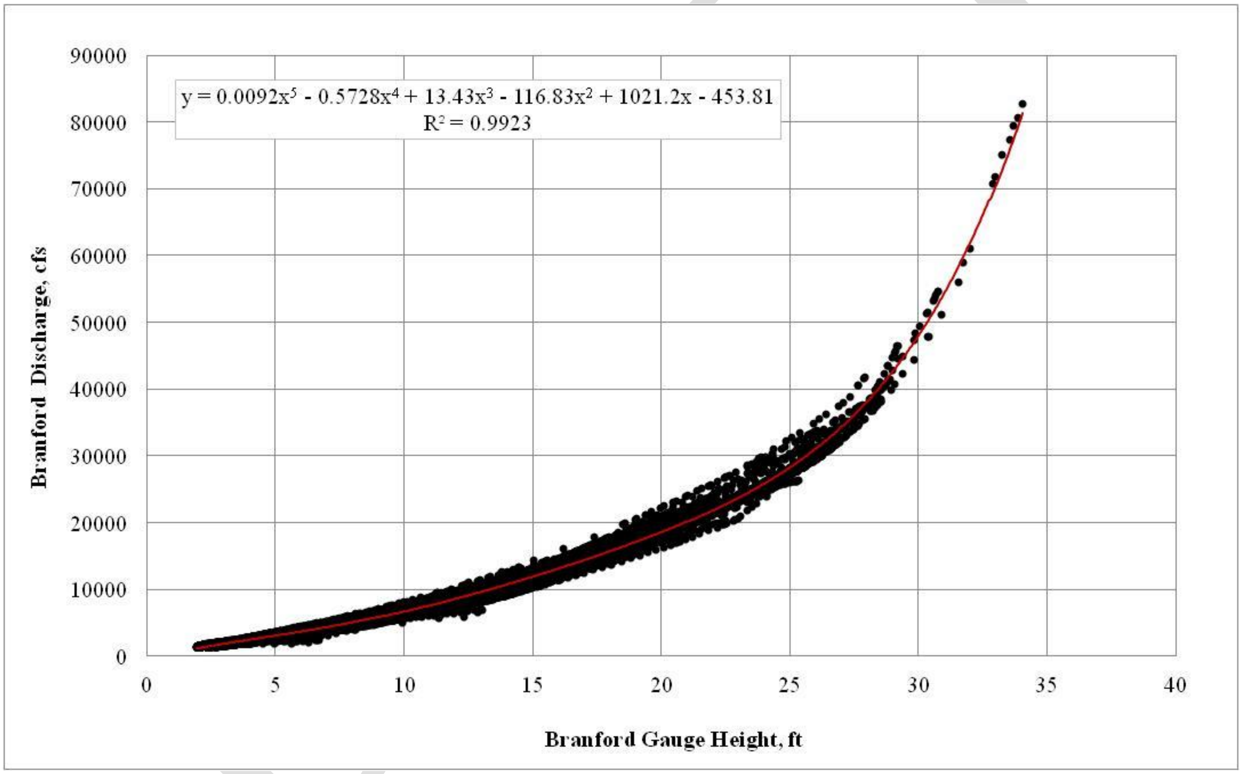



**Figure 8. Polynomial Fit: Branford Discharge vs. Branford Gauge Height**

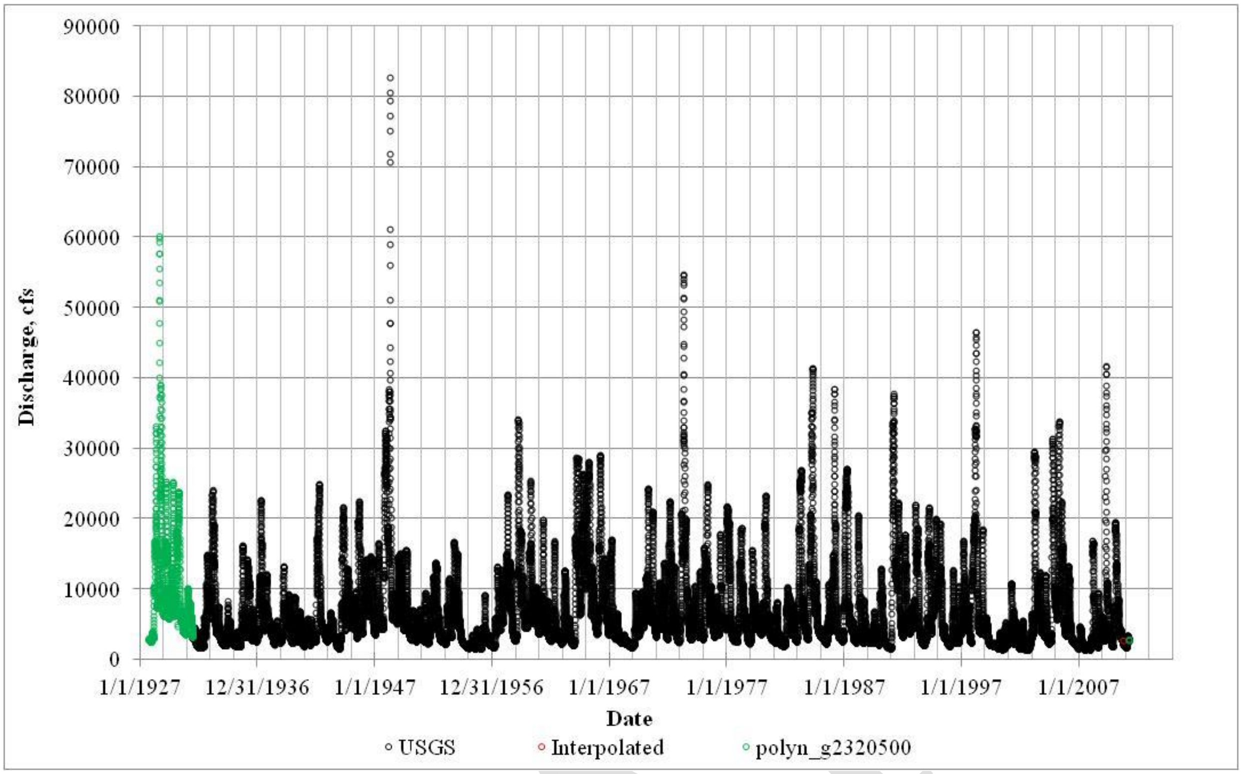



**Figure 9. Branford Discharge (cfs)**

## Suwannee River near Bell (2323000)

The gaps in the Bell gauge height record were filled using the Luraville and Branford gauge height records. A SLR model was developed for Bell and Branford stage, as shown in Figure 10. Two SLR models were developed for Bell and Luraville stage, shown in Figures 11 and 12. The Luraville models were divided between low and high stages, based on the natural inflection point of the data (Luraville stage of 30 ft). When gap filling, the Branford/Bell SLR was utilized when data was available at Branford. When data was not available at Branford, the Luraville SLRs were used for gap filling. Short intermittent gaps were filled using linear interpolation. The resulting gauge height record is shown in Figure 13. SLR model statistics are shown in Table 8.

**Table 8. Bell Gauge SLR Performance Summary**

| Independent Variable | RMSE | Average Residual | $R^2$ |
|---|---|---|---|
| Branford Gauge | 0.6743 | -5.62e-5 | 0.9833 |
| Luraville Gauge < 30 ft | 0.9088 | -7.5e-4 | 0.9101 |
| Luraville Gauge ≥ 30 ft | 1.196 | -1.75e-2 | 0.7940 |

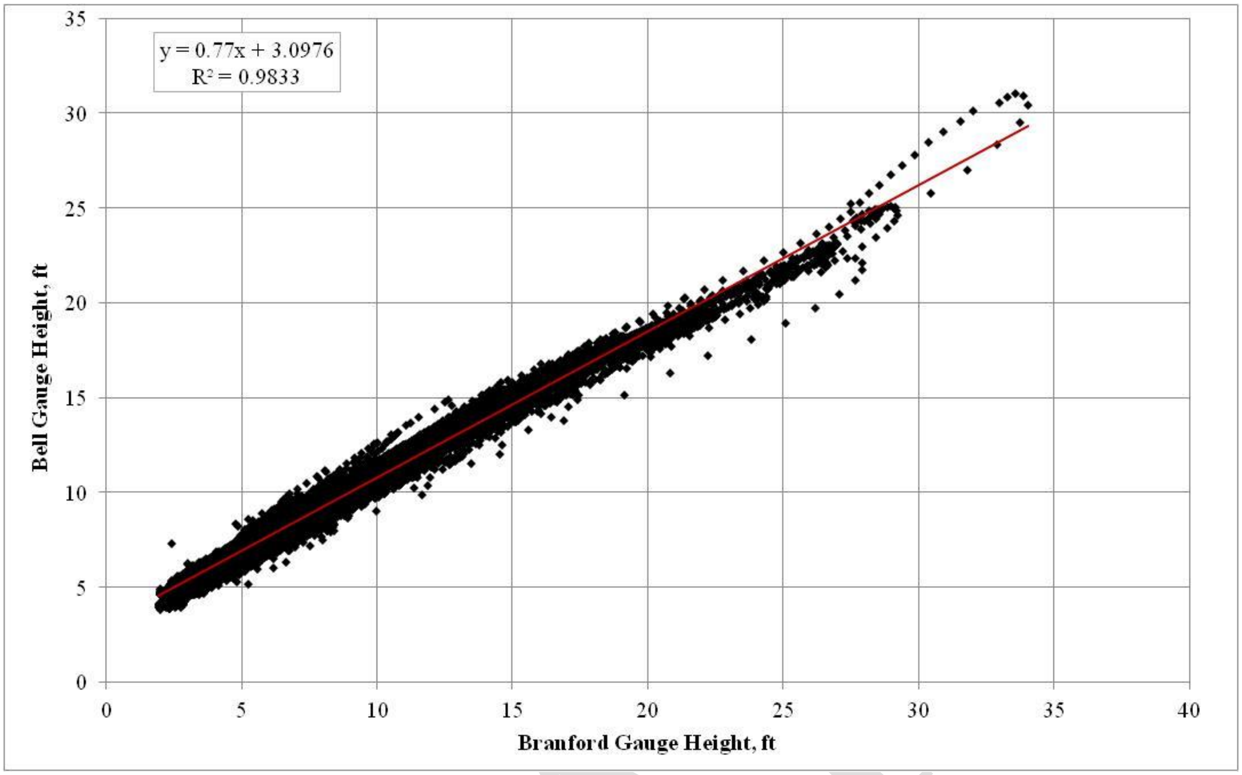



**Figure 10. SLR: Bell Gauge Height vs. Branford Gauge Height**

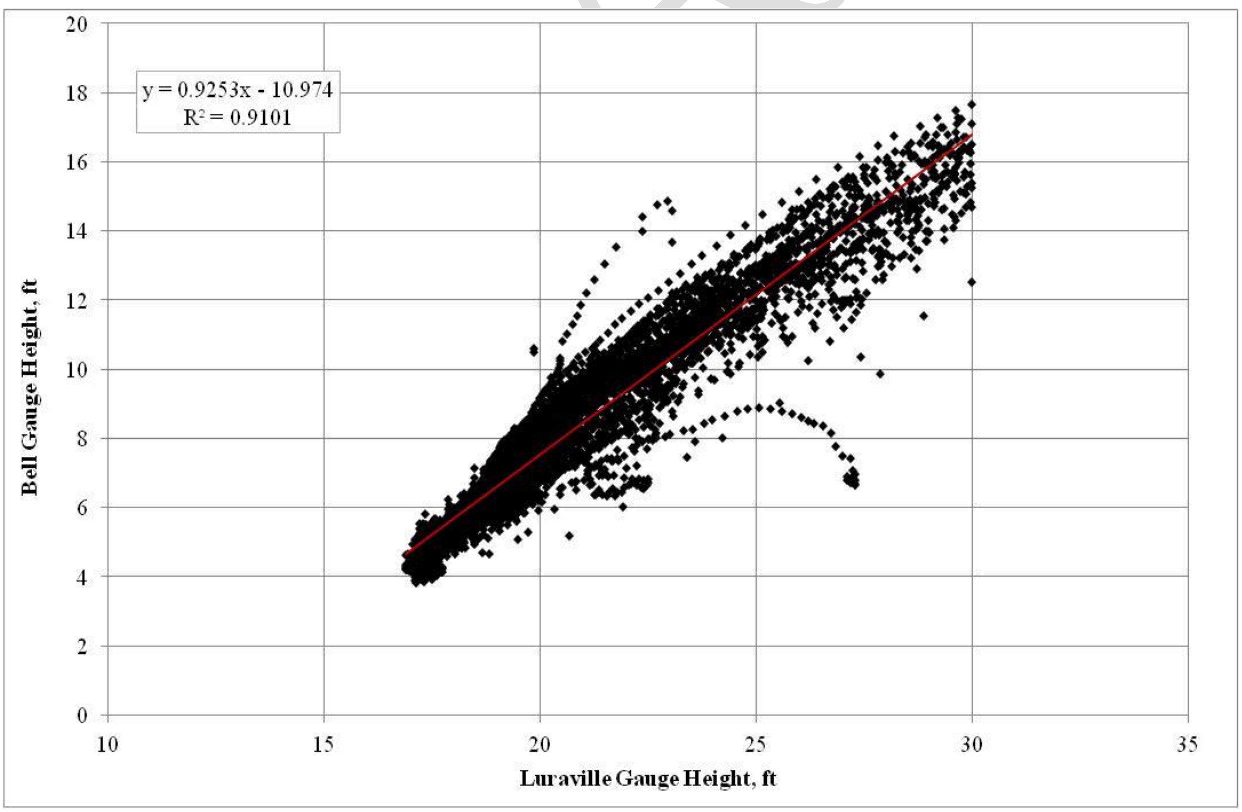



**Figure 11. SLR: Bell Gauge Height vs. Luraville Gauge Height (for Luraville Gauge Height less than 30 ft)**

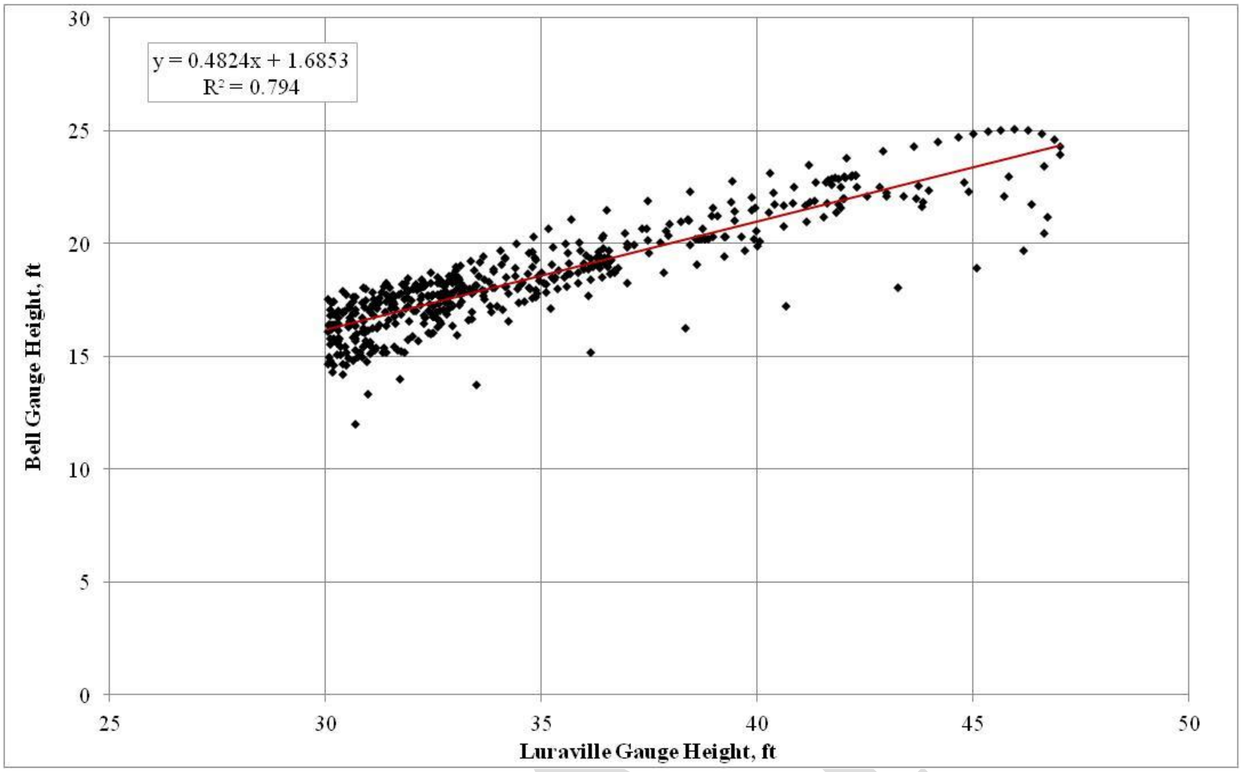



**Figure 12. SLR: Bell Gauge Height vs. Luraville Gauge Height (for Luraville Gauge Height greater than 30 ft)**

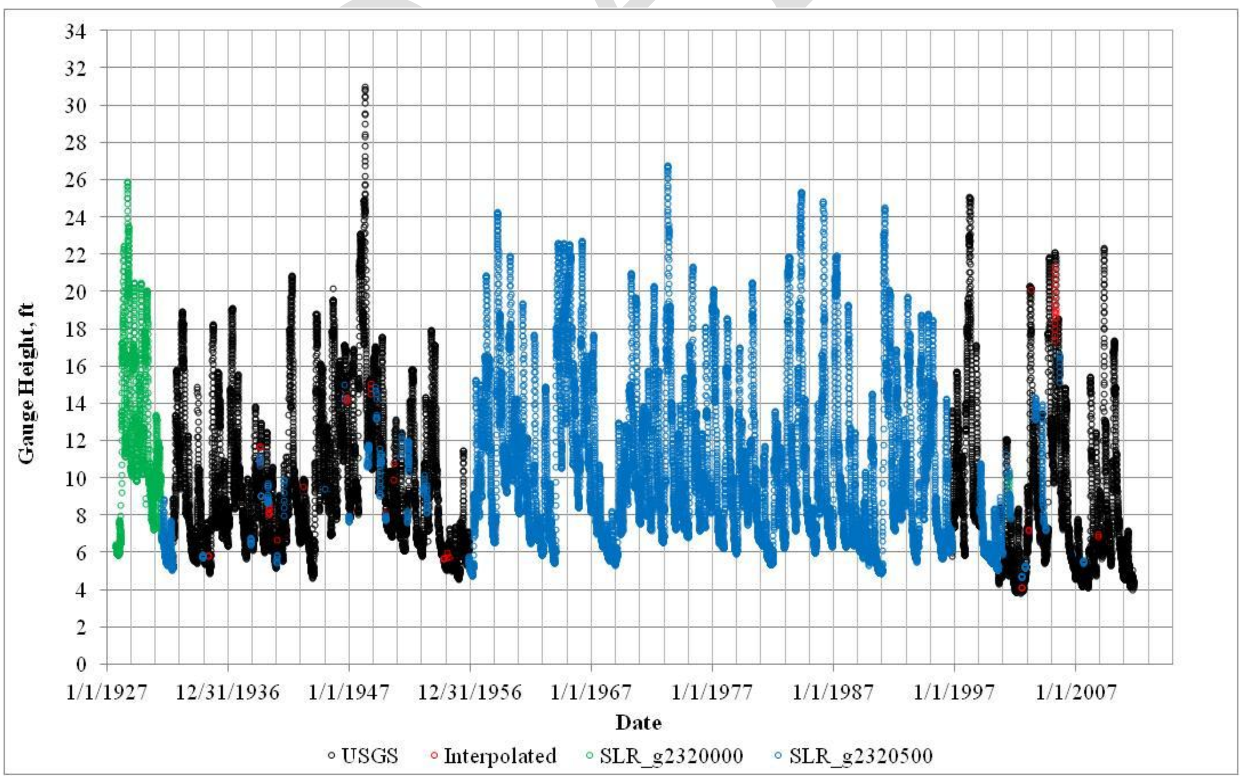



**Figure 13. Bell Gauge Height (ft, NGVD29)**

Gaps in the Bell discharge time series were filled using the stage-discharge relationship developed for Bell, as shown in Figure 14 and Table 9. The resulting filled time series is shown in Figure 15.

**Table 9.Bell Discharge Regression Performance Summary**

| Regression Performance Summary | |
|---|---|
| RMSE | 399.0 |
| Average Residual | 1.8910 |
| $R^2$ | 0.9957 |

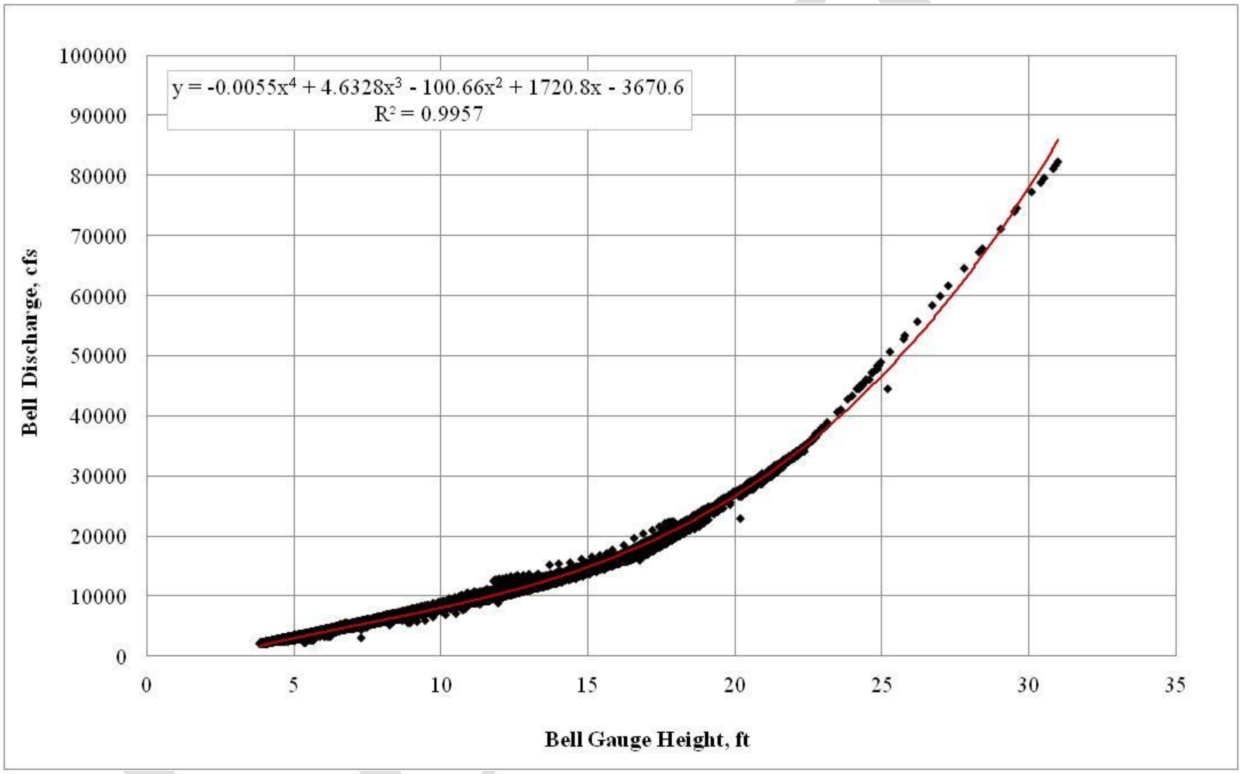



**Figure 14. Polynomial Fit: Bell Discharge vs. Bell Gauge Height**

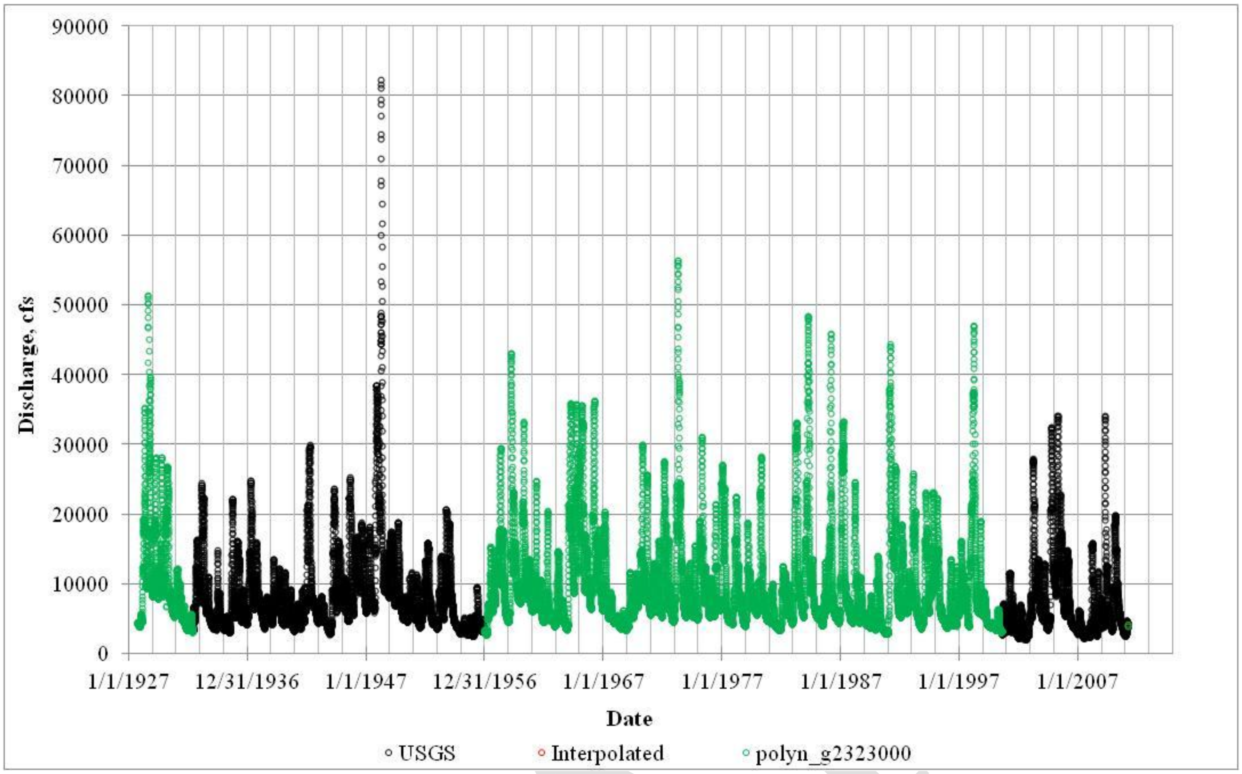



**Figure 15. Bell Discharge (cfs)**

## Santa Fe River Surface Water Stations

There are seven surface water stations on the Santa Fe River (Figure 16). Tables 2 and 3 show that the Santa Fe River near Fort White (2322500) surface water station has the longest discharge record and the Santa Fe River at Worthington Springs (2321500) has the second longest discharge record among the USGS stations on the Santa Fe River. The recently installed Santa Fe River at O'Leno State Park (2321898) surface water station has no available discharge measurements and a short gauge height record, from 6/8/2010 until 9/30/2011. The gaps in the USGS O'Leno State Park gauge height record were not filled.

The Santa Fe River at O'Leno State Park by footbridge is the gauge monitored by the District. The gauge was assigned a station number (Station ID) of 23218982 in the database. The Santa Fe at O'Leno State Park by footbridge (23218982) stage and flow data gaps were not filled.

For many of the Santa Fe River flow stations, additional data was available from the District.  There was no available USGS gauge height data at the Santa Fe River at Highway 441 near High Springs (2321975). Additional gauge height values were added from the District's database of stage measurements (SW_DVSTAGE.xls). Additional discharge values were added from the District's database of flow measurements (SW_DVQ.xls) to the Santa Fe River at Highway 441 near High Springs (2321975) record. Additional gauge height values were added from the District database (SW_DVSTAGE.xls) to the available USGS stage record of the Santa Fe River near Hildreth (2322800) surface water station. In addition, gauge height values were added from the District's stage databases (SW_DVSTAGE.xls and SW_POINTSTAGE.xls) to the available USGS stage record of the Santa Fe River above Ichetucknee near Hildreth (2322703) gauge, also known as the Three Rivers gauge. Discharge and gauge height records for the surface water stations on the Santa Fe River were generally filled from 6/4/1948. The period of record of the well-41705001 was a limiting factor.

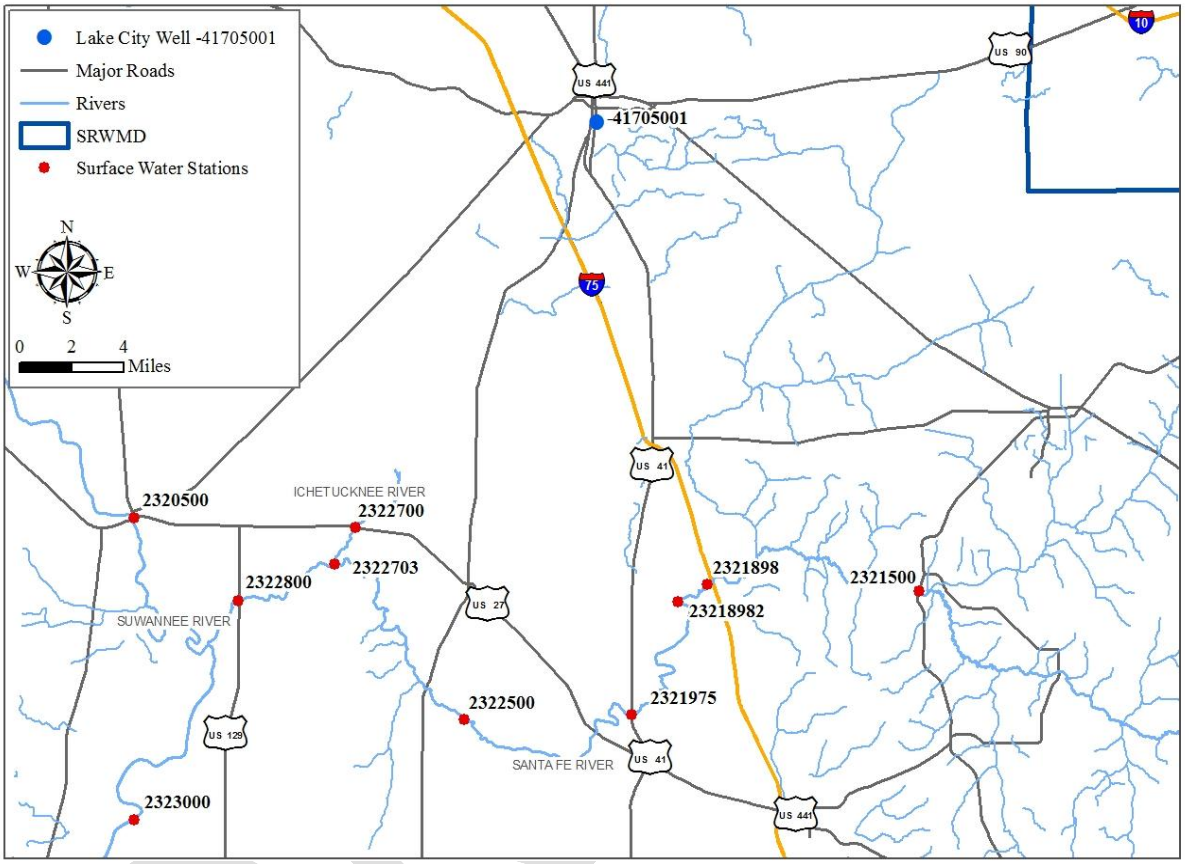



**Figure 16. Surface Water Stations on the Santa Fe River**

## Santa Fe River at Worthington Springs (2321500)

The Santa Fe River at Worthington Springs (2321500) stage data was gap filled from 6/4/1948. The small intermittent gaps were filled using linear interpolation. Two larger gaps in the Worthington stage record, from 2/13/1964 until 9/30/1964 and from 10/01/1966 until 9/30/1968, were filled using developed relationships between Worthington stage and discharge.

Several relations between stage and discharge at Worthington were fit, including a logarithmic fit and an exponential fit (Figure 17). Using the entire period of flow record and relating it to stage produced inadequate fits. The Worthington stage and flow data were sorted together and grouped by flow. The groups included: (a) flow values less than 400 cfs; (b) flow values greater than or equal to 400 cfs and less than 1000 cfs; (c) flow values greater than or equal to 1000 cfs and less than 6000 cfs; (d) flow values greater than or equal to 6000 cfs. The resultant relationships, after grouping by flow, produced adequate fits and $R^2$ values (Figure 18). Application of the resultant relationships, however, produced a simulated gauge height record with discontinuities between 11 and 12 feet (Figure 19). Additional correction was necessary to minimize the error; hence, shifts were generated to apply to the originally developed regressions (Table 10). The stage-discharge relationships for each of the 4 flow categories are shown in Figure 19. As shown, there are discontinuities present at each of the 3 transition points between flow groups (shown in blue). As shown, all equations, with the exception of the low flow equation, were shifted in order to create a continuous stage-discharge relationship (shown in black in Figure 19). These shifts made the transition from one regression curve to the next more seamless than the original relationships, and eliminated the discontinuities originally present in the filled data (Figure 20). The resultant gauge height record is shown in Figures 21 and 22.

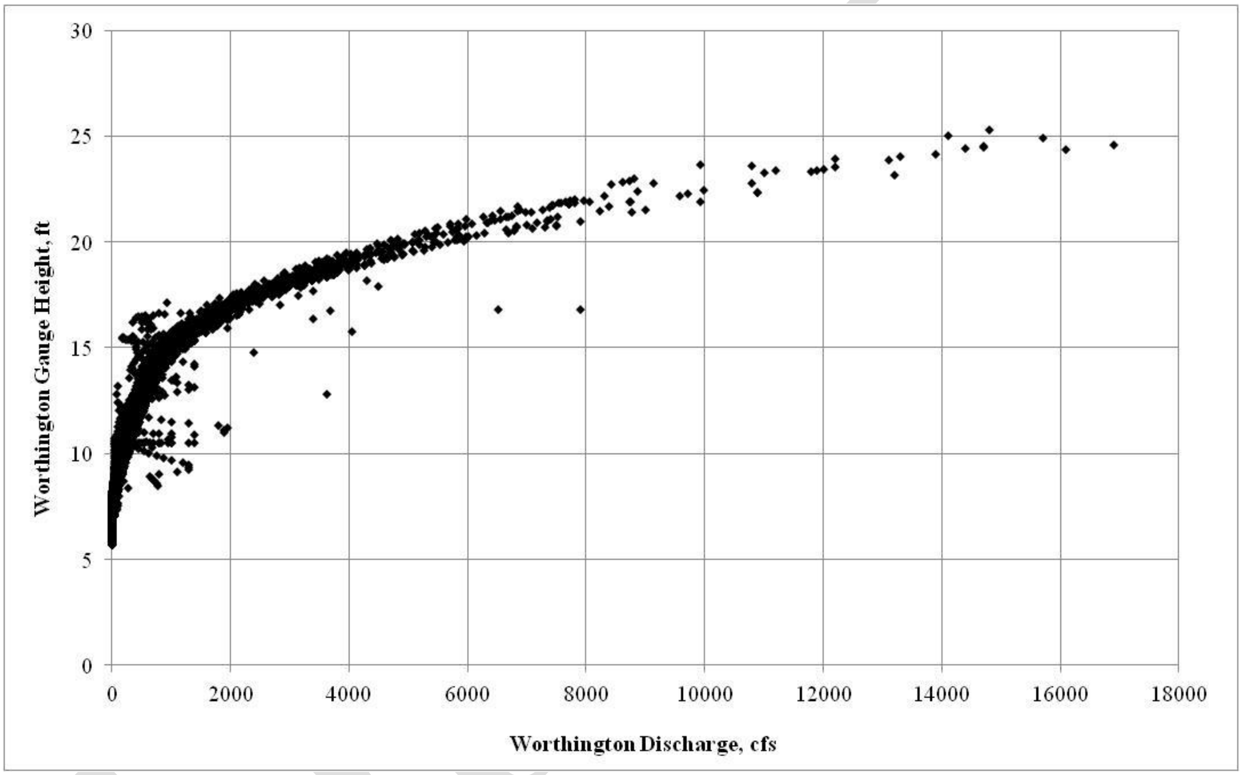



**Figure 17. Worthington Gauge Height vs. Worthington Discharge**

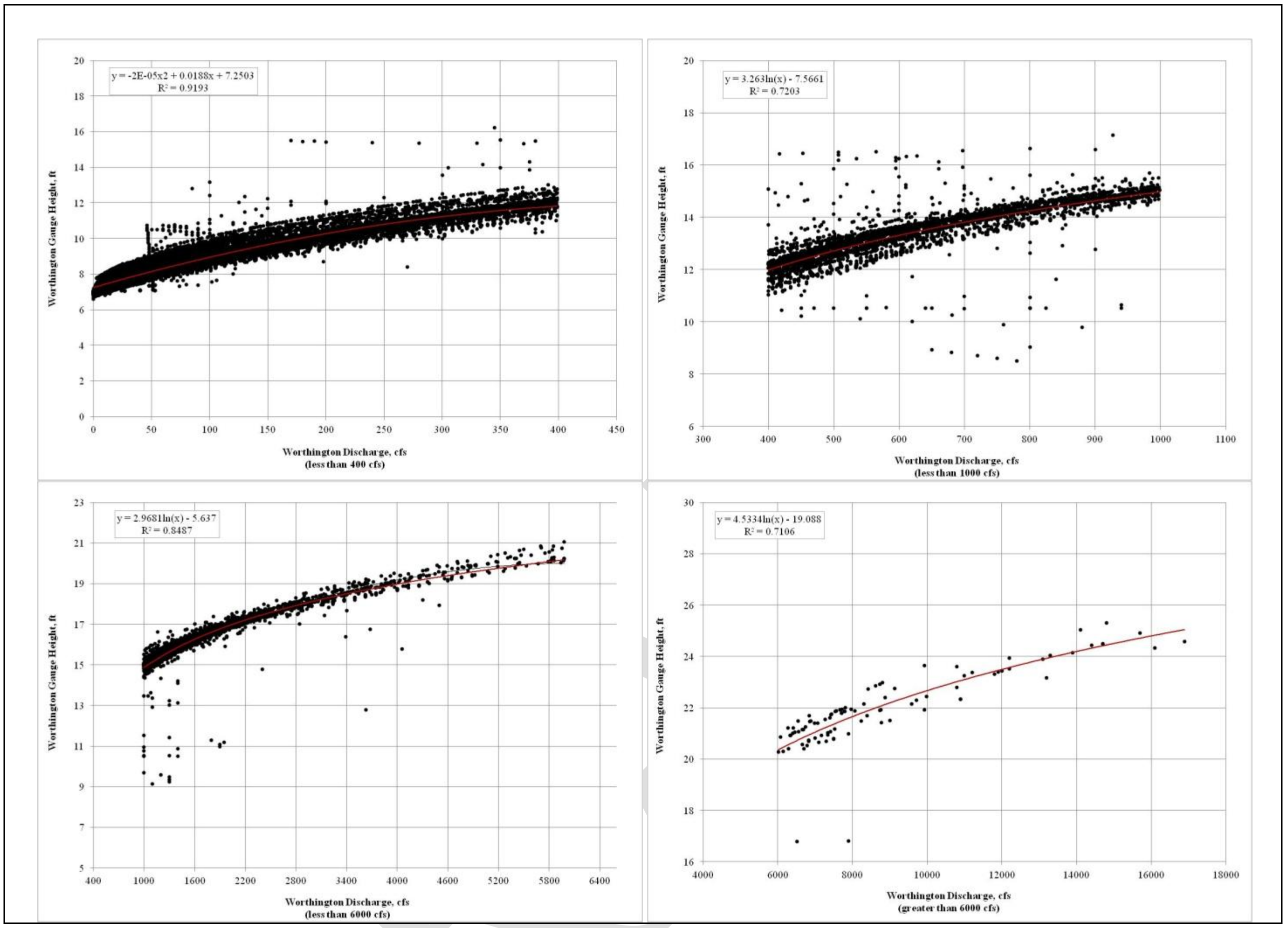



**Figure 18. Worthington Gauge Height vs. Discharge: (a) Q is less than 400 cfs; (b) Q is less than 1000 cfs; (c) Q is less than 6000 cfs; (d) Q is greater than 6000 cfs**

**Table 10. Worthington Stage-Discharge Regression Shifts**

| Initial Regression | Shift | Resultant Regression | Database Code |
|---|---|---|---|
| $y = -2E\text{-}05x^2+0.0188x + 7.2503$ | N/A | $y = -2E\text{-}05x^2+0.0188x + 7.2503$ | polyn_q2321500 |
| y = 3.263ln(x) -7.5661 | -0.44 | y = 3.263ln(x) -7.5661-0.44 | ln_q2321500_Shift |
| y = 2.9681ln(x) -5.637 | -0.375 | y = 2.9681ln(x) -5.637-0.375 | ln_q2321500_Shift |
| y = 4.5334ln(x) - 19.088 | -0.58 | y = 4.5334ln(x) - 19.088-0.58 | ln_q2321500_Shift |

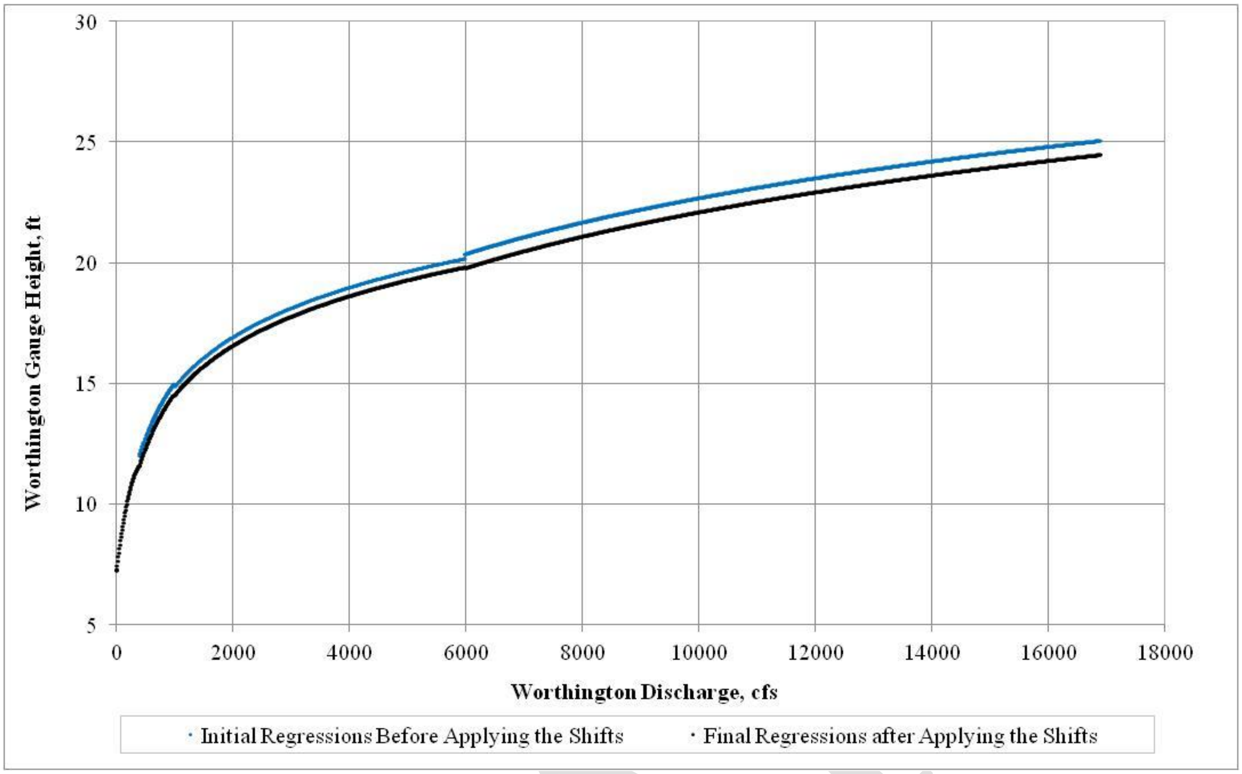



**Figure 19. Worthington Stage-Discharge Regression Shifts**

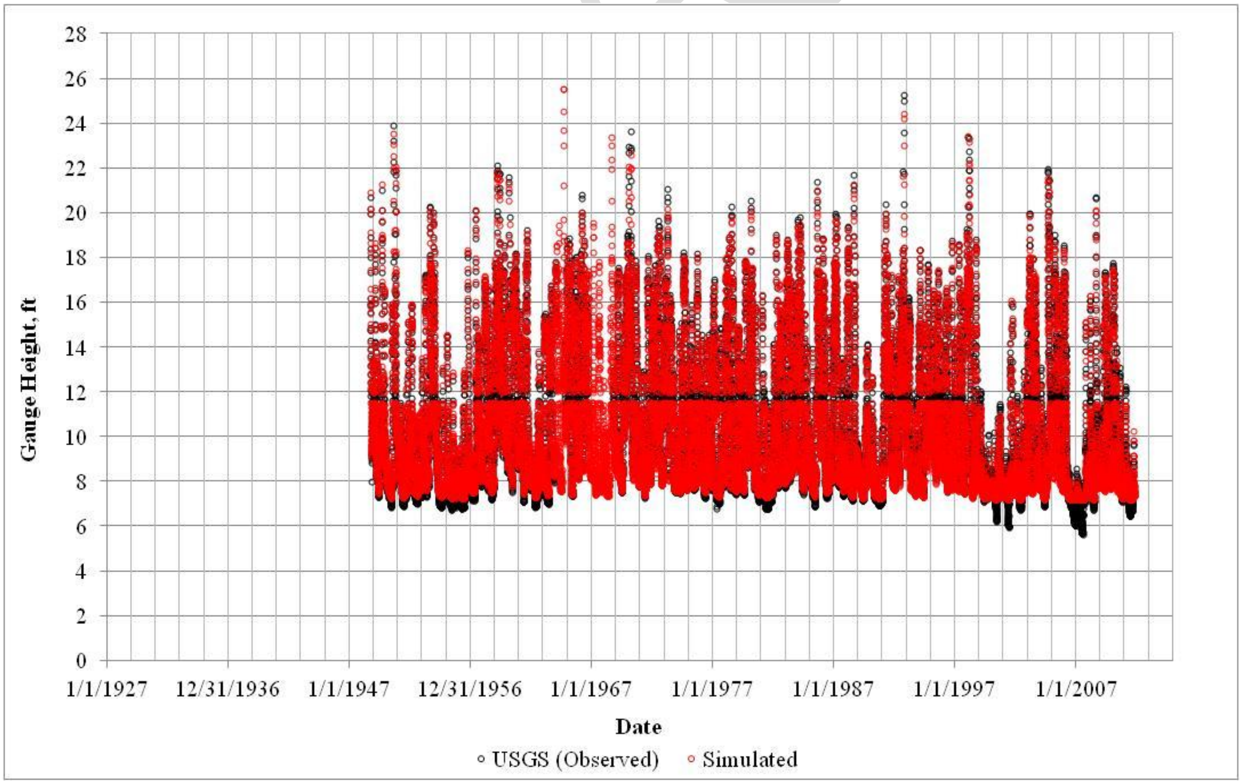



**Figure 20. Worthington Simulated and Observed Gauge Height before Applying Shifts**

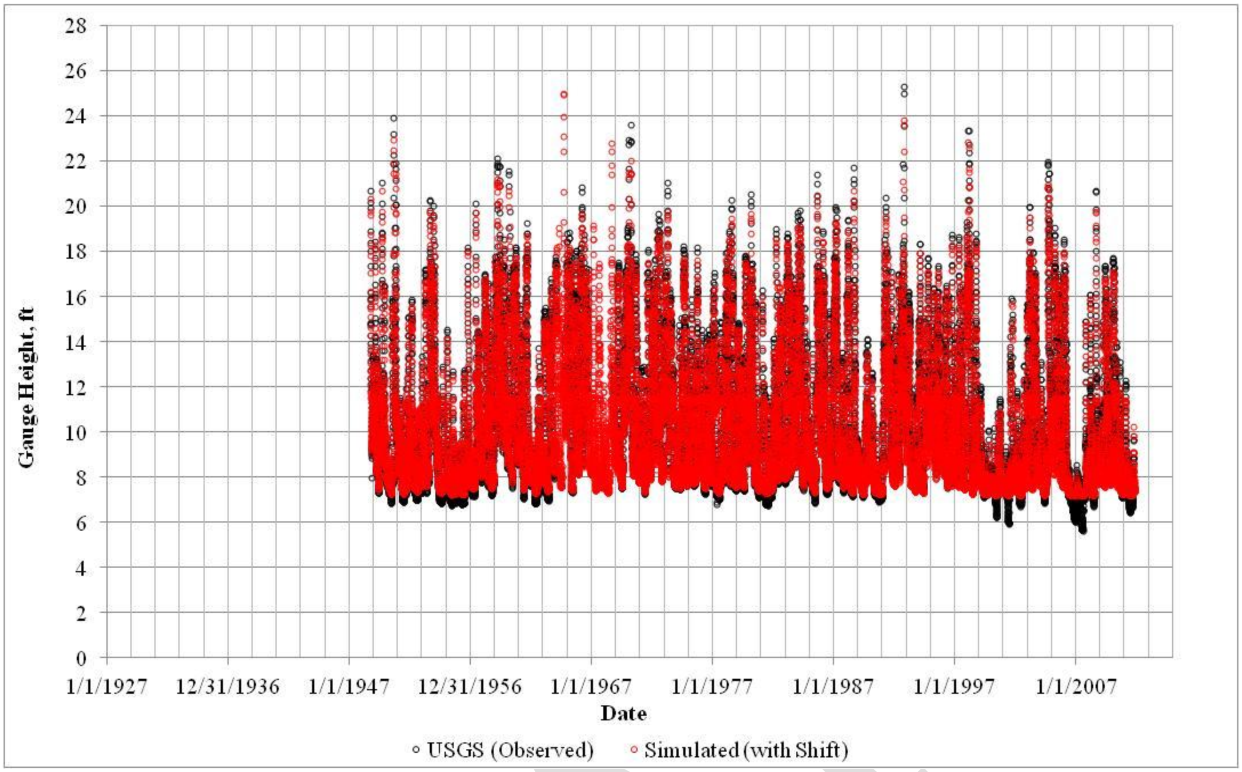



**Figure 21. Worthington Simulated and Observed Gauge Height after Applying Shifts**

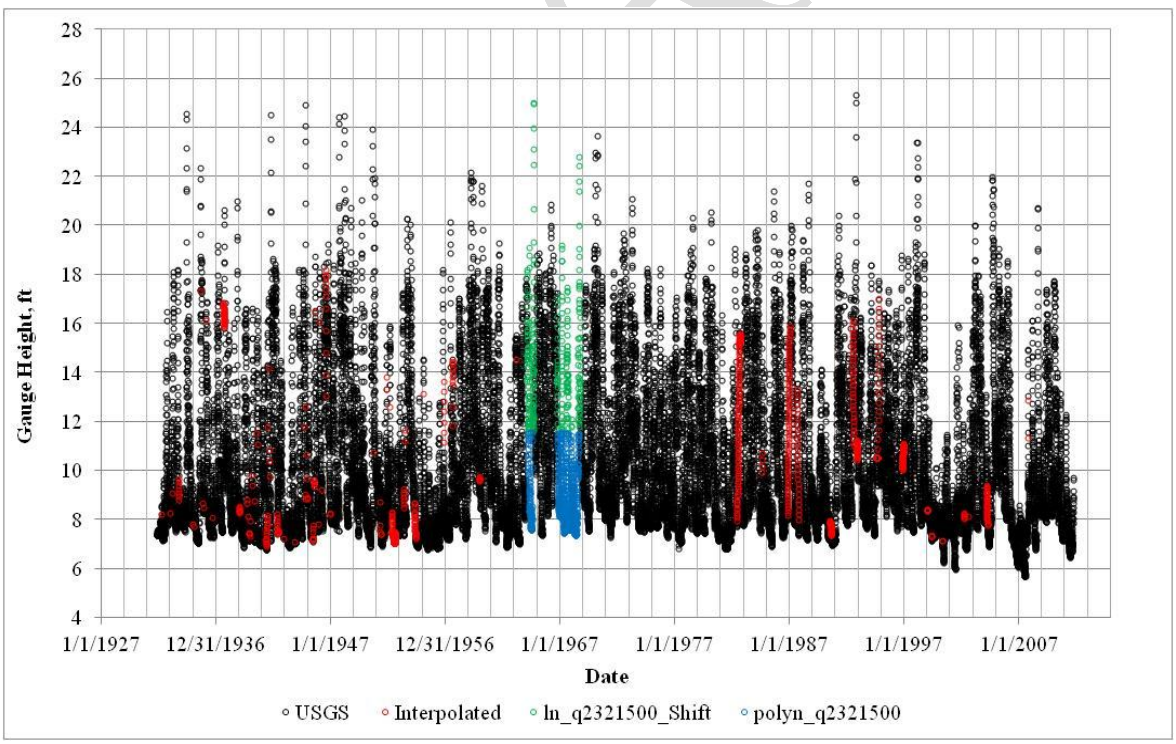



**Figure 22. Worthington Gauge Height (42.74 ft above NGVD29)**

The Worthington Springs discharge record included original USGS data from 10/1/1931 until 9/30/2011. The discharge record had no gaps in the period of interest (Figure 23).

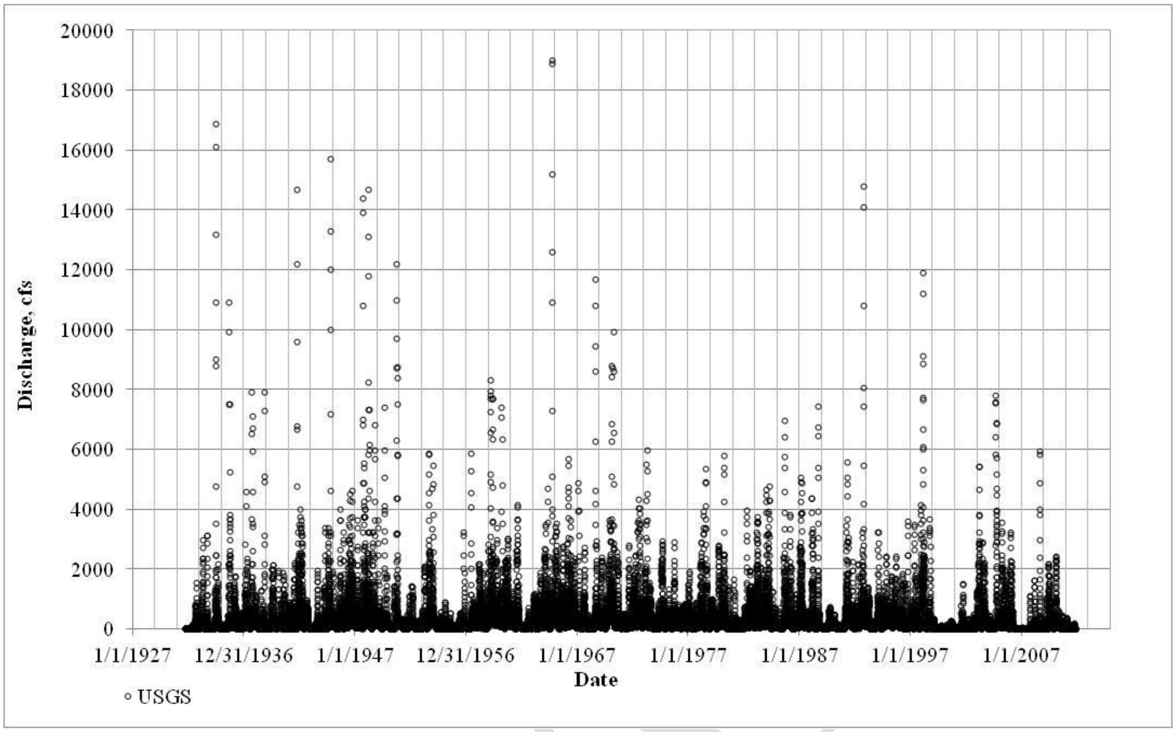



**Figure 23. Worthington Discharge (cfs)**

### Santa Fe River at O'Leno State Park (O'Leno by I-75) (2321898)

The Santa Fe River at O'Leno State Park (2321898) surface water station has no available discharge measurements and a short gauge height record, from 6/8/2010 until 9/30/2011. The gaps in the USGS O'Leno State Park gauge height record were not filled. Gauge height values are shown in Figure 24.

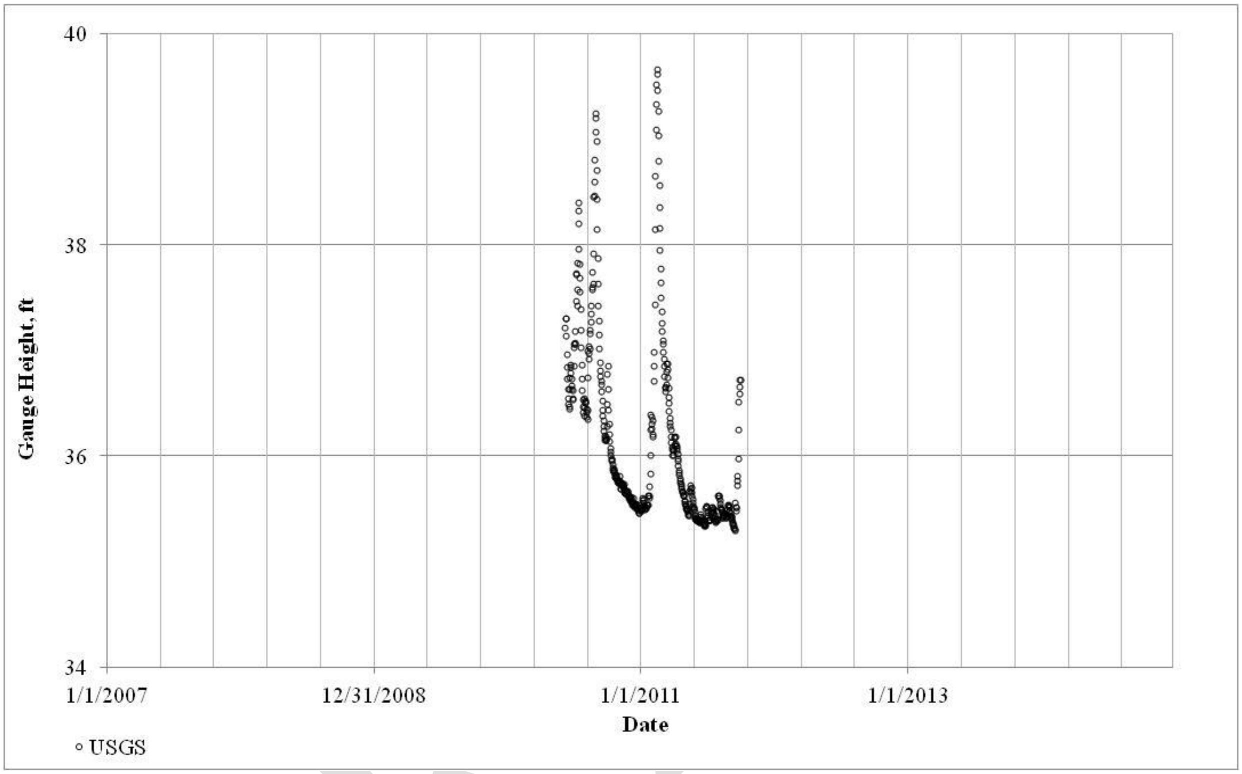



**Figure 24. O'Leno State Park Gauge Height (NAVD88)**

### Santa Fe River at O'Leno State Park by Footbridge (23218982 District Gauge)

The Santa Fe at O'Leno State Park by Footbridge (23218982) stage and flow data were obtained from the District. The data gaps were not filled. Gauge height values are shown in Figure 25. Discharge values are shown in Figure 26.

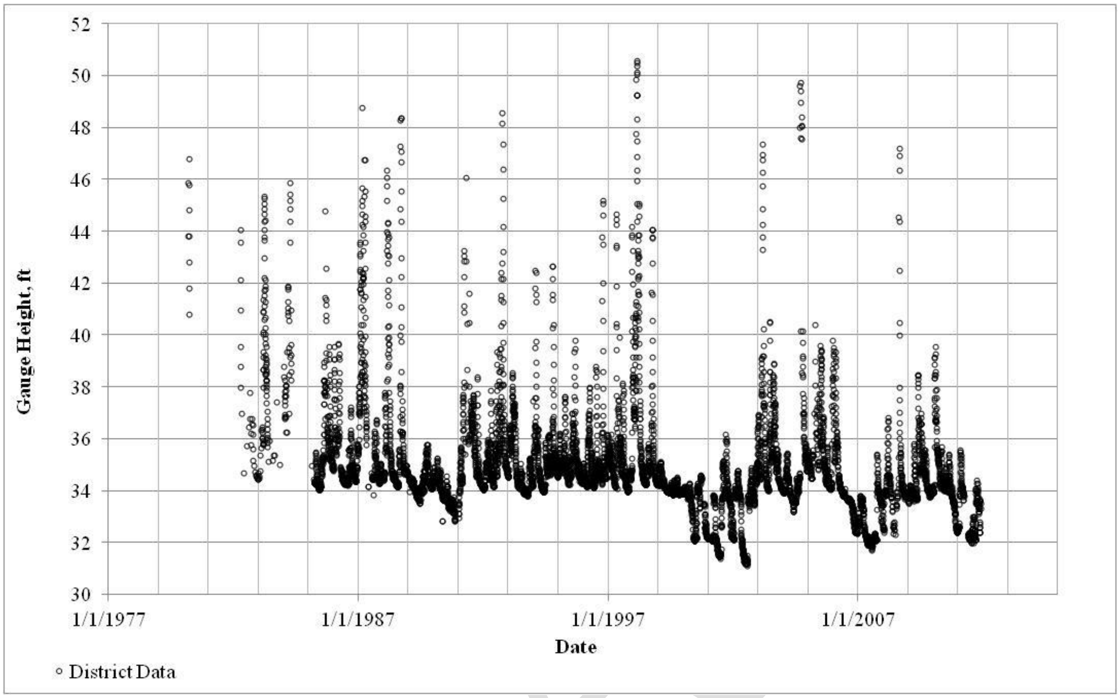



**Figure 25. O'Leno Footbridge Gauge Height (NGVD29)**

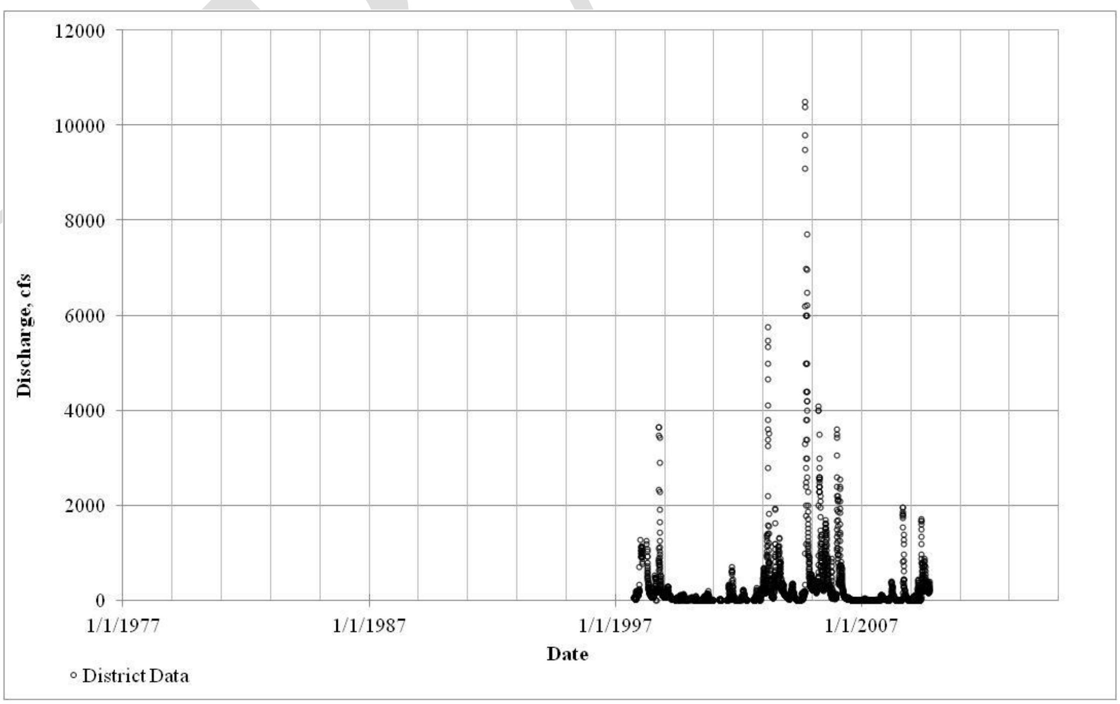



Figure 26. O'Leno Discharge (cfs)

### Santa Fe River at US Highway 441 near High Springs (2321975)

The Santa Fe River at US Highway 441 (441) gauge height record was not available from the USGS. District data was available from 11/2/2002 until 9/30/2011. The District data of flow and stage records from the Santa Fe at Worthington and Ft. White stations were utilized to develop a multiple linear regression (MLR) to predict 441 gauge height. The use of a MLR to predict a dependent variable is needed when the data cannot be described by a single variable and scientific knowledge indicates that more variables would be useful. These variables may or may not be correlated and independent of each other (Helsel and Hirsch, 2002). The MLR equation coefficients and model statistics are shown in Table 11 and Figure 27. Three intermittent gauge height values were filled using linear interpolation (3/26/2011, 4/8/2011, 4/9/2011). The gap filled gauge height record is shown in Figure 28.

**Table 11. Santa Fe at 441 Gauge MLR Summary**

| **MLR Coefficients and Significance** | | |
|---|---|---|
| | b | p-value |
| Intercept | 30.8380 | 0.000000 |
| Ft. White Gauge | 0.1476 | 0.000000 |
| Worthington Gauge | 0.01911 | 0.000001 |
| Ft. White Flow | 0.00087 | 0.000000 |
| Worthington Flow | 0.00030 | 0.000000 |
| **MLR Performance Summary** | | |
| RMSE | 0.319 | |
| Average Residual | 4.985e-15 | |
| $R^2$ | 0.9205 | |

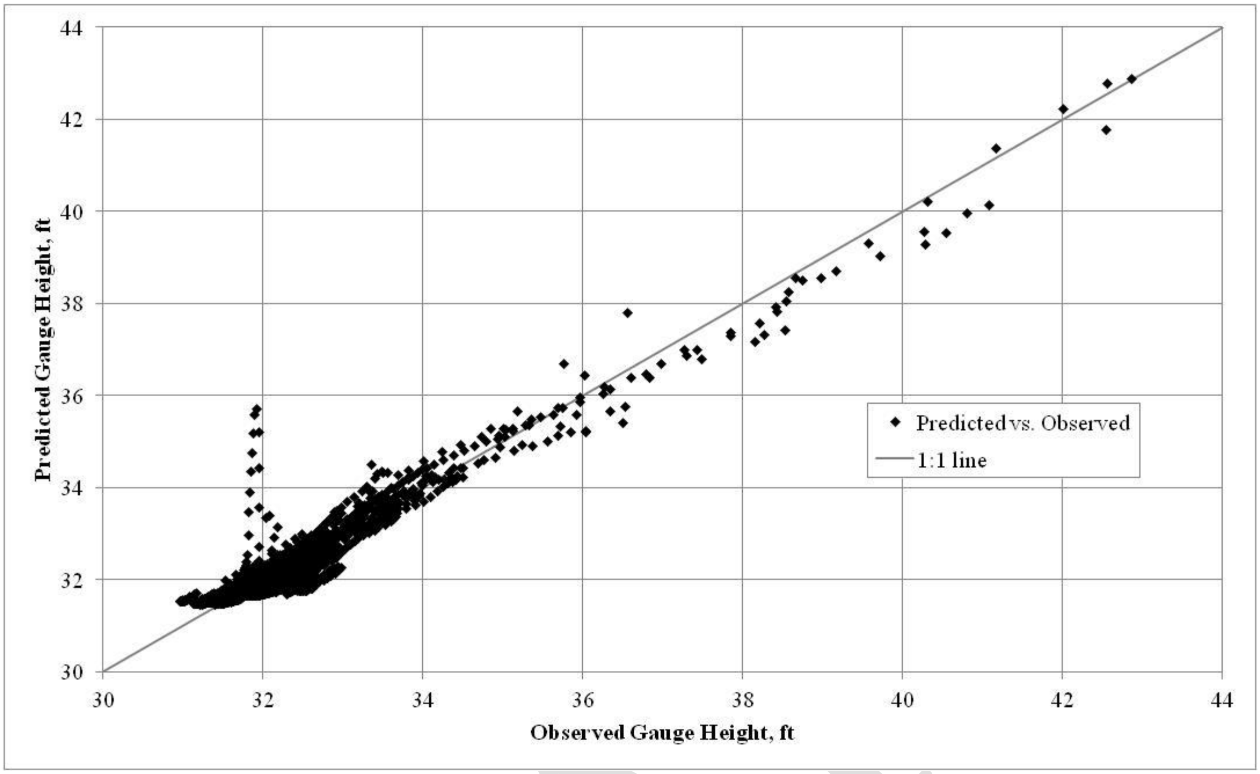



**Figure 27. MLR Highway 441, Predicted versus Observed Model Fit**

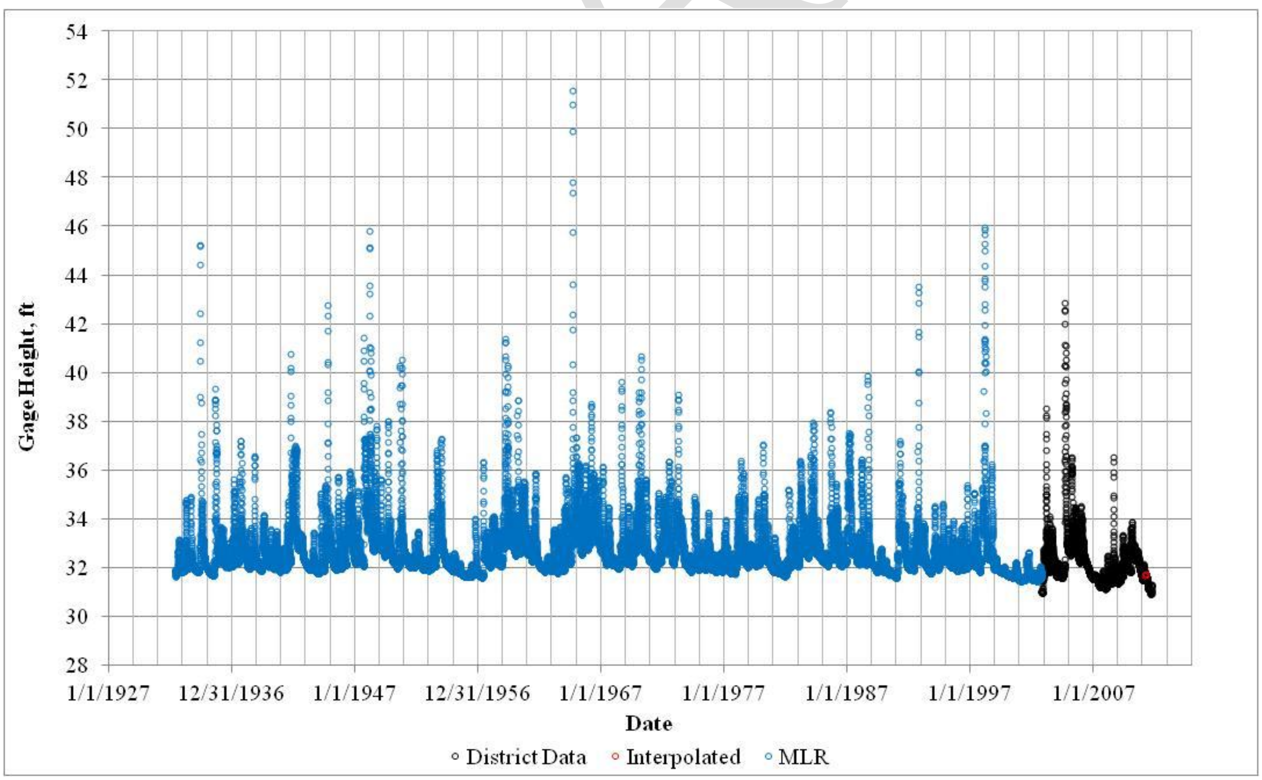



**Figure 28. 441 Gauge Height (NGVD29)**

The Santa Fe River at US Highway 441 discharge record from the USGS was sparse and included a period of record from 10/1/1992 through 9/30/2002. Additional District data was added to the Highway 441 USGS discharge record extending the record until 9/13/2010. A MLR was developed to predict Santa Fe at 441 flow using 2 explanatory variables: Santa Fe at Worthington Springs flow (upstream) and Santa Fe at Ft. White flow (downstream). The MLR summary and the MLR performance are shown in Table 12 and Figure 29, respectively. The resultant Highway 441 discharge record is shown in Figure 30.

**Table 12. Santa Fe at 441 Discharge MLR Summary**

| MLR Coefficients and Significance | | |
|---|---|---|
| | b | p-value |
| Intercept | -299.71 | 0.000000 |
| Ft. White Flow | 0.605 | 0.000000 |
| Worthington Flow | 0.396 | 0.000000 |
| **MLR Performance Summary** | | |
| RMSE | 276.3 | |
| Average Residual | 1e-12 | |
| $R^2$ | 0.8798 | |

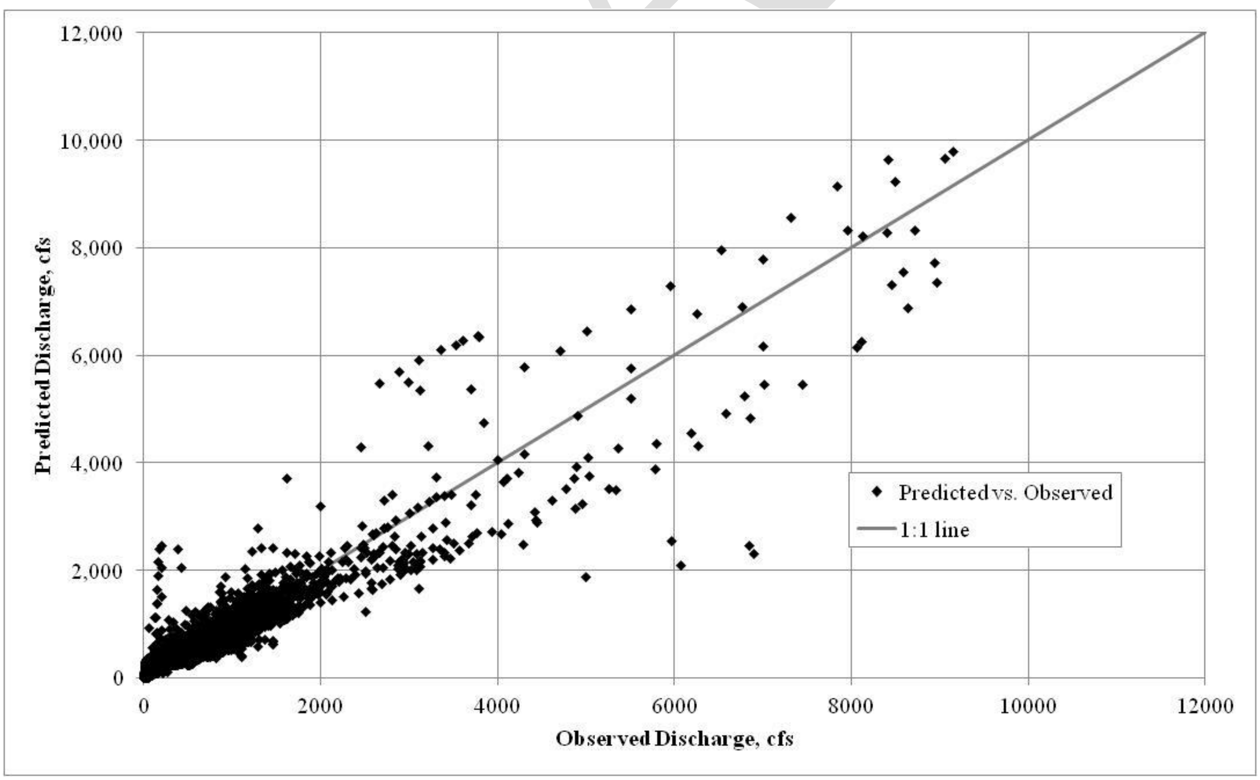



**Figure 29. Highway 441 MLR Performance Summary**

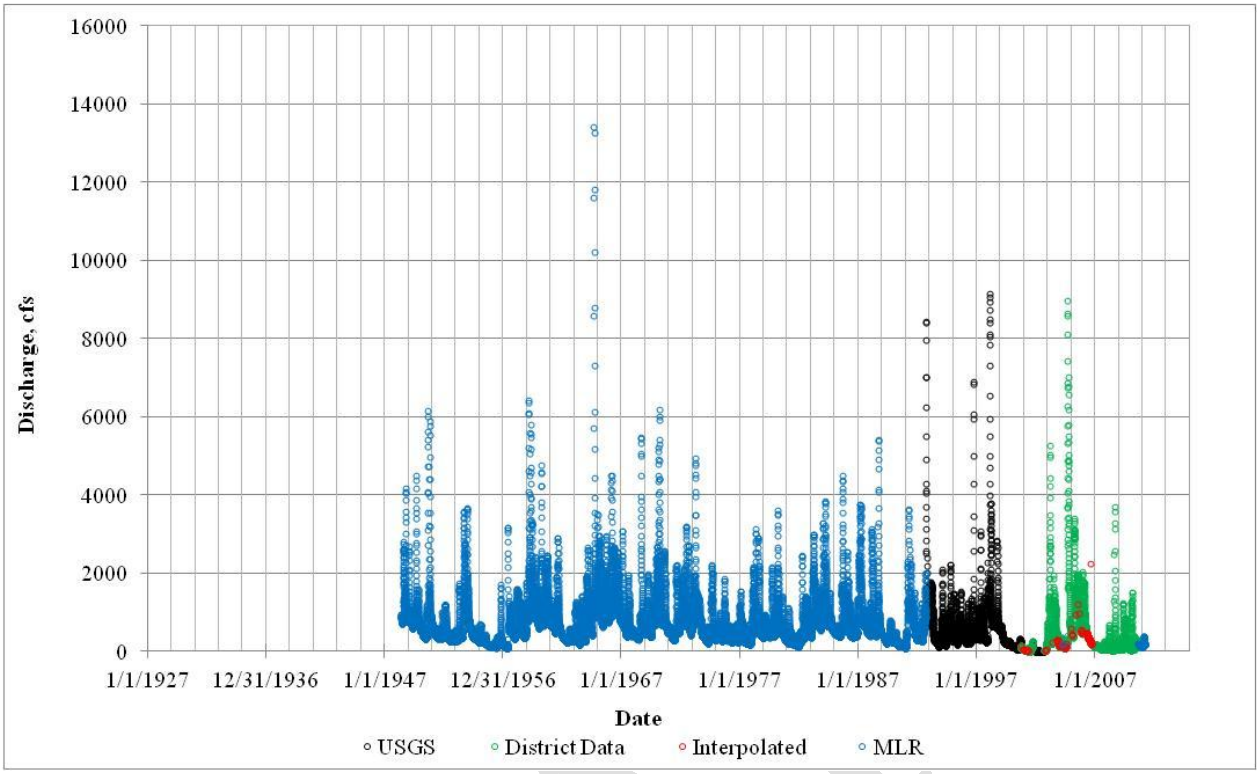



**Figure 30. Highway 441 Discharge (cfs)**

## Santa Fe River near Fort White (2322500)

The gaps in the Fort White USGS gauge height record were filled using linear interpolation and multiple linear regression (MLR). Linear interpolation was used to fill intermittent small gaps in the record. A multiple linear regression was developed using 2 explanatory variables: Santa Fe River at Worthington Springs surface water station flow (2321500) and Suwannee River near Bell surface water station gauge height (2323000). The fit of the MLR is shown in Figure 31, and the MLR equation and diagnostics are shown in Table 13. The use of a stage-discharge relationship for this station was also investigated. Because the fit of the stage-discharge relationship was poor, the MLR was utilized for gap filling. Although there is some scatter in the MLR fit, the model utilizes the best available data for gap filling. Additionally, the MLR was not utilized extensively since the Fort White stage record contained relatively small gaps as compared to other stations. The final filled gauge height is shown in Figure 32.

**Table 13. Santa Fe near Ft. White Gauge MLR Summary**

| **MLR Coefficients and Significance** | | |
|---|---|---|
| | b | p-value |
| Intercept | -0.712411 | 0.000000 |
| Bell Gauge | 0.208413 | 0.000000 |
| Worthington Springs Flow | 0.000565 | 0.000000 |
| **MLR Performance Summary** | | |
| RMSE | 0.7625 | |
| Average Residual | -0.00016 | |
| $R^2$ | 0.6892 | |

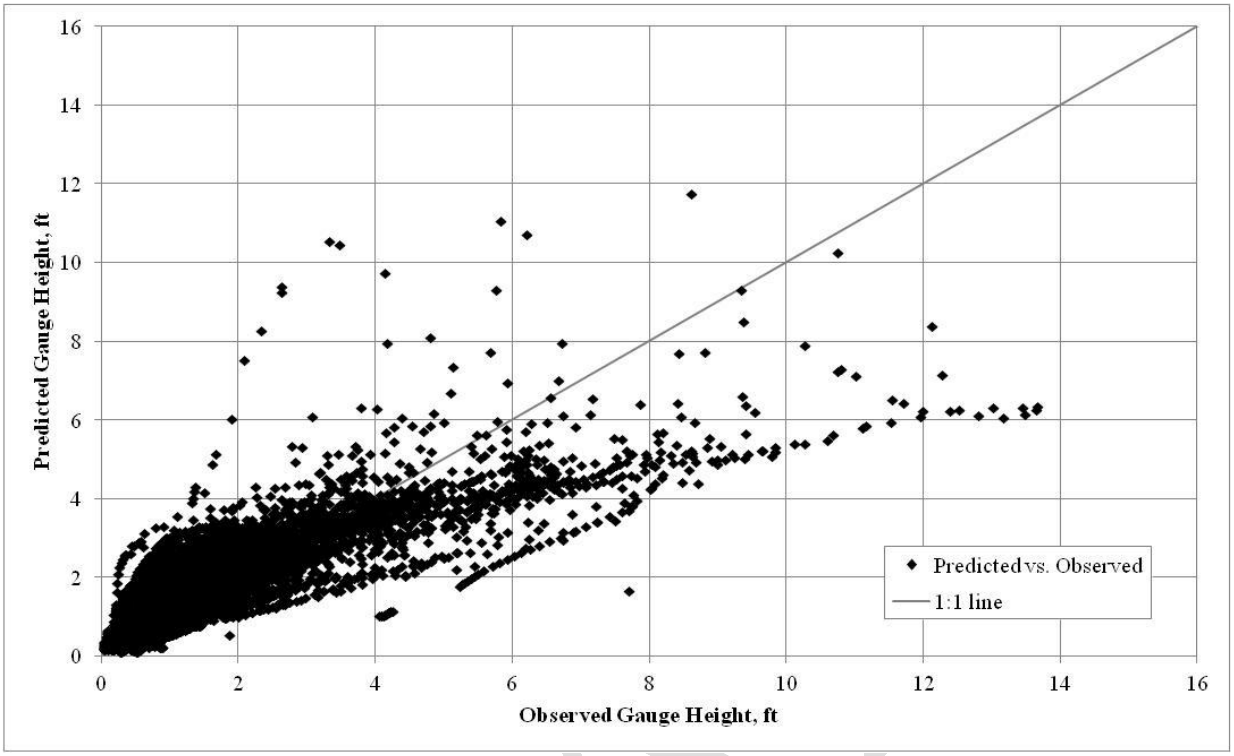



**Figure 31. Ft. White MLR Performance**

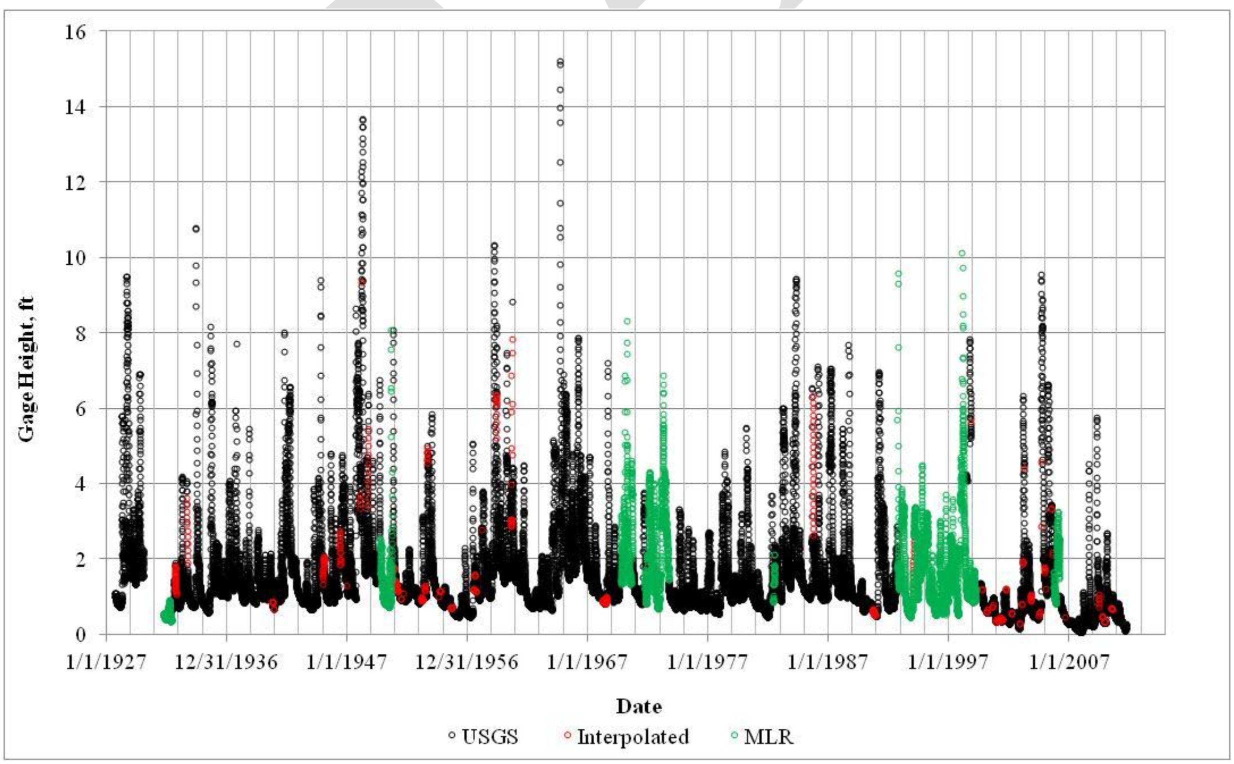



**Figure 32. Ft. White Gauge Height (20.86 ft above NGVD29)**

The Fort White discharge record consists of the USGS data from 10/1/1927 until 11/21/2011. There are no missing discharge values in the period of interest (6/4/1948-9/30/2011) (Figure 33).

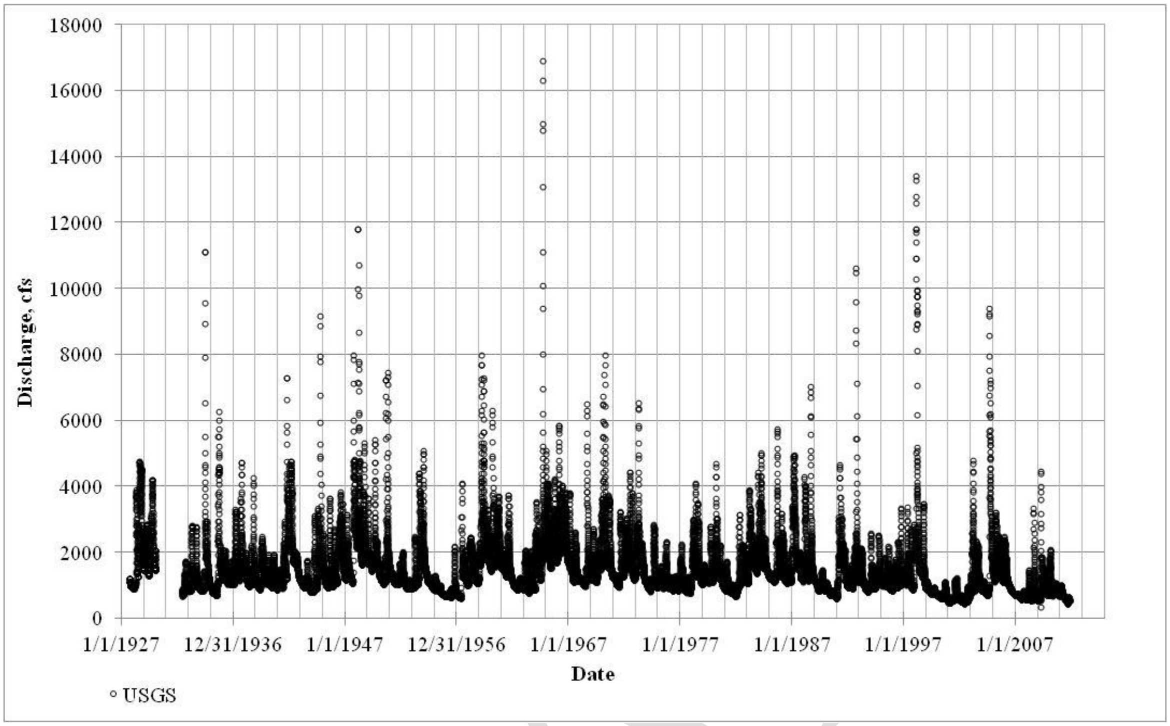



**Figure 33. Ft. White Discharge (cfs)**

## Santa Fe River above Ichetucknee River near Hildreth (Three Rivers) (2322703)

A simple linear regression, shown in Figure 34 and Table 14, was developed for gap filling of stage using stage data from gauge 2322800, Santa Fe River near Hildreth. The Three Rivers gauge observed stage data consisted of the USGS data and the stage data provided by the District. The resultant filled time series is shown in Figure 35.

**Table 14.Three Rivers Gauge Height SLR Performance Summary**

| Regression Performance Summary | |
|---|---|
| RMSE | 0.2303 |
| Average Residual | -0.1583 |
| $R^2$ | 0.9854 |

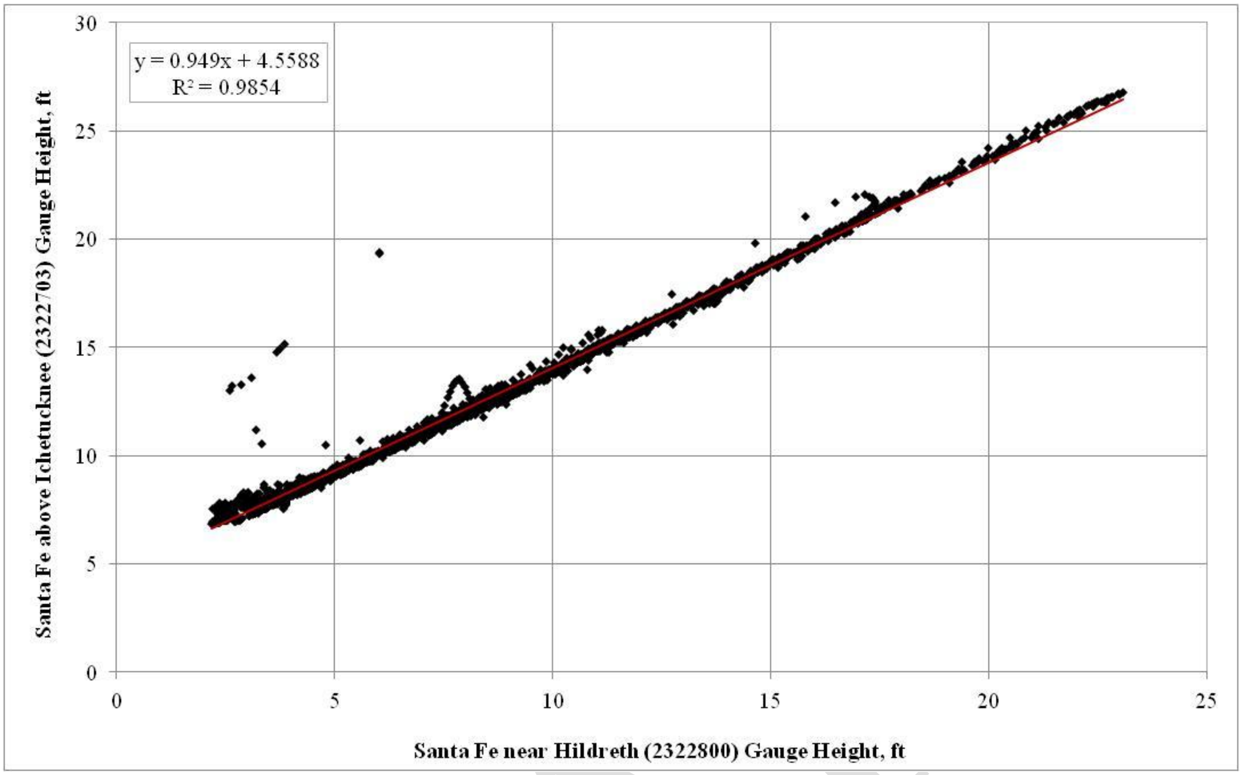



**Figure 34. SLR: Three Rivers (Above Ichetucknee) Gauge Height vs. Hildreth Gauge Height**

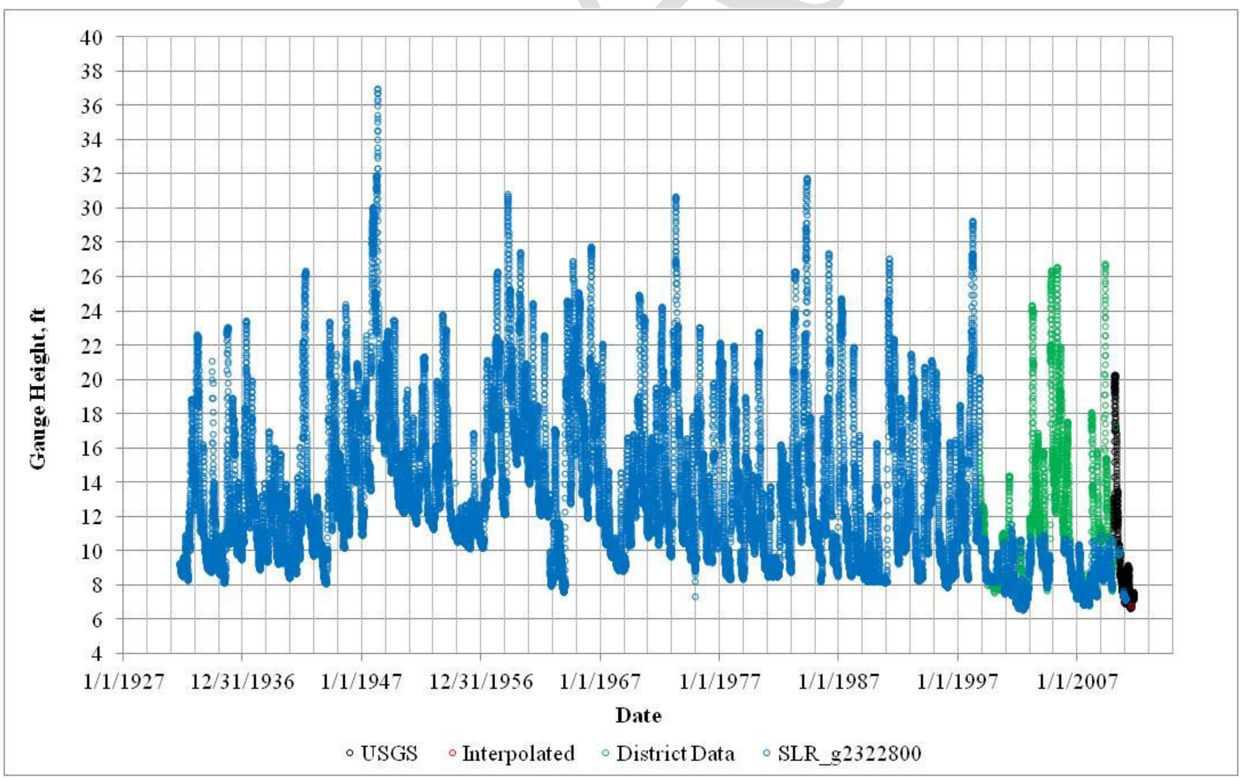



**Figure 35. Three Rivers (Above Ichetucknee) Gauge Height (NGVD29)**

There was no recorded discharge data for the Santa Fe River above Ichetucknee River near Hildreth station. Since no data was recorded, there was no available data for model construction or gap filling.

### Santa Fe River near Hildreth (2322800)

The gaps in the Hildreth USGS gauge height record, composed of USGS and District data, were filled using linear interpolation and multiple linear regression (MLR). A multiple linear regression was developed using 2 explanatory variables: Santa Fe River near Fort White surface water station gauge height (2322500) and Suwannee River at Branford surface water station gauge height (2320500). The fit of the MLR is shown in Figure 36 and Table 15. The resultant gauge height record is shown in Figure 37.

**Table 15. Santa Fe at Hildreth Gauge MLR Summary**

| MLR Coefficients and Significance | | |
|---|---|---|
| | b | p-value |
| Intercept | 0.7039 | 0.000000 |
| Ft. White Gauge Height | 0.2813 | 0.000000 |
| Branford Gauge Height | 0.7833 | 0.000000 |
| **MLR Performance Summary** | | |
| RMSE | 0.3485 | |
| Average Residual | 3.50E-14 | |
| $R^2$ | 0.948 | |

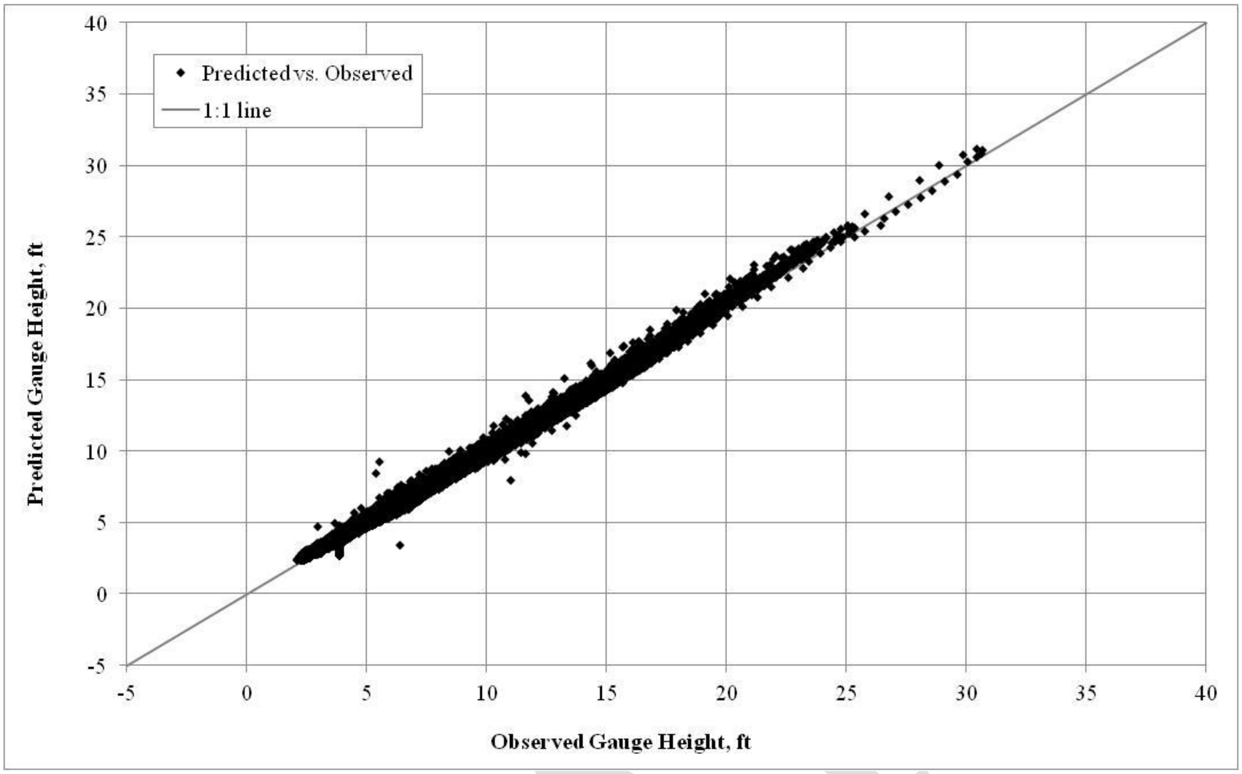



**Figure 36. Santa Fe Hildreth Gauge Height MLR Fit**

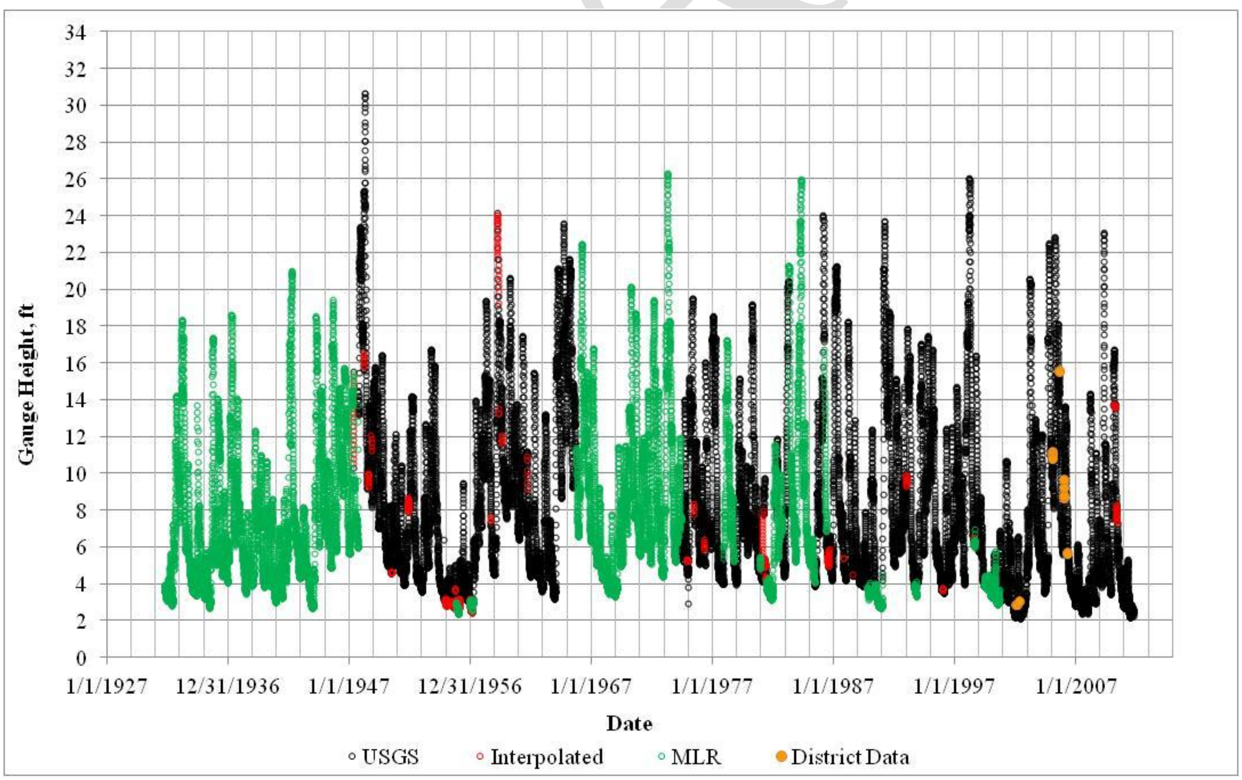



**Figure 37. Santa Fe Hildreth Gauge Height (3.5 ft above NGVD29)**

The Santa Fe River near Hildreth discharge record from USGS was sparse and included data from 11/1/2000 through 11/21/2011. Several relationships between gauge height and flow on Hildreth were fit, including a logarithmic and an exponential fit. In addition, grouping of data by flow and averaging of gauge height produced inadequate results. A polynomial fit of all available discharge and gauge height data at Hildreth produced a reasonable $R^2$ value, given the sparse discharge data at this station. The plot of the polynomial fit and the resultant Hildreth discharge record are shown in Figures 38 and 39. Regression statistics are shown in Table 16.

**Table 16. Santa Fe at Hildreth Discharge MLR Summary**

| **MLR Coefficients and Significance** | | |
|---|---|---|
| | b | p-value |
| Intercept | 384.0725 | 0.000000 |
| Hildreth Gauge Height | 25.5988 | 0.000000 |
| Ft. White Flow | 0.9200 | 0.000000 |
| **MLR Performance Summary** | | |
| RMSE | 298.4625 | |
| Average Residual | -1.6E-12 | |
| $R^2$ | 0.8803 | |

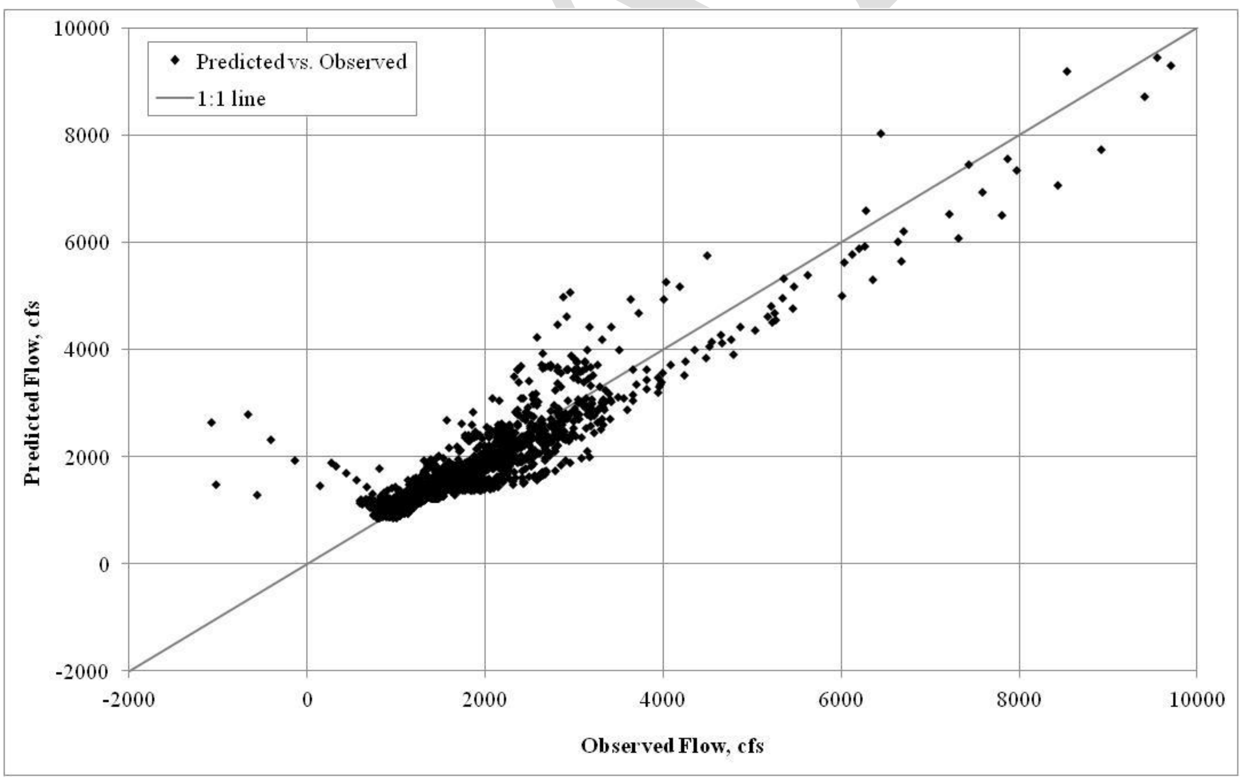



**Figure 38. Santa Fe Hildreth Discharge MLR Fit**

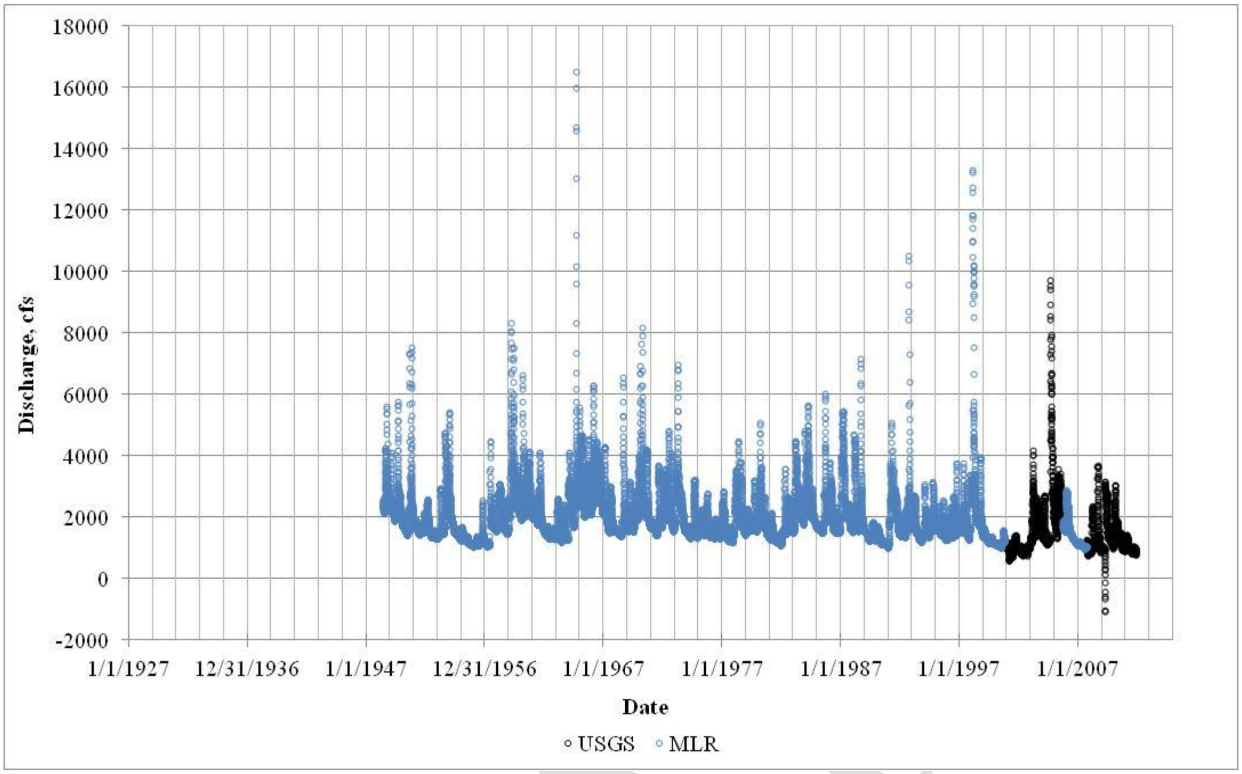



**Figure 39. Santa Fe Hildreth Discharge (cfs)**

## Ichetucknee River Surface Water Stations

There are two USGS surface water stations located on the Ichetucknee River: Ichetucknee River at Dampier's Landing (2322698) and Ichetucknee River at Highway 27 (2322700) (Figure 40). As Tables 2 and 3 illustrate, the period of record was short for both stations, and there was little available USGS data. Additional discharge and gauge height measurements were available for the Ichetucknee River at Highway 27 station (2322700) as USGS field measurements. The USGS field measurements for the Ichetucknee River at Highway 27 consisted of manual monthly and bi-monthly streamflow and gauge height readings for a period of record from 1/23/1931 until 2/4/2002.

Missing discharge and gauge height values for the Ichetucknee River at Highway 27 gauge were filled starting in 6/4/1948. The period of record of the well - 41705001 was a limiting factor. Since well - 41705001 had the longest and the most complete record, it was utilized for model development when groundwater levels were needed.

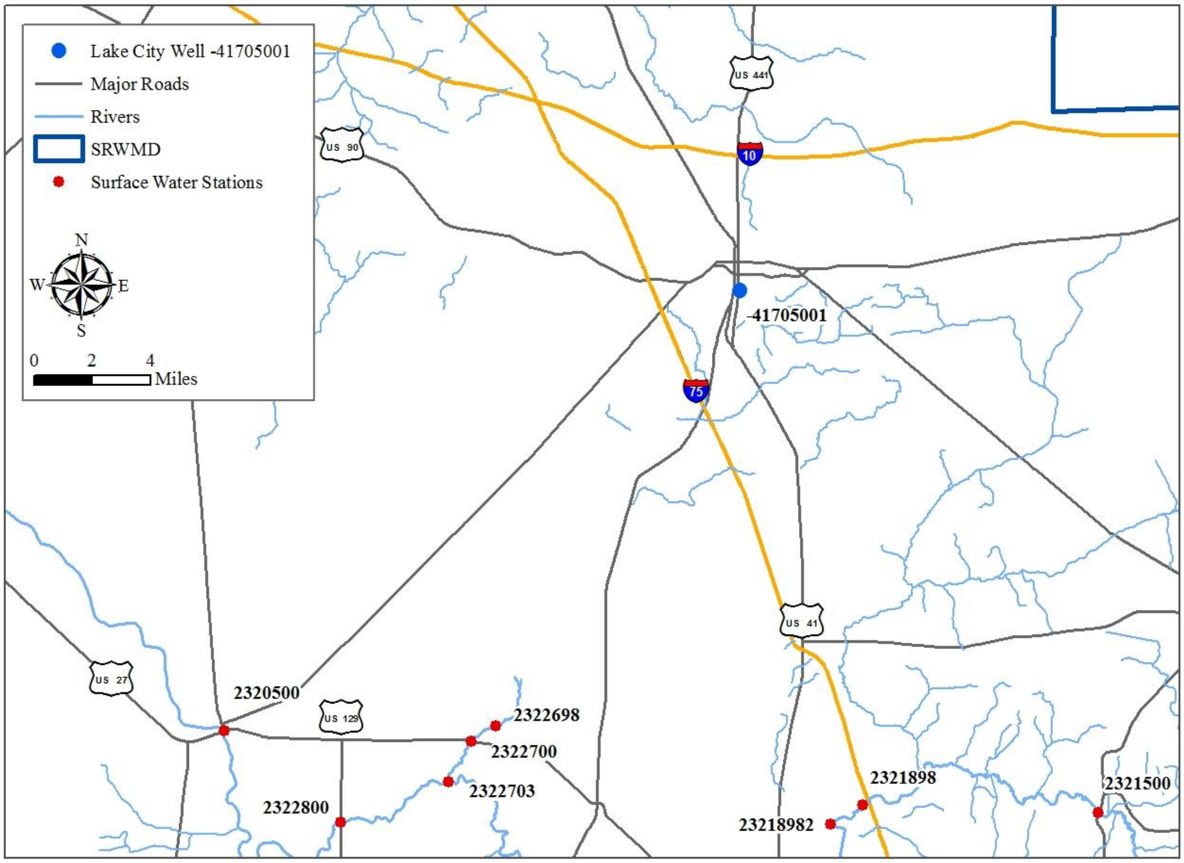



**Figure 40. Surface Water Stations on the Ichetucknee River**



## Ichetucknee River at Dampier's Landing near Hildreth (2322698)

Gauge height data was available for the Dampier's Landing near Hildreth station from 2/15/2002 until 9/30/2011 (Figure 41). The available flow record of the Dampier's Landing gauge spanned from 2/15/2002 until 4/3/2011 (Figure 42). No gap-filling was performed for this station for several reasons. The period of record for the gauge was limited, which would result in extensive gap-filling. More importantly, the Dampier's Landing gauge is located just upstream of the Highway 27 gauge, which has been rated as higher quality than Dampier's Landing. Flows at the two Ichetucknee stations are very similar. Since Dampier's Landing is upstream of Highway 27, it would be expected that the mean and median flows at Dampier's Landing would be slightly less than those at Highway 27. There were, however, occurrences of flow loss from upstream to downstream in the historical record. This flow loss was investigated by the District with the following conclusion (Coarsey, 2012):

> *"After speaking with USGS Staff it was determined that the most likely cause of the loss of flow was from a bias in the measurements at the upstream sites. All the upstream sites have a rating of poor due to additional error attributed to vegetation in the area of measurement. The only site without a significant*

*amount of vegetative interference is 02322700 (US27). Therefore 02322700 (US27) best approximates the actual flow in the Ichetucknee River."*

Since the US Highway 27 station serves as the best flow approximation of Ichetucknee River flow, gap-filling and analysis were conducted on the Highway 27 record.

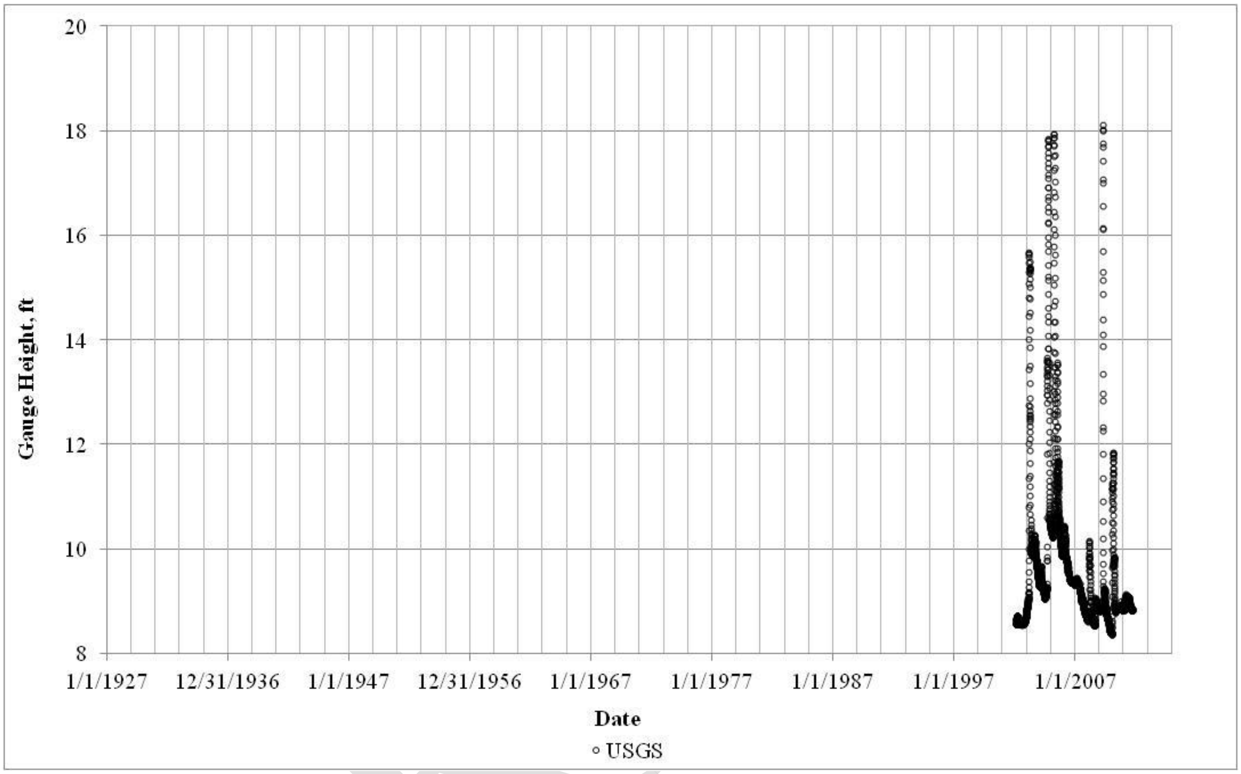



**Figure 41. Dampier's Landing Observed Gauge Height (8.62 ft above NGVD29)**

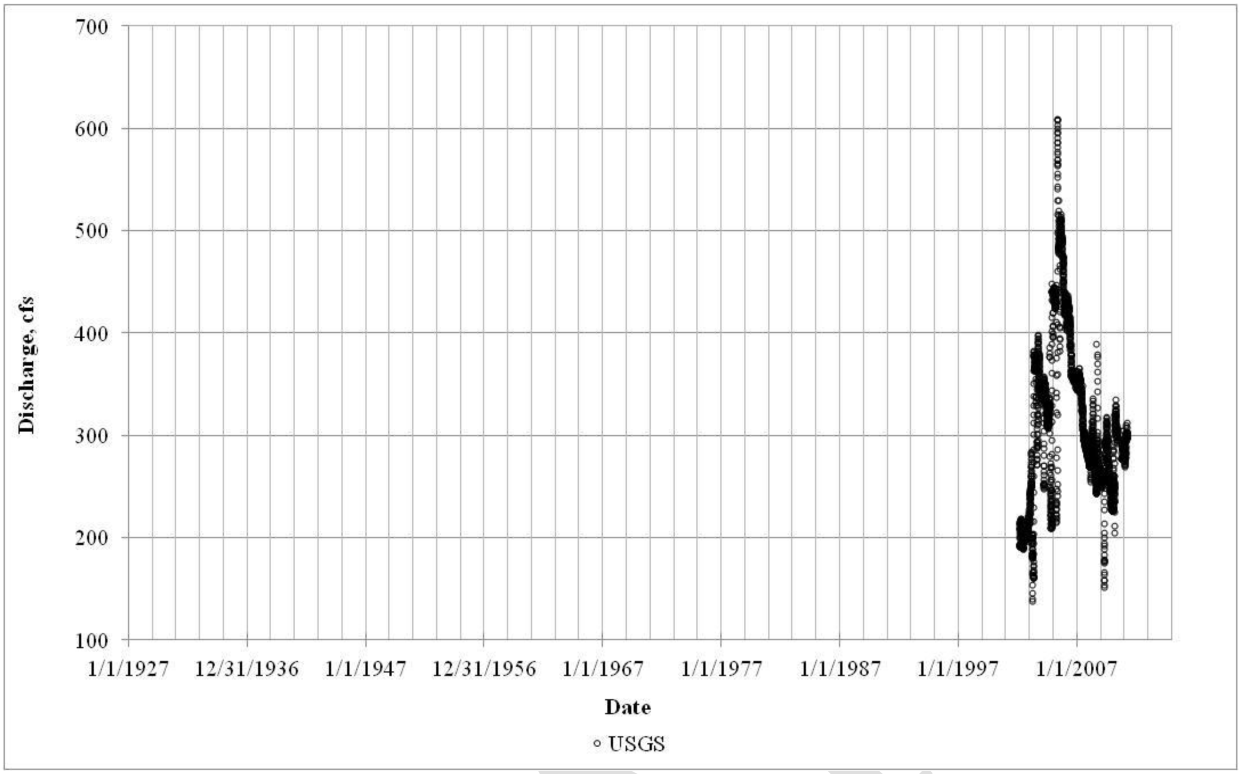



**Figure 42. Dampier's Landing Observed Discharge (cfs)**



## Ichetucknee River at Highway 27 near Hildreth (2322700)

An artificial neural network was utilized to develop a filled stage record for the Ichetucknee River at the Highway 27 gauge. The use of an ANN provided a better fit to the observed data than a multiple linear regression (MLR). The ANN was a multilayer perception (MLP) with 2 input nodes: well stage at FDOT well -41705001 in Lake City and gauge height at Santa Fe River near Hildreth (2322800). The fit for the ANN is shown in Figure 43, and performance statistics are shown in Table 17. The ANN was utilized to fill the gauge height record, as shown in Figure 44.

**Table 17. Highway 27 Gauge ANN Performance Summary**

| ANN Performance Summary | |
|---|---|
| RMSE | 0.1753 |
| Average Residual | -0.0043 |
| $R^2$ | 0.9934 |

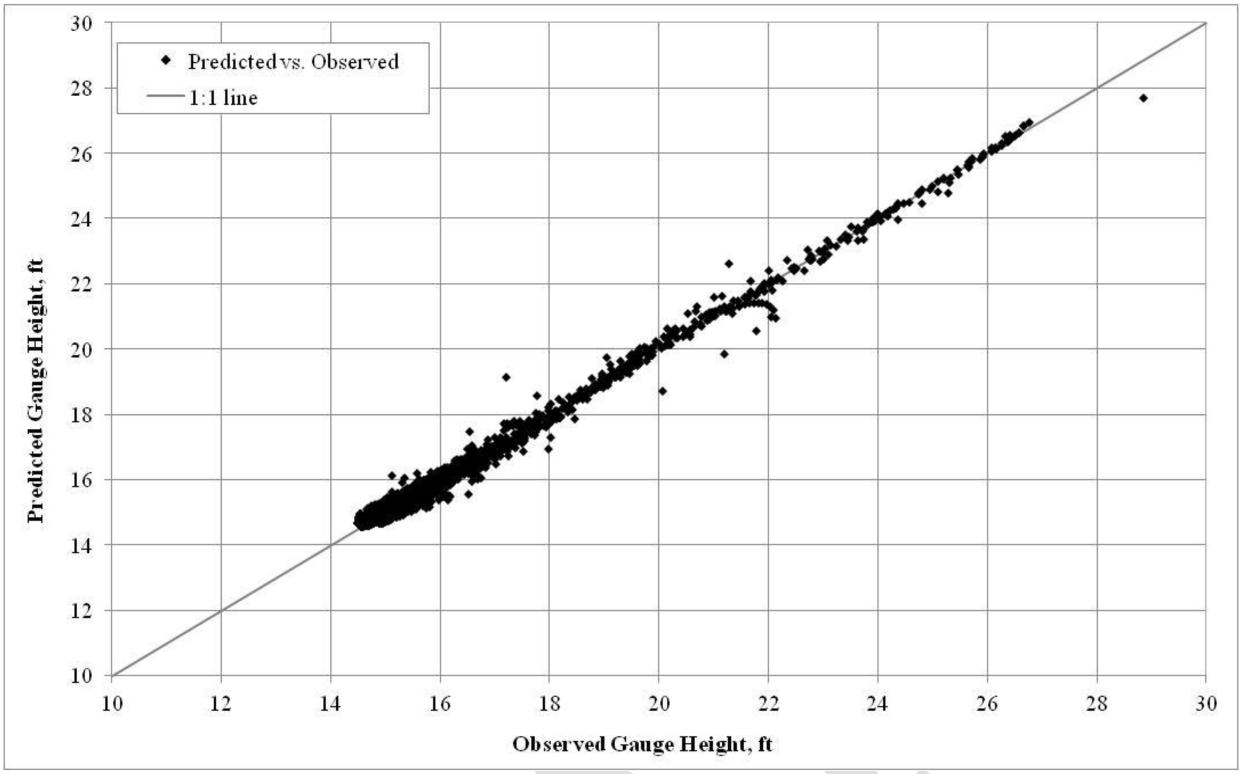



**Figure 43. Ichetucknee at Highway 27 Gauge Height ANN Model Fit**

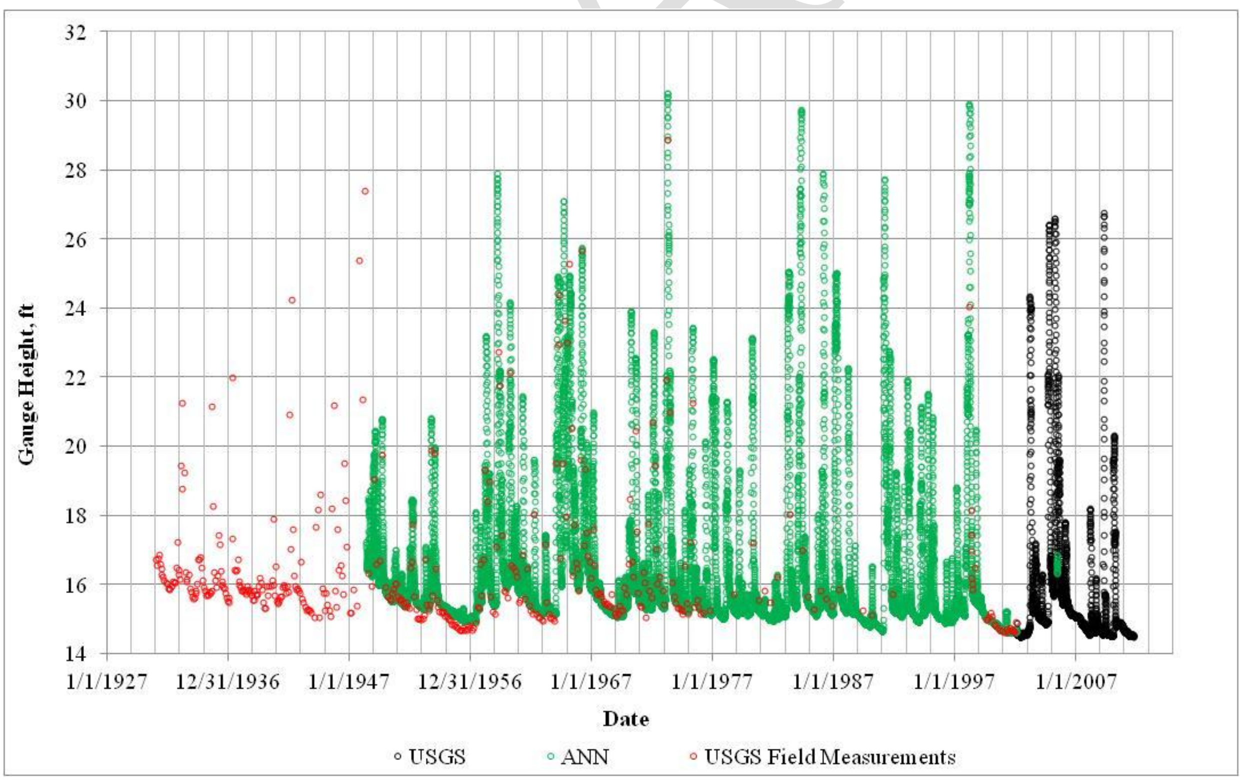



**Figure 44. Ichetucknee at Highway 27 Final Filled Gauge Height (NGVD29)**

The flow record for this gauge was also filled using several methods. The available flow period of record for daily flow was from 2/5/2002 until 4/4/2011. Additional sporadic measurements were available prior to 2002 from USGS field measurements. The USGS field measurements ranged in frequency from monthly to longer gaps spanning multiple months. Based on the availability of the observed data, the data set was broken up into three periods:

- Prior to 1975, when USGS field measurements were frequent,
- From 1975 through 2000, when measurements were infrequent (with 54 measurements during the measurement period), and
- After 2000, when measurement frequency was high (daily means).

For the periods with high measurement frequency (prior to 1975 and after 2000), linear interpolation was utilized to produce a daily flow time series. The largest gap filled with interpolation was a 125-day gap from July 20, 1948 to November 21, 1948. For the 1975 through 2000 period, a simple linear regression was developed using a local well level (well -41705001). The fit of the SLR is shown in Figure 45. The SLR coefficients and performance summary are shown in Table 18. The filled flow record is shown in Figure 46.

**Table 18. Ichetucknee at Highway 27 Flow SLR Performance Summary**

| SLR Performance Summary | |
|---:|:---:|
| RMSE | 39.0525 |
| Average Residual | 1.79e-13 |
| $R^2$ | 0.6048 |

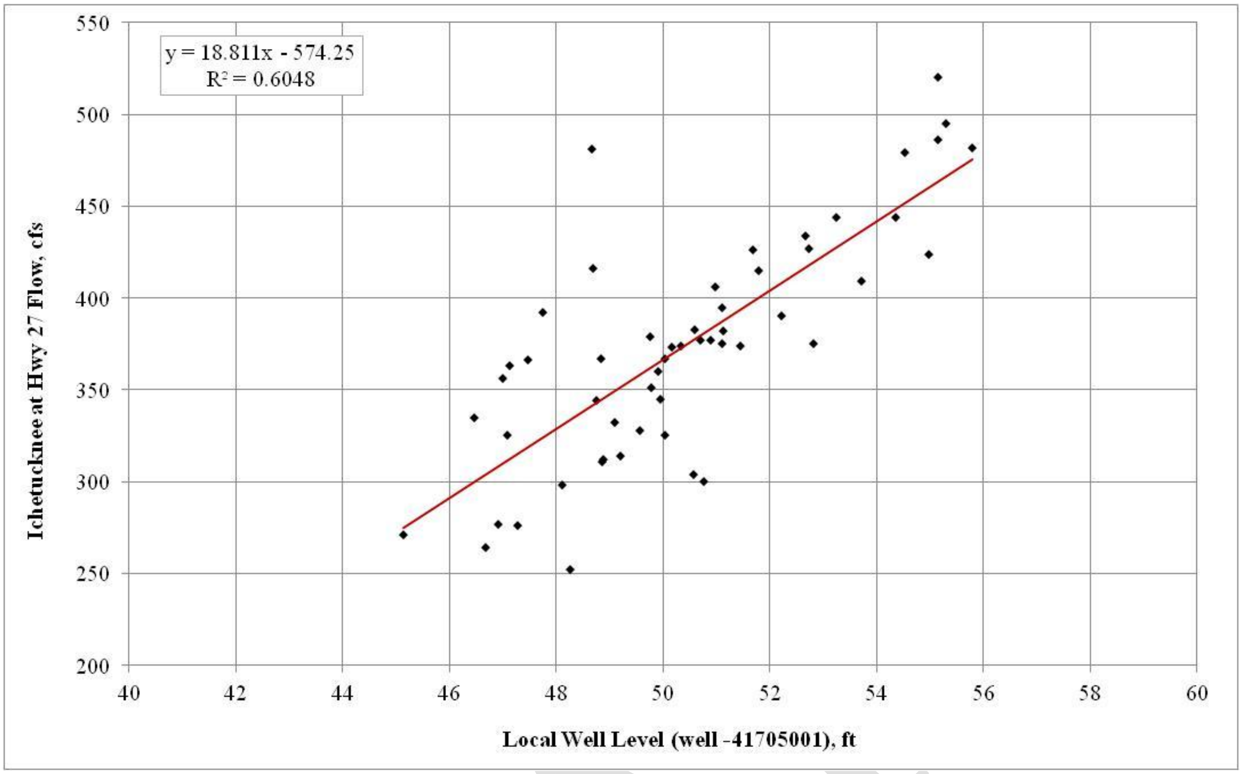



**Figure 45. SLR: Ichetucknee at Highway 27 Discharge vs. Local Well Level (well -41705001)**

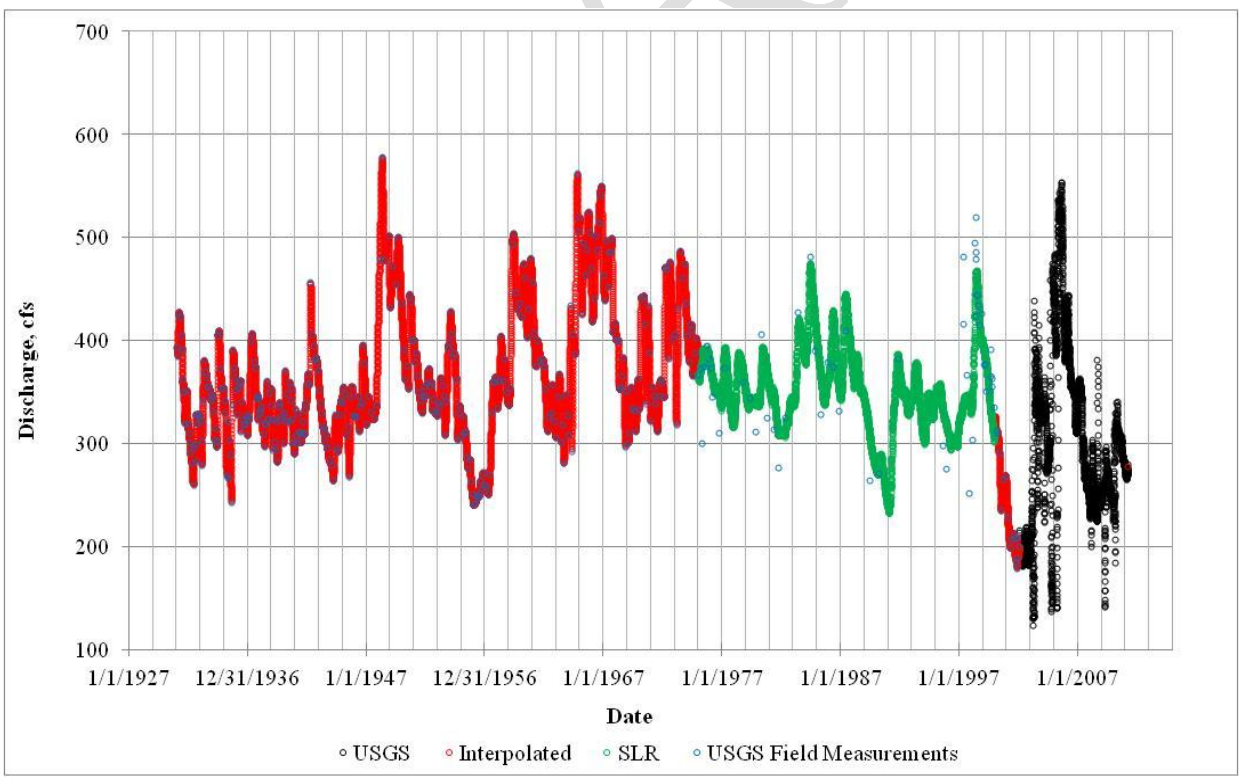



**Figure 46. Ichetucknee at Highway 27 Discharge (cfs)**

## Database

An Access database consisting of stage and flow records was created based on the USGS data, the District data, and the filled data. The database contains 10 tables which include a station list of the USGS surface water sites, gauge height data, discharge data, well data, rainfall data for two nearby stations, discharge and gauge height data history, and quality flags. After use of this data for exploratory and trend analysis, estimated baseflow and adjusted baseflow were computed and added to the database, each in their own table. Methods and results for the baseflow and unimpaired flow analyses are provided in subsequent sections of this report. Descriptions of the tables' fields are provided below.

*USGS_12_Stations* table includes a list of the USGS and the District's surface water stations, stations' names, longitude and latitude, horizontal and vertical datums. Table 19 shows the table fields and fields' descriptions. Descriptions of the fields for each table are also accessible via design view mode for each table in the Access database.

**Table 19. *USGS_12_Stations* Table Fields**

| Field Name | Description |
|---|---|
| StationID | USGS or SRWMD gauging station number |
| USGSNAME | USGS or SRWMD gauging station name |
| LATDEC | Latitude |
| LONGDEC | Longitude |
| HORIZDATUM | Horizontal datum |
| VERTDATUM | Vertical datum |

The *DISHARGE_NO_NULL_VALUES* and *GAGE_NO_NULL_VALS* tables store the completed (filled) discharge and gage height data. Field descriptions for each table are shown in Tables 20 and 21.

**Table 20. *DISHARGE_NO_NULL_VALUES* Table Fields**

| Field Name | Description |
|---|---|
| StationID | USGS or SRWMD gauging station number |
| Date | Date |
| Discharge | Observed or filled discharge, cubic feet per second |
| ErrorFlow | Qualifying data flag; refer to DATA_QUAL table |

**Table 21. GAGE_NO_NULL_VALS Table Fields**

| Field Name | Description |
|---|---|
| StationID | USGS or SRWMD gauging station number |
| Dat | Date |
| GageHeight | Gage height, feet; refer to USGS_12_Stations [VERTDATUM] for vertical datum of gage |
| ErrorGage | Qualifying data flag; refer to DATA_QUAL table |

The *DATA_QUAL* table stores data flags of observed and filled discharge and gage height values. A description of the field names is shown in Table 22. The qualifying data flag abbreviation and description are shown in Table 23. Data qualifiers "A", "e", and "P" were flags in the original USGS data. In general, data with any other qualifier falls into 1 of 3 groups: filled data, a District measurement, or a USGS field measurement.

**Table 22. DATA_QUAL Table Fields**

| Field Name | Description |
|---|---|
| DataQual | Qualifying data flag; found in GAGE_NO_NULL_VALS [ErrorGage] and DISCHARGE_NO_NULL_VALUES [ErrorFlow] |
| QualDescrip | Description of [DataQual] |

**Table 23. Qualifying Data Flag Description**

| DataQual | Use | QualDescrip |
|---|---|---|
| A | Flow, Gage Height | USGS - Approved for publication |
| ANN | Gage Height | Artificial Neural Network |
| District data | Flow, Gage Height | Data provided by the SRWMD |
| e | Flow, Gage Height | USGS - Value has been estimated |
| interp | Flow, Gage Height | Linear interpolation |
| ln_q#_Shift | Gage Height | Natural log fit with a shift with discharge record of indicated station |
| M | Flow, Gage Height | Multiple linear regression |
| P | Flow, Gage Height | USGS - Provisional data subject to revision |
| polyn_g# | Flow | Polynomial fit with gage height record of indicated station |
| polyn_q# | Gage Height | Polynomial fit with discharge record of indicated station |
| SLR | Flow | Simple linear regression |
| SLR_g# | Gage Height | Simple linear regression with gage height record of indicated station |
| USGS Field Measurement | Flow, Gage Height | USGS - field measurement |

The *EstBaseflow* and *UnimparedBF* tables store estimated baseflow and adjusted baseflow time series. Their respective fields are described in Tables 24 and 25.

**Table 24. *EstBaseflow* Table Fields**

| Field Name | Description |
|---|---|
| Dat | Date |
| Baseflow | Estimated baseflow, cubic feet per second |
| Sta | USGS gauging station number |

**Table 25. *UnimparedBF* Table Fields**

| Field Name | Description |
|---|---|
| Date | Date |
| BF_Unimpaired | Adjusted estimated baseflow, cubic feet per second |
| Station | USGS gauging station number |

The *w-41705001* table contains local well data. Field descriptions of the *w-41705001* table are shown in Table 26.

**Table 26. *w-41705001* Table Fields**

| Field Name | Description |
|---|---|
| StationID | District well ID |
| DAT | Date |
| VAL | Piezometric head, feet |
| QUAL | Data flag |

Field descriptions of the *GAINESVILLE_FILLED_DAILY_RAIN_RECORD* and *LakeCity_2E_FILLED_DAILY_RAIN_RECORD* tables containing observed rainfall data are provided in Tables 27 and 28, respectively.

**Table 27. GAINESVILLE_FILLED_DAILY_RAIN_RECORD Table Fields**

| Field Name | Description |
|---|---|
| Dat | Date |
| DailyRain | Observed and filled rainfall, inches |

**Table 28. LakeCity_2E_FILLED_DAILY_RAIN_RECORD Table Fields**

| Field Name | Description |
|---|---|
| Dat | Date |
| ID20_ID103_ID21 | Observed and filled rainfall, inches |

The *History* table contains the history of changes made to the database (Table 29).

**Table 29. History Table Fields**

| Field Name | Description |
|---|---|
| Dat | Date of entry |
| Description | Entry description |
| Initials | Initials |

## Exploratory Data Analysis and Trend Analysis

Examination of the flow time series for trends allows for characterization of the streamflow and baseflow as monotonic or non-monotonic. The majority of trend analysis tests are based on the assumption that the trend is monotonic over the time period of interest. Monotonic trends can be defined as a gradual and continued trend over time that is either positive (increasing) or negative (decreasing) (Helsel and Hisrch, 2002). Mathematical techniques can be used in conjunction with visual inspection to determine if the monotonic assumption is valid, and this should be done before any other type of trend analysis. Locally weighted scatterplot smoothing (LOWESS) is a technique that can help visualize overall trends in a time-series, and can mathematically identify times of change or "break-points" (Aly and Biggs, 2010).

Exploratory data analysis was conducted on all filled flow time series in order to determine if there were similarities between the flow stations in terms of periods of negative and positive trends. The same analysis period (1948 through 2011) was utilized for all stations in order to allow for direct comparison of the flow data trends between stations. LOWESS was plotted for each flow station using mean annual flows, as shown in Figures 47 and 48. The LOWESS line is shown in blue through each of the time series. This line was based on a span parameter of 0.6. The span parameter, which varies from 0 to 1, is defined as the fraction of the dataset that is utilized to calculate the regression over the moving window. A span of 0.6 means that the data window for the regression will include 60% of the data. A larger span will result in a smoother LOWESS line, whereas a smaller span will fit the data more closely and result in the identification of many breakpoints. For this case, it was desired to identify multi-decadal trends, and hence, a larger span was selected. The selection of a larger span is preferable because it avoids over-fitting the data and generating multiple breakpoints. As shown, many stations exhibit an increasing trend during the early period of record. Conversely, all stations exhibit decreasing trends in the latter portion of the record. All stations exhibited non-monotonic (piecewise) trends which changed from increasing to decreasing around 1970.

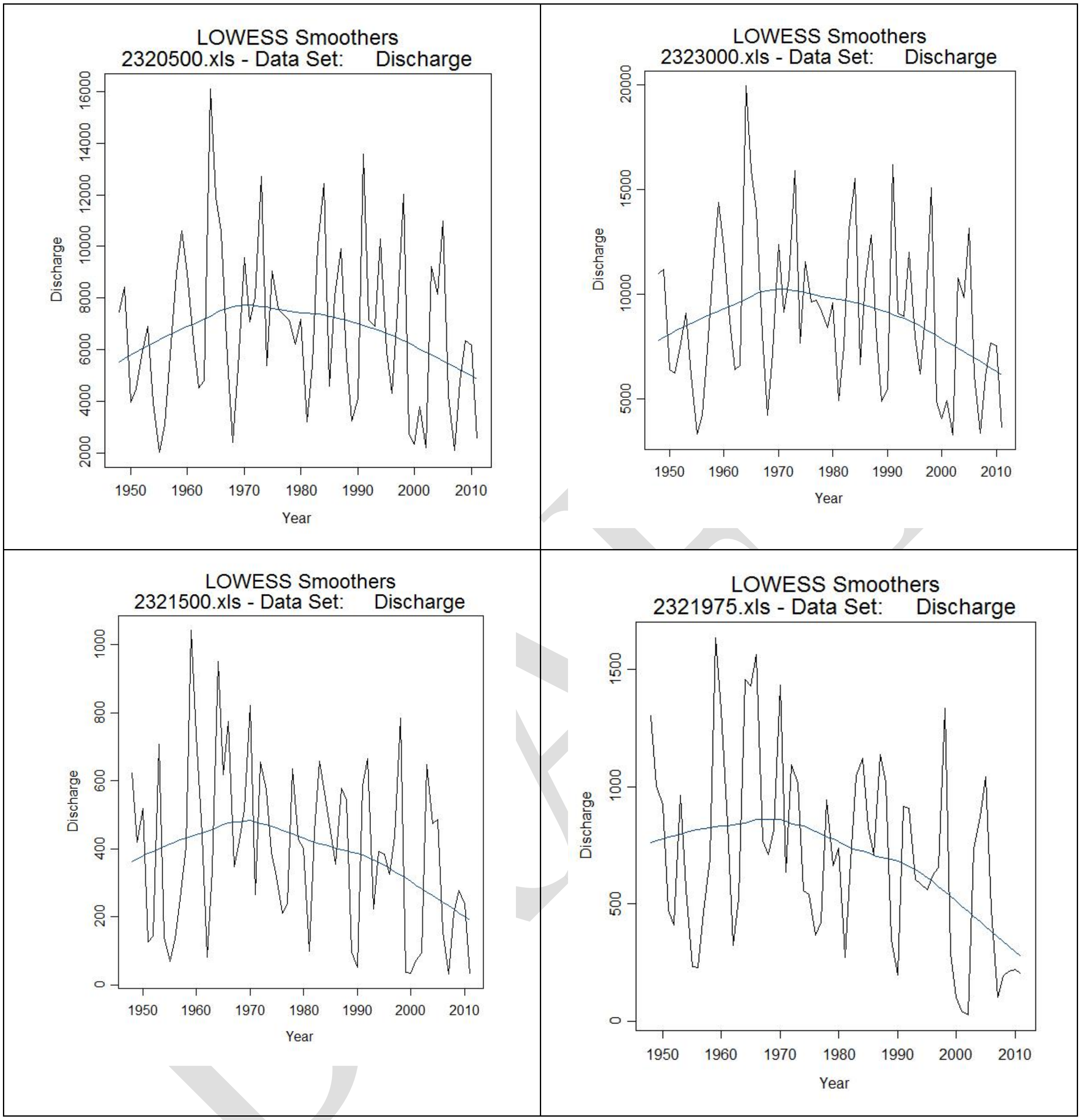



**Figure 47. LOWESS: Stations 2320500, 2323000, 2321500, and 2321975**

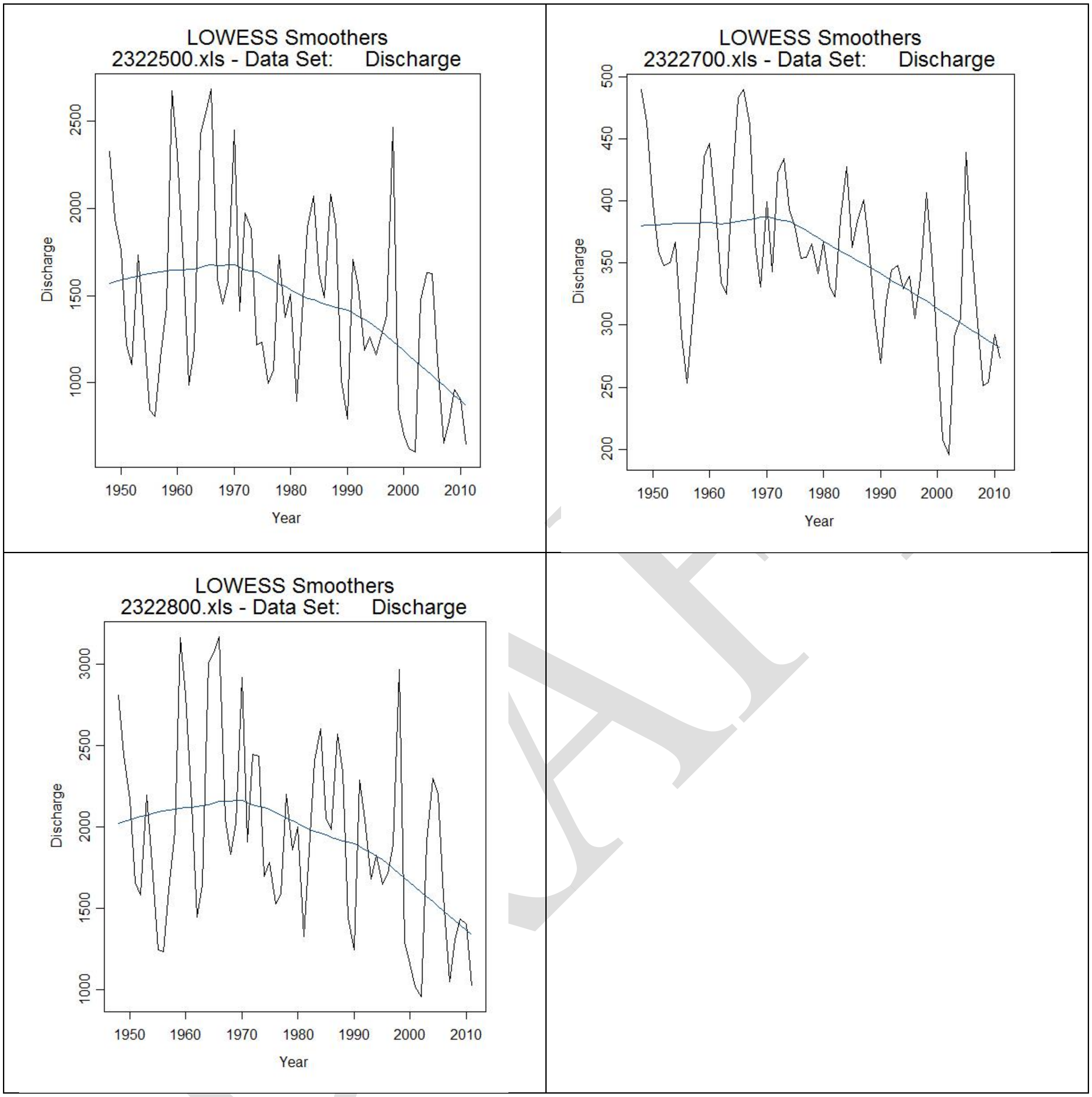



**Figure 48. LOWESS: Stations 2322500, 2322700, and 2322800**

The Mann Kendall (MK) test is extensively utilized for the examination of trends in hydrologic and hydro-climatic time series (Birsan et al., 2005; Kahya and Kalayci, 2004; Tao et al., 2011). Mann Kendall trend analysis was conducted on all stations in order to determine the significance of the trends over the common period of record (1948 through 2011). The aggregation of data to annual average minimized the serial correlation present in the time series. The presence of serial correlation (or dependency of the data at time t on time t-1), can lead to a false positive test (i.e. concluding that there is a trend when in fact the trend is due to serial correlation). The outcome of the MK test is the decision of whether or not to reject the null hypothesis, ho. Failure to reject the null hypothesis does not indicate that there is no trend in the data, but rather that there is not sufficient evidence to conclude that there is a trend (Helsel and Hirsch, 2002). The aggregation of data into an annual time series for this analysis is one technique which minimizes the presence of serial correlation. An additional technique for the minimization and/or elimination of serial correlation is the Mann Kendall bootstrapping technique, which involves the implementation of bootstrapping sampling to create N sample data sets, calculate N MK test statistics, and determine the bootstrap empirical distribution function of the MK test statistic. Yue and Pilon (2004) found that, although the MK bootstrapping procedure is more computationally intensive than the MK test, the statistical power of the test (the probability of not concluding there is no trend when there is a trend) was the same as the MK test. Additionally, as the magnitude of the trend increased, the statistical power of both the MK test and the bootstrap MK test increased; p-values obtained for the MK-test and bootstrapped MK test were very similar and generally did not change the conclusion of the hypothesis test. Therefore, since there was minimal advantage running the more computationally intensive test MK bootstrapped test, the MK test was utilized for all trend analysis.

In addition to trend analysis on total flow, trend analysis was also performed on baseflow. Baseflow was estimated using a low pass filter with a 120-day average minimum flow for each station (Perry, 1995). Results for the trend analysis are summarized in Table 30, and individual station trends are shown in Figures 49 through 55. As shown in Table 30, all stations exhibited statistically significant decreasing trends in baseflow, while a majority of stations exhibited statistically significant decreasing trends in total flow at the 80% significance level. In general, p-values were lower for baseflow than for total flow. The correlation coefficient, tau, is also shown. Tau, which ranges from -1 to 1, is a measure of the correlation between the data and time: a negative tau indicates that the data is decreasing as time increases, and a positive tau indicates that the data is increasing as time increases. For the MK test, the p-value corresponds to the probability of obtaining a tau value at least as extreme as the observed tau, assuming that the null hypothesis is true. The null hypothesis of the test is that there is no trend in the data. The null hypothesis is rejected when the p-value is less than the significance level, alpha. When the null hypothesis is rejected, it is concluded that the results of the test are statistically significant. Thus, given a 90% confidence level and the fact that the test is 2-sided (because trends can be either positive or negative), the null hypothesis can be rejected when the p-value is less than 0.05. Based on a 90% confidence level, all stations have statistically significant negative trends in baseflow, while 3 of the 7 stations have statistically significant trends in total flow.

**Table 30. Trend Analysis Results**

| Station Name | Total Flow | | | Baseflow | | |
|---|---|---|---|---|---|---|
| | Mann Kendall p-value | Mann Kendall Sen Slope cfs/year | Mann Kendall tau | Mann Kendall p-value | Mann Kendall Sen Slope cfs/year | Mann Kendall tau |
| Santa Fe Worthington | 0.073 | -2.892 | -0.154 | 0.005 | -0.470 | -0.240 |
| Santa Fe 441 | 0.006 | -7.900 | -0.237 | 0.000 | -5.658 | -0.334 |
| Santa Fe Ft. White | 0.002 | -11.503 | -0.263 | 0.000 | -8.511 | -0.371 |
| Ichetucknee Highway 27 | 0.000 | -1.799 | -0.352 | 0.000 | -1.977 | -0.397 |
| Santa Fe Hildreth | 0.004 | -11.525 | -0.246 | 0.000 | -10.415 | -0.360 |
| Suwannee Branford | 0.483 | -18.169 | -0.061 | 0.028 | -20.057 | -0.188 |
| Suwannee Bell | 0.196 | -32.862 | -0.111 | 0.008 | -33.276 | -0.227 |

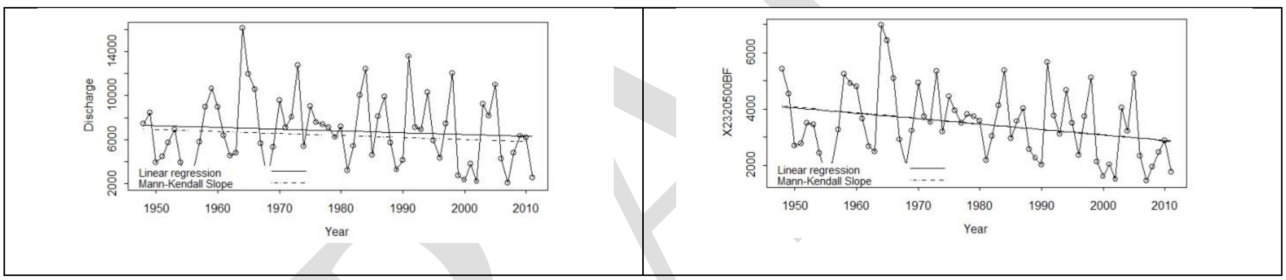



**Figure 49. Suwannee River at Branford Total Flow (left) and Baseflow (right)**

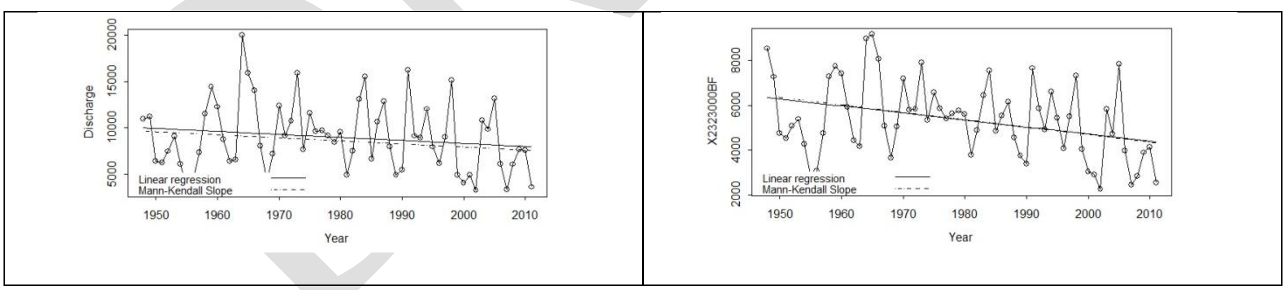



**Figure 50. Suwannee River at Bell Total Flow (left) and Baseflow (right)**

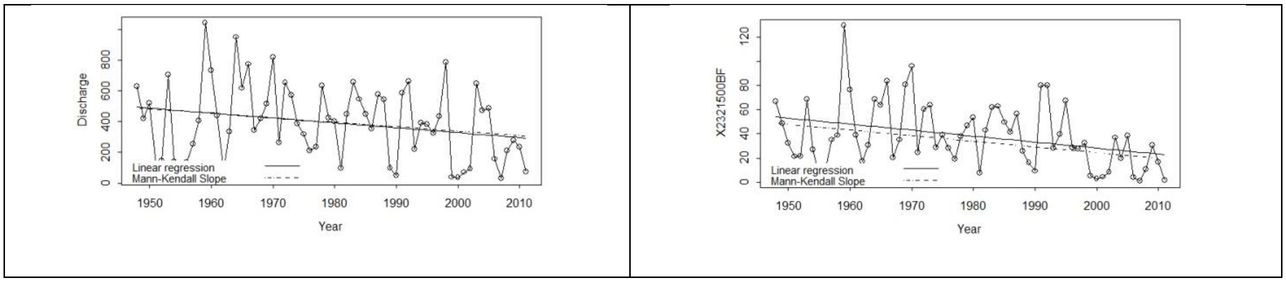



**Figure 51. Santa Fe at Worthington Springs Total Flow (left) and Baseflow (right)**

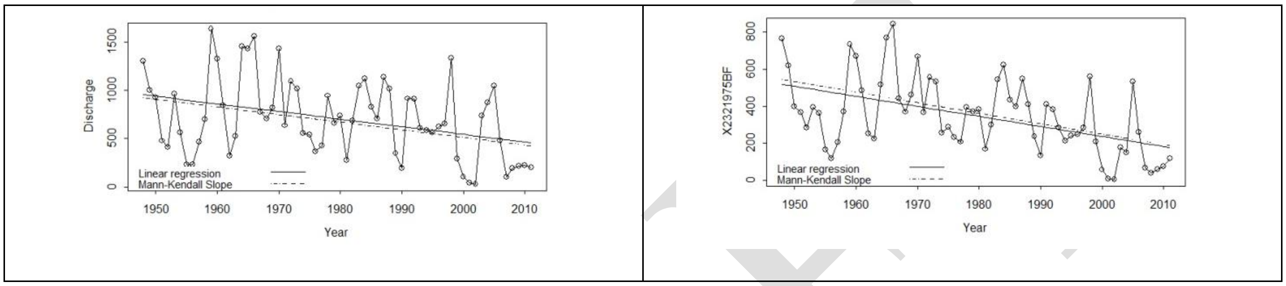



**Figure 52. Santa Fe at Hwy 441 Total Flow (left) and Baseflow (right)**

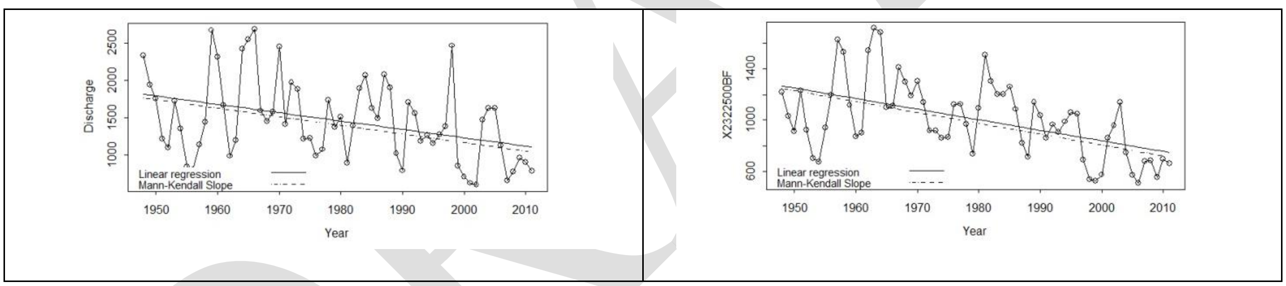



**Figure 53. Santa Fe near Ft. White Total Flow (left) and Baseflow (right)**

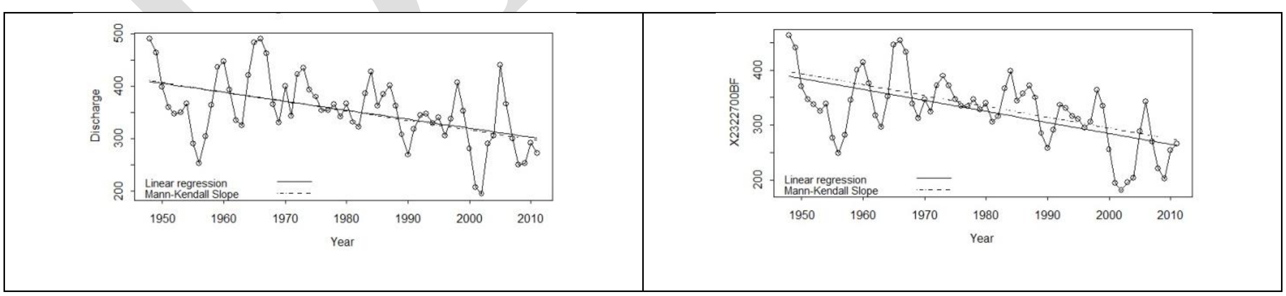



**Figure 54. Ichetucknee River HWY 27 Total Flow (left) and Baseflow (right)**

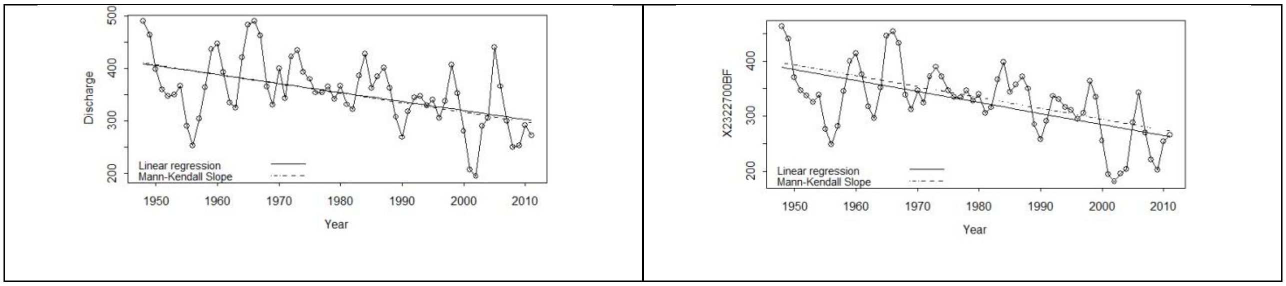



**Figure 55. Santa Fe near Hildreth Total Flow (left) and Baseflow (right)**

Although there were monotonic trends over the analysis period plotted in Figures 49 through 55 (1948-2011), based on the prior LOWESS results there appeared to be 2 well defined periods of monotonic trends in the majority of the time series, especially in the baseflow. In order to determine if these periods were periods of significant trends in the data, piecewise trend analysis was conducted on the total flow. Since the LOWESS results showed break points commonly around 1970, 1970 was used as the common break point for all piecewise trend analysis. Piecewise trend analysis divides the period of record into 2 time segments: the first from common beginning point of the record (1948) to the break point, and the second from the break point to the end of the record, as shown graphically for the example in Figure 56. As shown, the Santa Fe River at 441 station exhibits an increase in flows of 16.1 cfs/year in period 1, and a decrease in flow of 11.8 cfs per year in period 2. Results of the trend analysis for each station and segment are shown in Table 31. As shown, all stations experience decreases in slope from the first period to the second. The p-value also increases in significance from period 1 to period 2 for many stations. It should be noted that all period 2 trends are negative, indicating that flows are decreasing annually, while some period 1 trends (Suwannee River at Branford, Santa Fe at Worthington Springs, Santa Fe at 441, and Santa Fe at Ft White) are increasing.

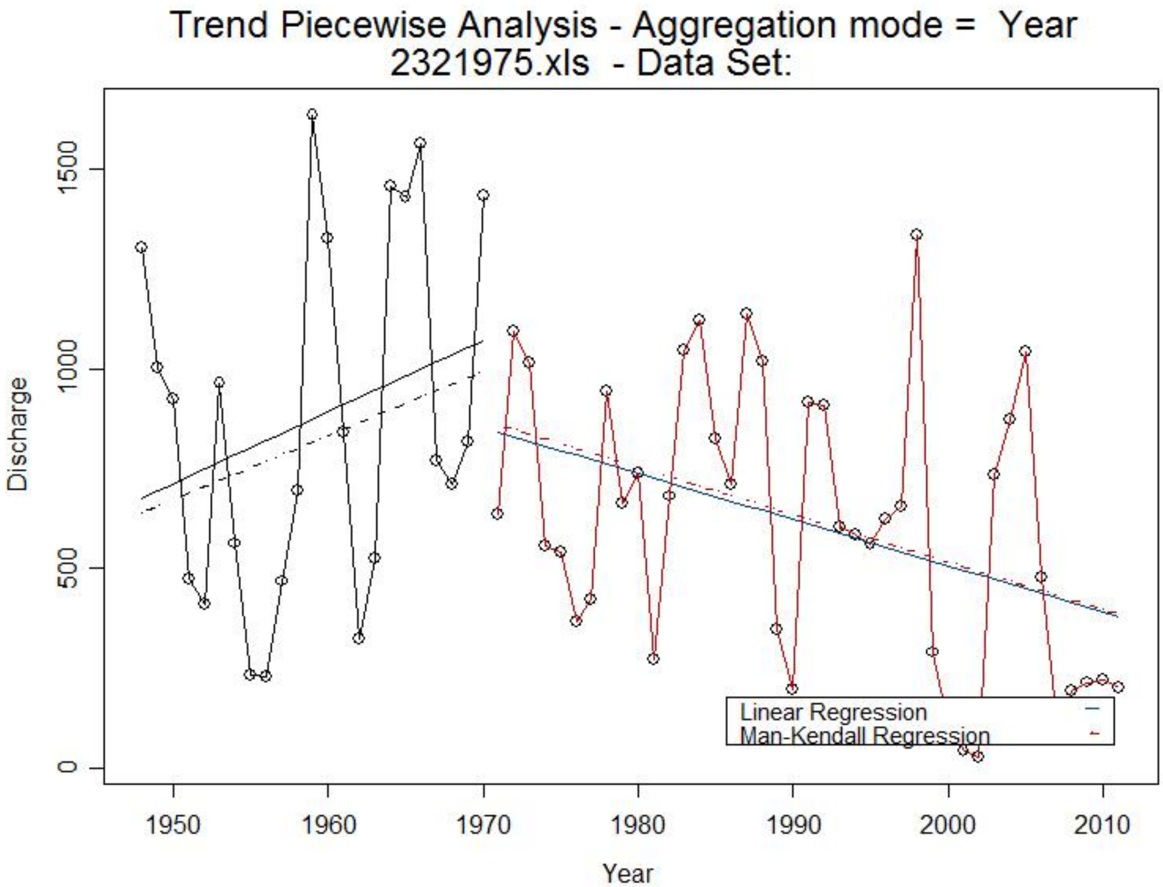



**Figure 56. Piecewise Trend Analysis Example: Santa Fe River at 441 Total Flow**

**Table 31. Non-Monotonic Trend Analysis**

| Station Name | Period 1 (1948-1969) | | | Period 2 (1970-2011) | | |
|---|---|---|---|---|---|---|
| | Sen Slope, cfs/year | MK tau | MK p-value | Sen Slope, cfs/year | MK tau | MK p-value |
| Santa Fe Worthington | 15.201 | 0.217 | 0.154 | -5.703 | -0.215 | 0.049 |
| Santa Fe 441 | 16.078 | 0.154 | 0.316 | -11.844 | -0.266 | 0.015 |
| Santa Fe Ft. White | 20.275 | 0.138 | 0.369 | -16.391 | -0.293 | 0.007 |
| Ichetucknee Highway 27 | 0.993 | 0.051 | 0.751 | -2.777 | -0.429 | 0.000 |
| Santa Fe Hildreth | 20.026 | 0.130 | 0.398 | -17.873 | -0.298 | 0.006 |
| Suwannee Branford | 92.469 | 0.146 | 0.342 | -83.239 | -0.222 | 0.042 |
| Suwannee Bell | 95.837 | 0.154 | 0.316 | -115.311 | -0.285 | 0.009 |

## Rainfall Analysis

Since rainfall is typically a primary explanatory variable that affects the hydrologic response of a region, changes to the relationship between rainfall and streamflow can indicate changes in the relationship between streamflow and other explanatory variables such as water use, impoundments, and land use changes. Of all the explanatory variables which contribute to streamflow, climatic variables such as rainfall offer the most complete historical data set. Historical anthropogenic changes are difficult to obtain with sufficient resolution to examine the influences of a single variable on streamflow over time.

An in-depth analysis of nearby rainfall was conducted in order to fully characterize a major variable for changes in flow. Other potential explanatory variables include pumping, land use changes, and other anthropogenic changes. Of these factors, rainfall is the most complete data set, and therefore, an in-depth examination of rainfall can yield important conclusions regarding the causal relationship between rainfall and streamflow. Two long-term rain gauges, located in Lake City and Gainesville, were available for the analysis (Figure 57). Analysis included moving window average rainfall in order to examine long-term changes in total rainfall at each of the gauges, and trend analysis on the moving window average time series.

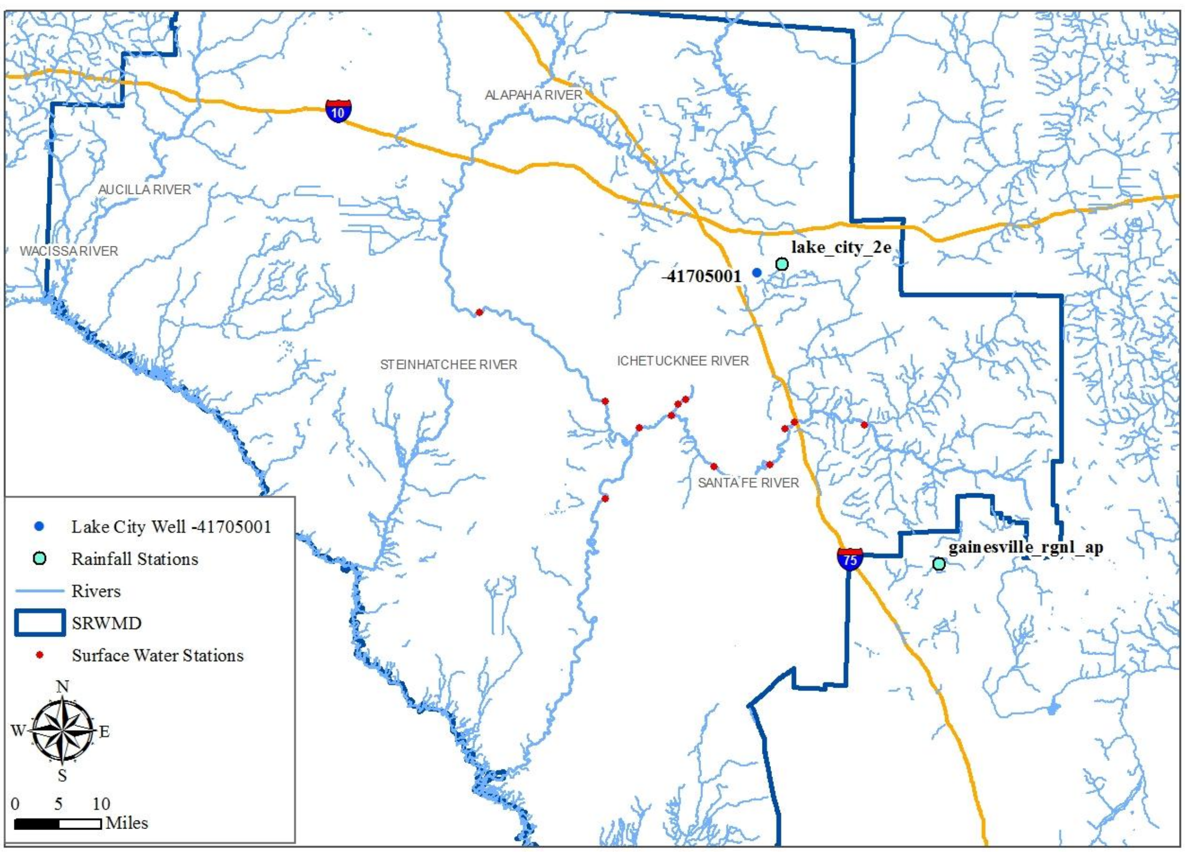



**Figure 57. Rain Gauge Locations**

## Lake City Rainfall

The Lake City daily rainfall was aggregated into monthly totals. For each month, the total rainfall for the current month plus the 11 months prior was calculated, resulting in an annual total rainfall which was assigned to that month. These totals were averaged based on the desired window length in order to look for long term trends in the rainfall data. Varying the window length is an exploratory technique that makes it possible to detect long term changes in rainfall trends, which was the objective of this analysis. The resulting rainfall time series are shown in Figure 58. As shown in the figure, as the window length increases, the long term trends in the data become more evident due to the averaging of the high frequency noise in the data. Both the 5-year and 10-year averages exhibit clear changes in trend (from increasing to decreasing). Locally weighted scatterplot smoothing (LOWESS) was run on the data as an exploratory technique to determine if non-monotonic trends were present in the data.

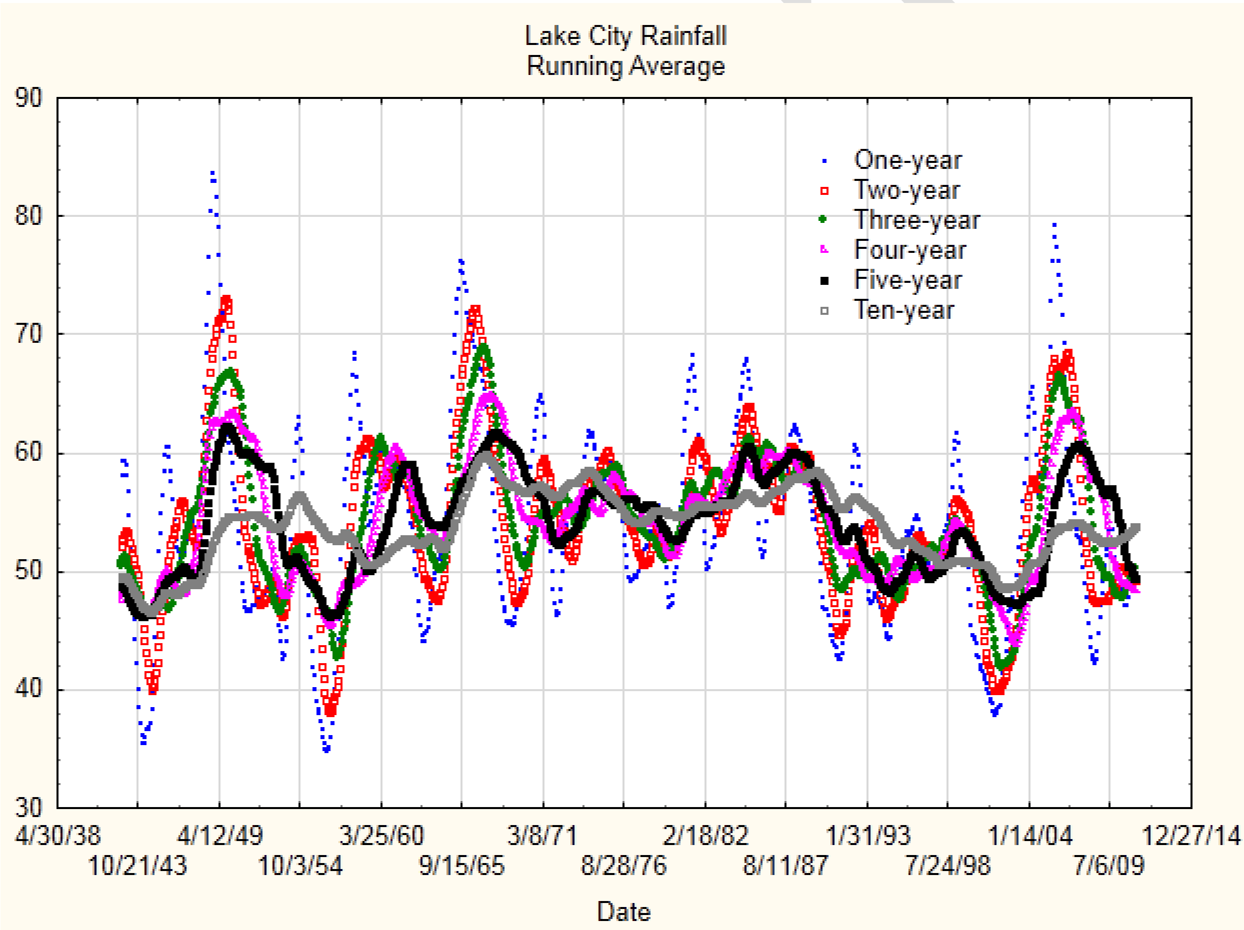



**Figure 58. Lake City Rainfall, Moving Averages**

As shown in Figure 59, a clear change in the trend of the data is visible in both the 5-year and 10-year smoothed data. This trend is less pronounced, yet still visible, in the 1-year and 3-year rainfall data. The 10-year average data was further analyzed in order to determine if there was a statistically significant monotonic trend over the period of record. The results are shown in Figure 60 and Table 32. As shown in the table, based on a p-value of 0.58 and an assumed confidence level of 90%, a statistically significant monotonic trend is not present in the data. Based on the results of the LOWESS analysis, this was expected. The time series was further divided into two segments in order to determine if there were periods in the time series with statistically significant trends. The one-year aggregated data set was used. Trend analysis was completed on each segment, as shown in Figure 61 and Table 33. As shown, when the time series is divided into segments, the trends in each segment are statistically significant at the 90% confidence level, and show a change in trend from increasing to decreasing around the late 1970s.

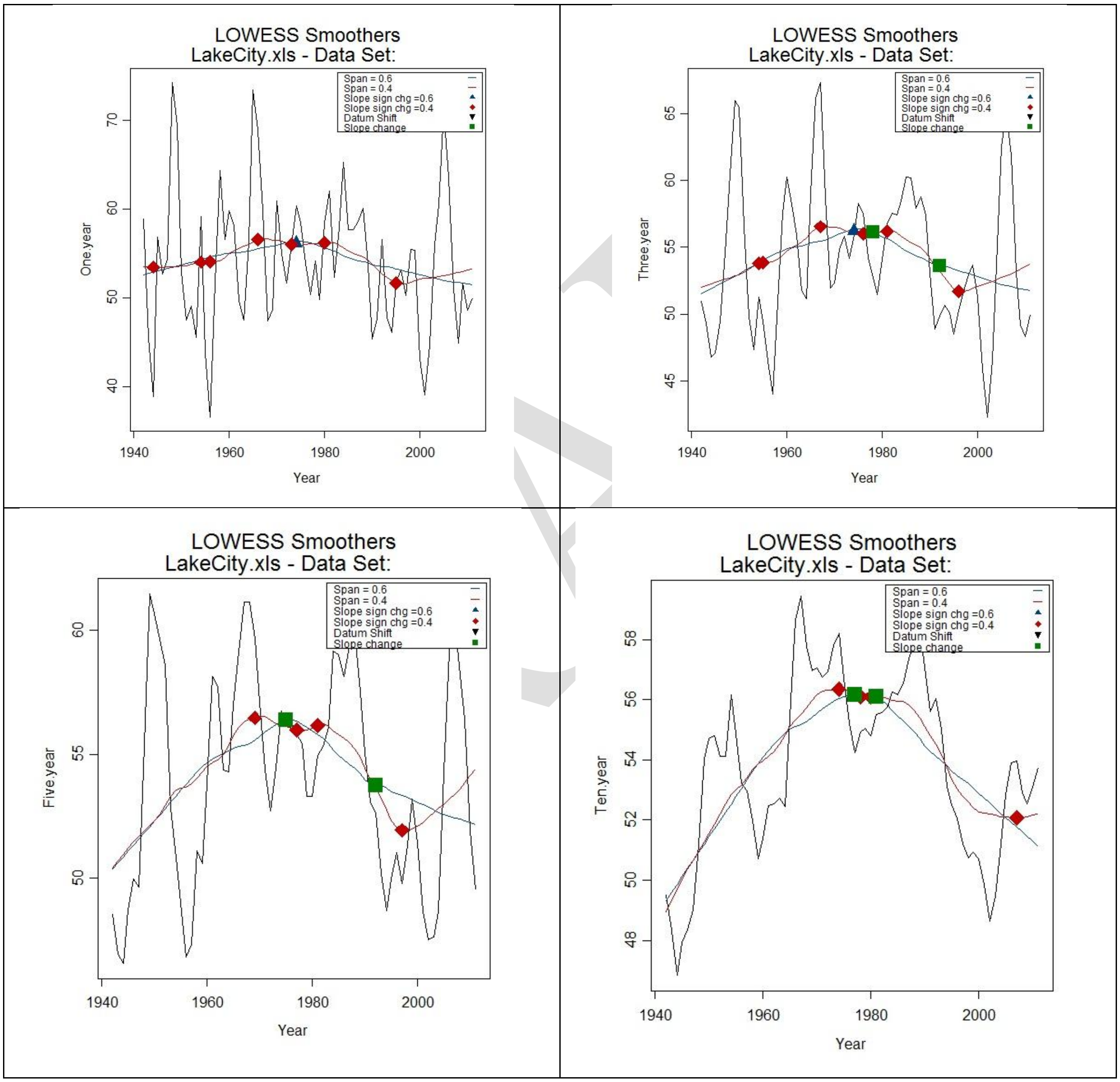



**Figure 59. LOWESS, 1-year, 3-year, 5-year and 10-year**

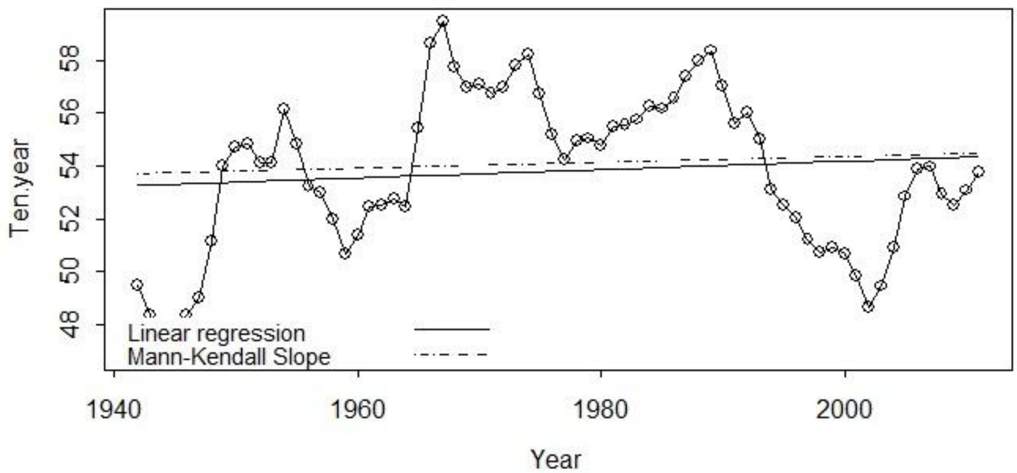



**Figure 60. Ten-year Average Monotonic Trend Analysis**

**Table 32. Ten-year Monotonic Trend Analysis**

| Mann Kendall p-value | Mann Kendall Sen Slope, inches/year | Mann Kendall tau |
|---|---|---|
| 0.5840 | 0.0111 | 0.0451 |

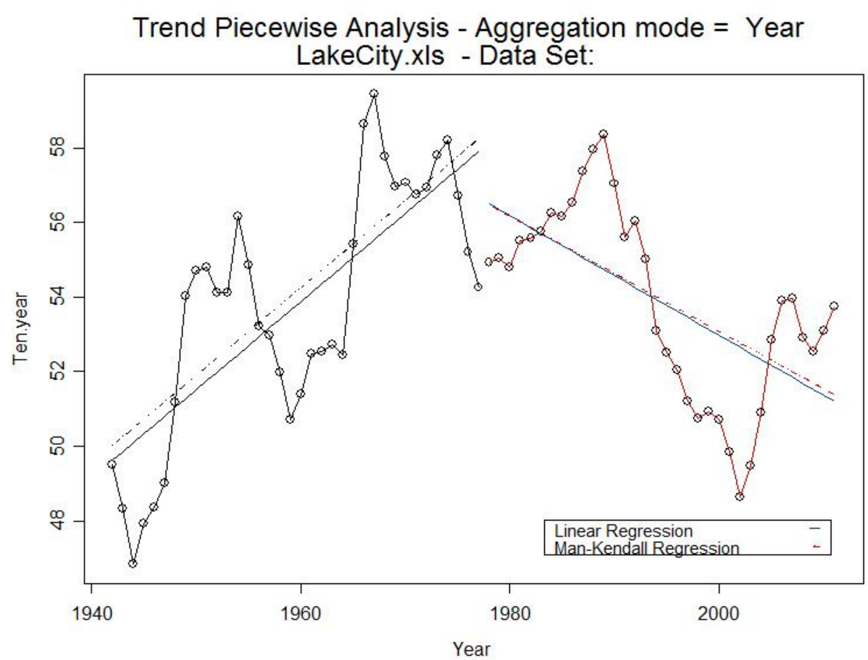



**Figure 61. Segmented Trend Analysis, Lake City Rainfall**

**Table 33. Segmented Trend Analysis, Lake City Rainfall**

| | Segment 1 (Prior to 1/1/1977) | Segment 2 (After 1/1/1997) |
|---|---|---|
| Mann Kendall p-value | 1.159e-07 | 7.622e-05 |
| Mann Kendall Sen slope (inches/year) | 0.23487 | -0.15437 |
| Mann Kendall tau | 0.50159 | -0.36899 |

## Streamflow and Baseflow Monthly Analysis

Based on the results of the exploratory data analysis and trend analysis, there are two distinct periods of monotonic trends in the flow data. From approximately the mid-1960s through 1970, the majority of the stations are within a transition period, changing from increasing to decreasing flows. In general, flows are increasing or have little to no trend from the beginning of the record until approximately 1970. Flows are generally decreasing from 1970 through the end of the record. The data was divided into 2 sets for analysis: pre-1970 data and post-1970 data. Monthly average flows for these 2 periods were compared.

Trends in streamflow and baseflow have been studied extensively by others (Birsan et al., 2005; Kahya and Kalayci, 2004; Lins, 2005; Tao et al., 2011; Zhang and Schilling, 2006). Streamflow and baseflow can vary significantly during a year due to seasonal weather patterns (Zhang and Schilling, 2006). The seasonal variations in streamflow and baseflow are likely to change due climate, changes in precipitation, land use changes and anthropogenic impacts.

The seasonal (monthly) streamflow averages at the surface water stations of the Santa Fe River Basin were examined in order to determine if there were changes in monthly average streamflow over the period of record. For each surface water station, average streamflow and baseflow of each month in 1970 through 2011 were calculated. Average streamflow and baseflow of each month in 1970 through 2011 were compared to average streamflow and baseflow of each month from the beginning of the record through 1970. Cumulative average streamflow and baseflow of each month before and after 1970 were also plotted. The results for streamflow and baseflow are shown in Figures 62 through 68.

Generally, the plots of monthly mean discharge show that there has been a change in seasonality or in how streamflow is distributed throughout a year. Generally for the Suwannee River stations and the Ichetucknee River stations, total cumulative streamflow in the early period is very similar to the late period. In comparison to the earlier time series, there is a decrease in monthly streamflow during later months of a year (August though December). The Santa Fe River stations exhibit a decline in total cumulative streamflow in the post-1970 period. These changes can be attributed to climate differences and human activity.

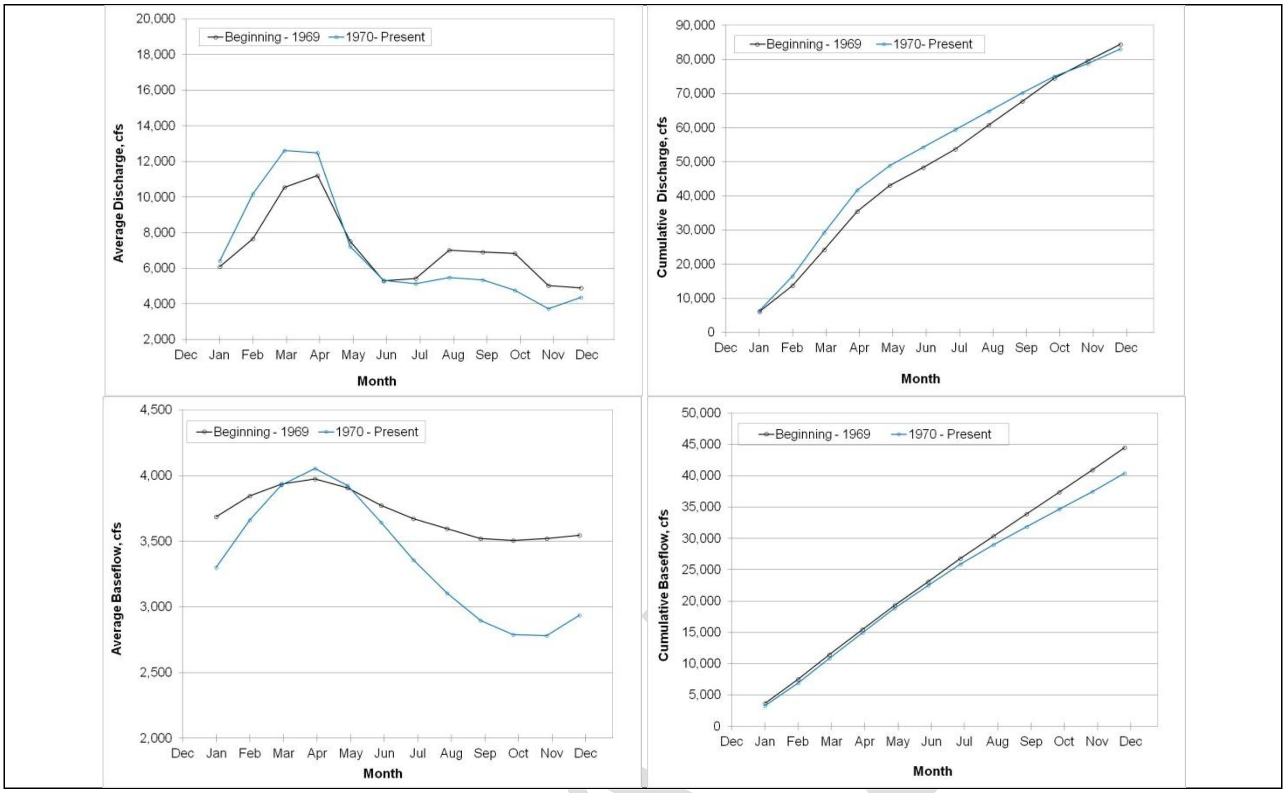



**Figure 62. Suwannee River at Branford (2320500): Monthly Statistical Analysis**

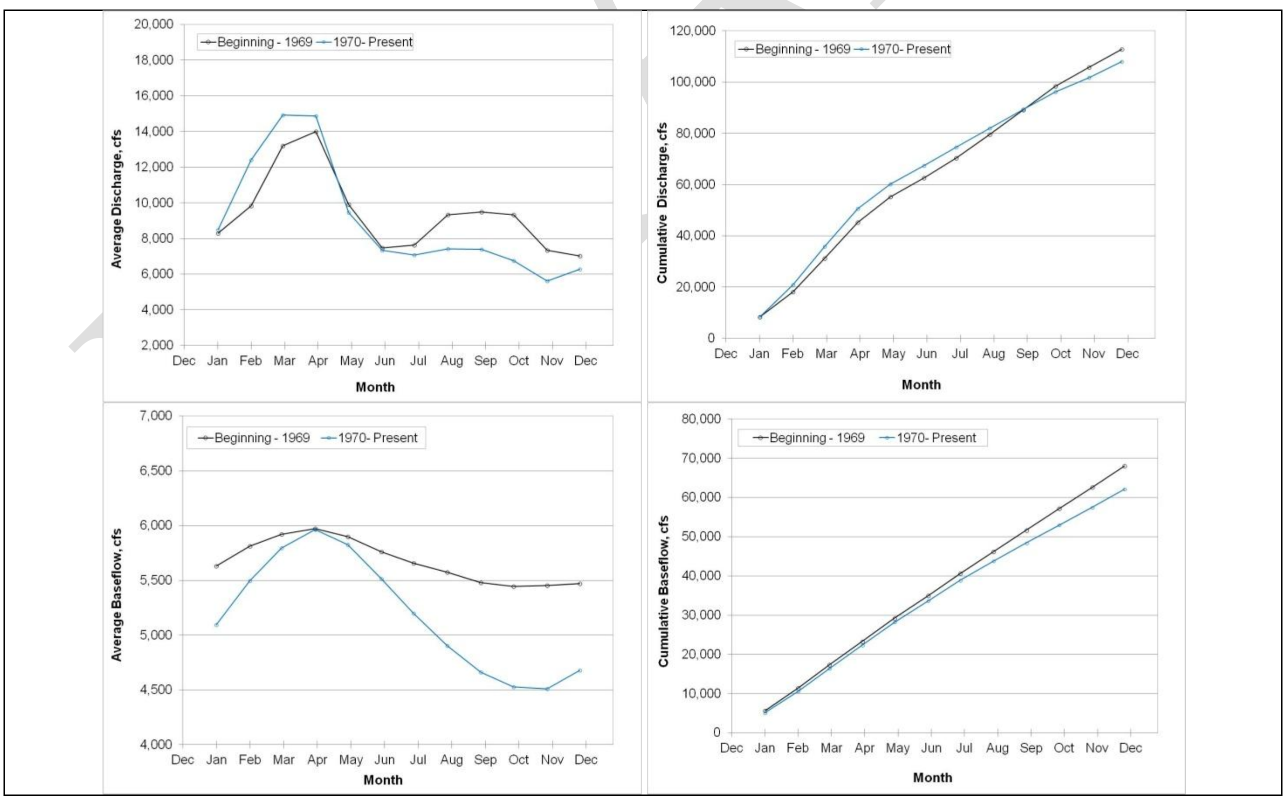



**Figure 63. Suwannee River near Bell (2323000): Monthly Statistical Analysis**

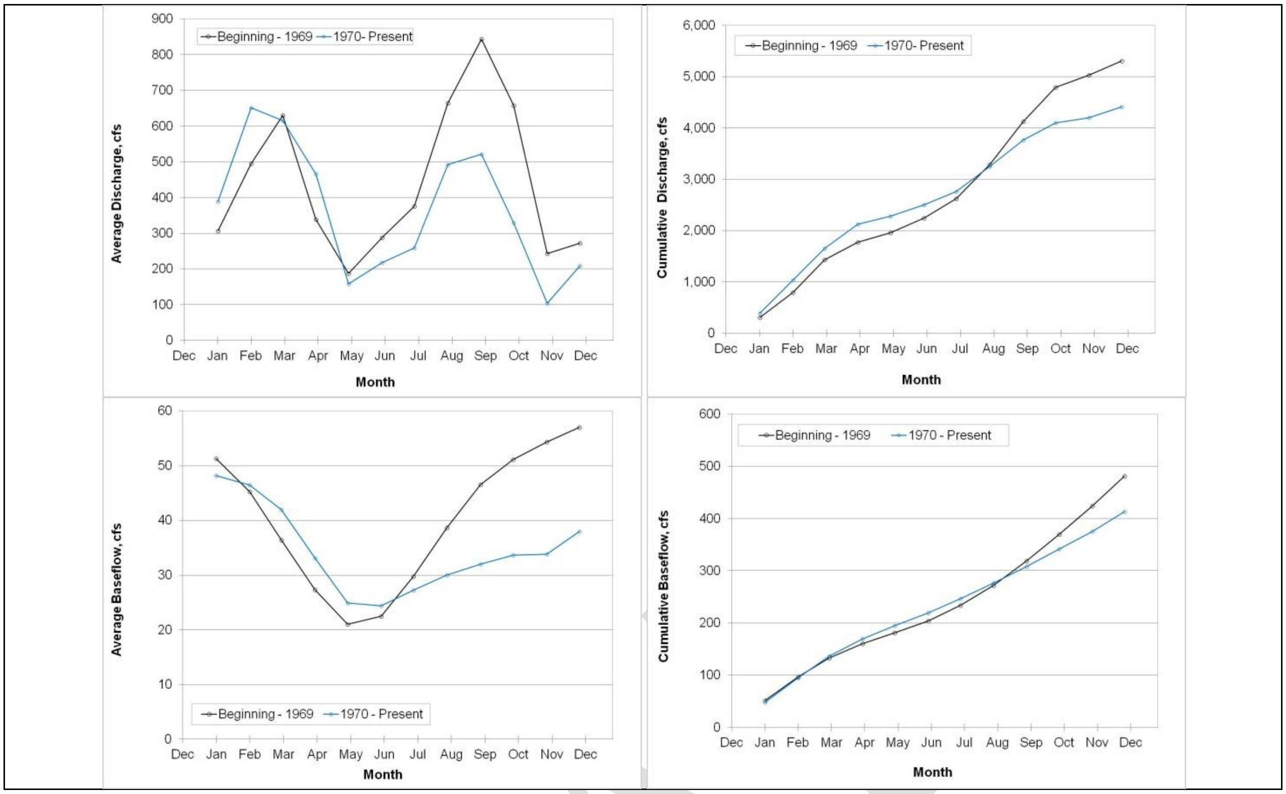



**Figure 64. Santa Fe at Worthington Springs (2321500): Monthly Statistical Analysis**

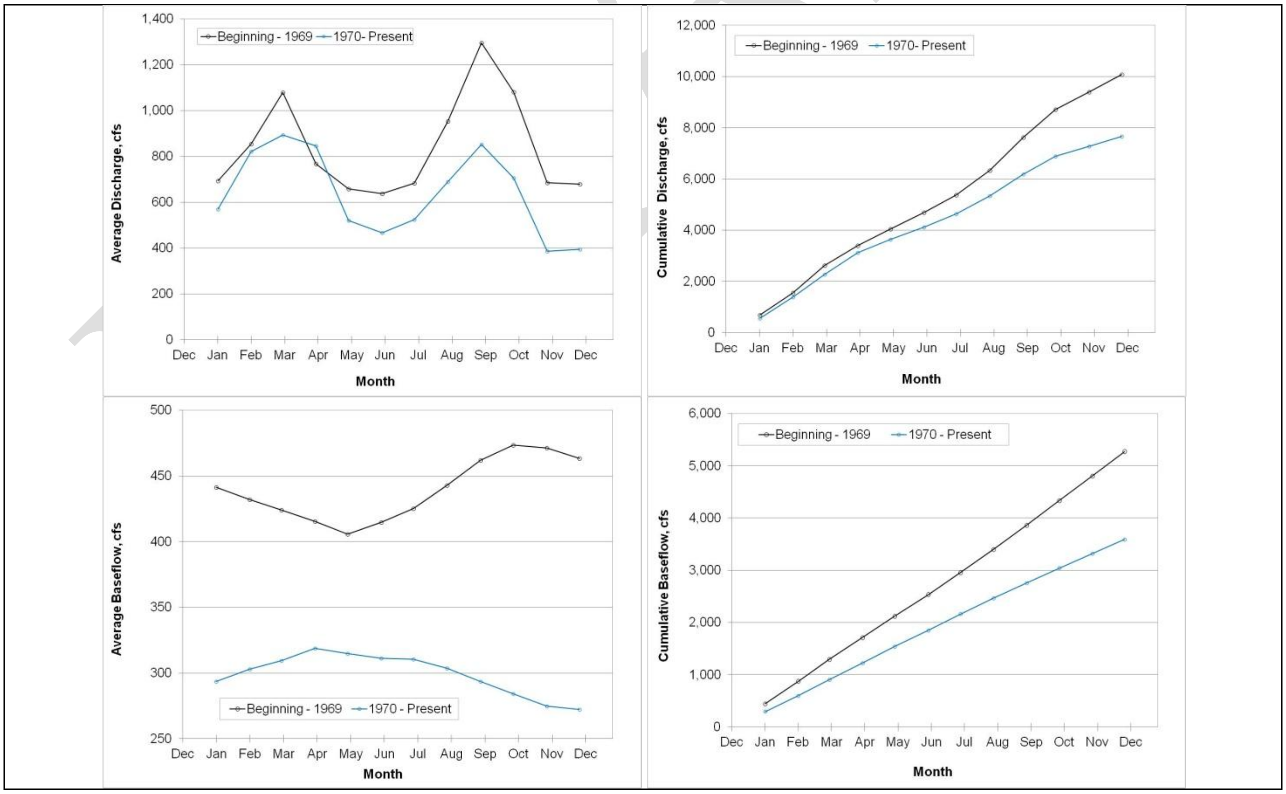



**Figure 65. Santa Fe at Highway 441 (2321975): Monthly Statistical Analysis**

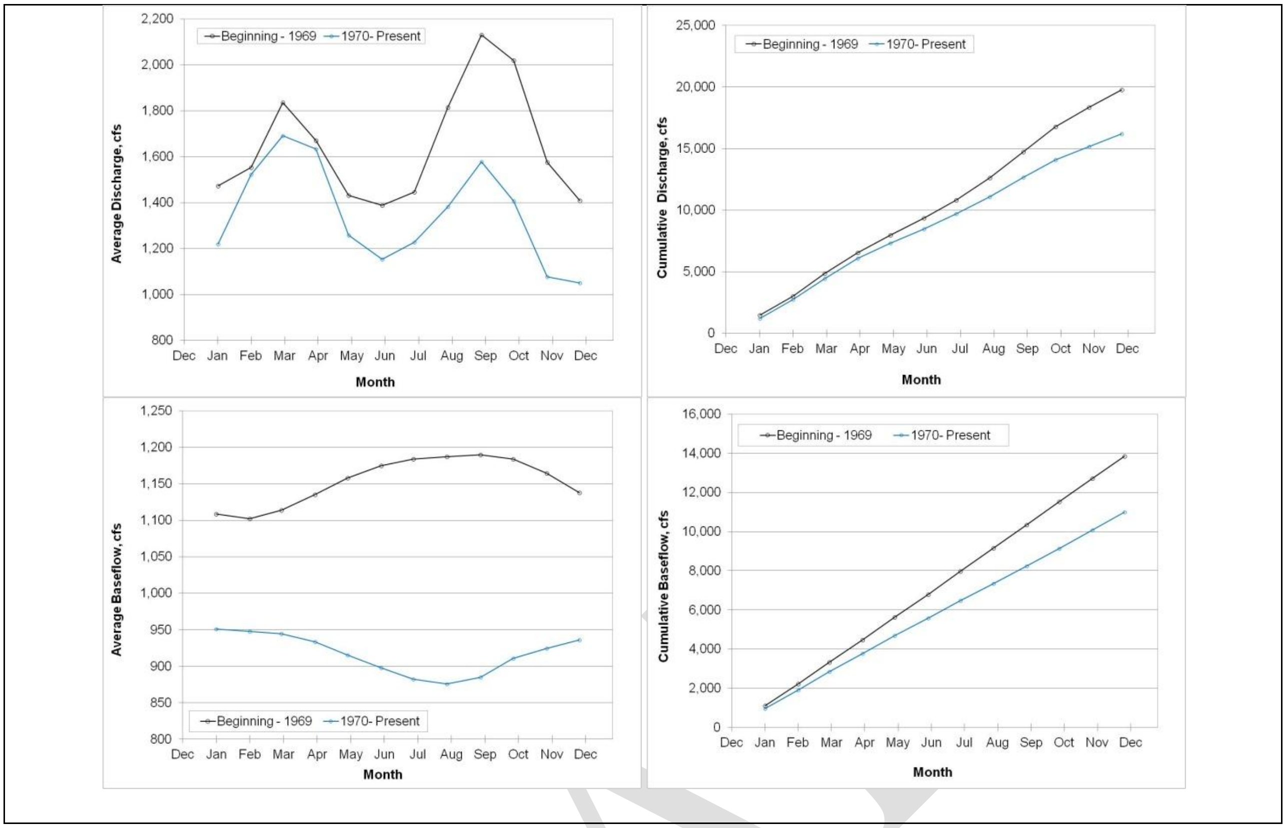



**Figure 66. Santa Fe near Ft. White (2322500): Monthly Statistical Analysis**

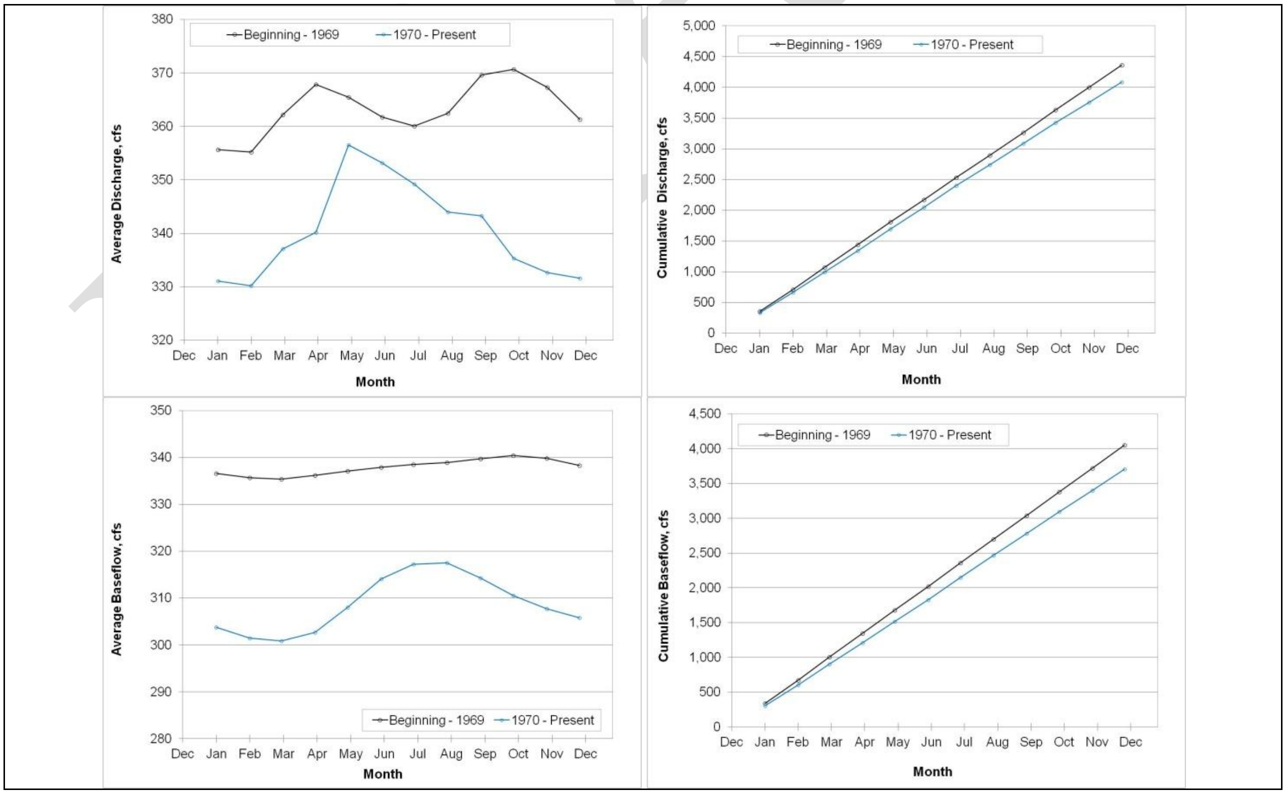



**Figure 67. Ichetucknee at Highway 27 (2322700): Monthly Statistical Analysis**

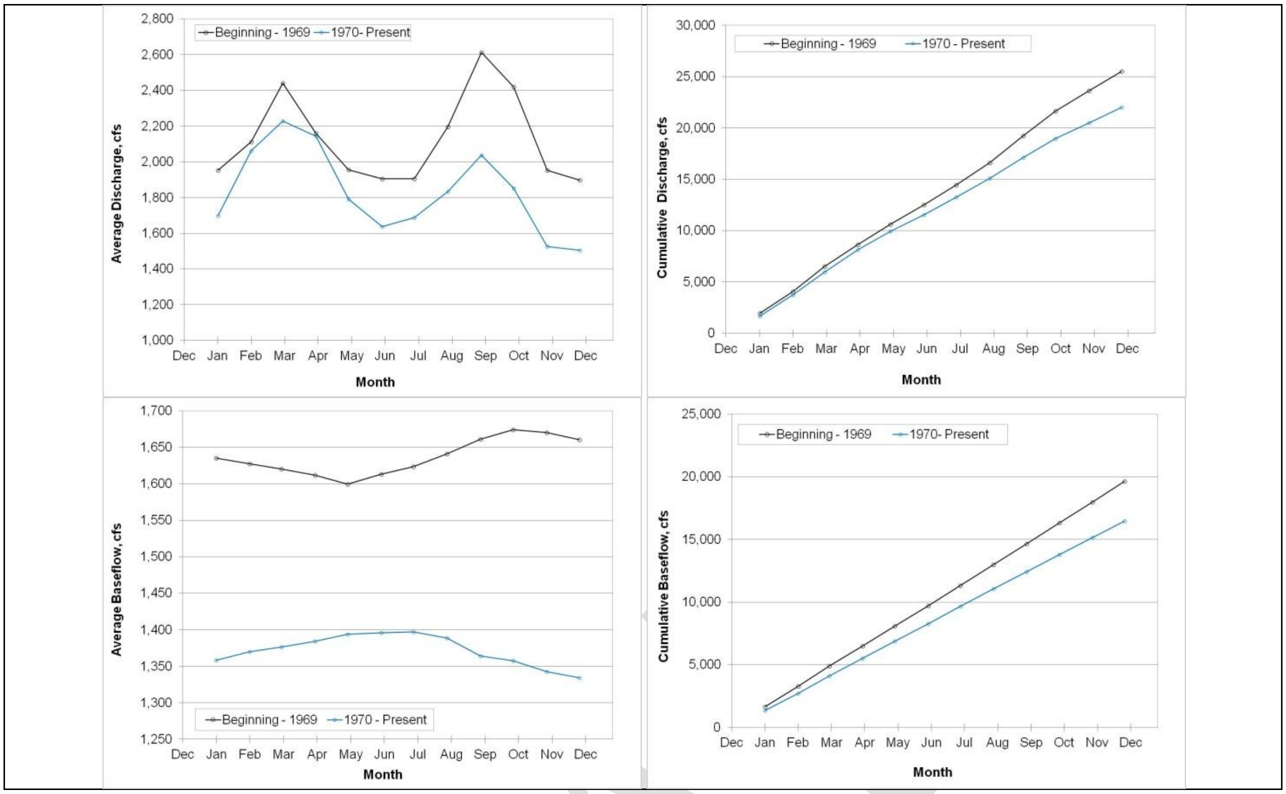



**Figure 68. Santa Fe near Hildreth (2322800): Monthly Statistical Analysis**

## Adjusted Historical Flows

A methodology for the adjustment of flows for groundwater influences was developed and applied to baseflow time series for selected long-term stations on the Santa Fe and Ichetucknee Rivers. For the purposes of this analysis, the term adjusted historical flows refers to the flows in a river that have be adjusted to account for anthropogenic activities, namely, groundwater pumping. Flows in the Lower Santa Fe River Basin and Ichetucknee River Basin are baseflow-dominated due to the large contribution spring flow to total flow, the local karst formations and the close proximity to the Floridan Aquifer (Schneider et. al, 2008).

This section of the report describes background, assumptions, and methods applied to the development of the adjusted historical flow time series for the selected Santa Fe and Ichetucknee stations. Adjusted historical flow time series were developed for the Santa Fe River at Highway 441 near High Springs (2321975), Santa Fe River near Fort White (2322500) and Ichetucknee River at Highway 27 (2322700).

### Background and Literature Review

A literature review was conducted to identify methods and procedures previously used in the development of adjusted historical flows. Unimpacted and unimpaired flows are used in literature interchangeably. Unimpaired flows are the flows that would have occurred in the main stem of a river if the flows had not been changed by water withdrawals, discharges, and operation or installation of dams – human influences that can be numerically quantified (ARCADIS, 2010). Tarboton (1994) defines unimpaired streamflow as "measured streamflow adjusted for anthropogenic consumptive use and reservoir operations." Similar to unimpaired flows, unimpacted flows are the flows that would have occurred in a river if the flows had not been impacted by any anthropogenic activities, both numerically quantifiable and non-quantifiable, such as water withdrawals, discharges, and changes to land use.

*Streamflow Trends*

Streamflow variability in a river has been described in the past by many various process models. However, because of the site-specific nature of these process models, physical understanding of flow variability on a large scale is still limited (Milly and Wetherald, 2002).

Low-flow characteristics are commonly used to establish minimum flows. Quantitative estimates of low-flow characteristics were performed on 216 continuous-record and 1,100 partial-record USGS gauging stations through water year 1987. Low-flow daily-mean discharge records were utilized for an analysis of low-flow characteristics of Florida streams including the following stations: Suwannee River at Branford (2320500), Suwannee River near Bell (2323000), Santa Fe River at Worthington Springs (2321500), Santa Fe River at O'Leno State Park (2321898), Santa Fe River near Fort White (2322500), and Santa Fe River near Hildreth (2322800). The data was examined for trends in the daily-mean discharge records, and the largest concentration of trends were found at the Santa Fe River Basin stations. In addition, a significant upward trend in annual low flows at the Worthington Springs gauging station was reported (Rumenik and Grubbs, 1996).

The USGS studied trends in streamflow using various methods and concluded that streamflow had been generally increasing in the United States since 1940 (Lins, 2005). The USGS found that between 1940 and 1999 most of the increases occurred in low to moderate stream flows (Lins, 2005). Decreases in the annual minimum and annual median stream flows were found in only 8 percent and less than 1 percent of the Hydro-Climatic Data Network (HCDN) stations respectively (Lins, 2005). The HCDN is the national data set of streamflow records "that are relatively free of confounding anthropogenic influences" (Slack and Landwehr, 1992). The daily mean streamflow records are available at Worthington Springs, Fort White, and Branford surface water stations for select water years in the HCDN period of record, which spans 1874 through 1988 (Slack and Landwehr, 1992). The USGS also concluded that increasing trends were prevalent in the Upper Mississippi, Ohio Valley, Texas-Gulf, and the Mid-Atlantic; whereas, the decreasing streamflow trends were observed in the Pacific Northwest and the South Atlantic-Gulf (Lins, 2005).

*Assessment of land use and climate impacts on streamflow and analysis of patterns in rainfall-streamflow relationships*

Attempts have been made to assess the impacts of climatic differences, and land use and other anthropogenic changes on streamflow. Climate impacts are difficult to discern or isolate from impacts due to land use change and other human activity (Tomer and Schilling, 2009). Li et al. (2007) attempted to assess impacts of climate and soil conservation measures (land use impacts) on streamflow of the Wuding River Basin in China. Li et al. (2007) estimated that soil conservation measures were responsible for 87% reduction in mean annual streamflow, whereas climate impacts were responsible for 13% reduction. The assessment method was based on the relationship between annual streamflow and precipitation and produced consistent results (Li et al., 2007). However, the methodology to assess the land use changes on streamflow did not produce consistent results.

Peterson et al. (2011) postulate that in order to assess the impacts of land use change on streamflow, a clear understanding of the past and present correlations between streamflow and stationary landscape characteristics are to be established using a regionalized approach. Multivariate statistical analysis and principal component analysis are commonly used to perform regionalization. Johnston and Shmagin (2008) used factor analysis, a multivariate statistical technique, to identify the regions in the Great Lakes Basin with common and unique discharge patterns.

Milly and Dunne (2002) attempted to estimate changes in streamflow due to changes in precipitation and potential evaporation. They developed a semi-empirical relation that could be used to derive estimates of the sensitivity of annual runoff and evaporation to annual precipitation. Rossi et al. (2009) studied the response of the Mississippi River to climate fluctuations and suggested that precipitation data was the most representative hydrological signal of climate fluctuations. Overall, the impacts of human activities on streamflow have been studied extensively (Christian et al., 2011; Hao et al., 2009; Zhang et al., 2006).

*Methods to Establish Causal Relationships*

Many statistical and numerical methods have been utilized by others in order to establish causal relationships between basin flow and explanatory variables such as water use and precipitation. Wang and Cai (2009) utilized base flow recession analysis to demonstrate that direct human interferences within a basin, such as groundwater pumping and discharge of effluent, affect the low-flow hydrograph, while indirect human interferences such as land use change, alter the rising and falling limbs of the hydrograph and the peak flow. Gao et. al (2010) utilized trend analysis and double mass analysis between stream flow and precipitation to detect changes in the Yellow River flow. Downward trends were largely attributed to human intervention. Mongan and Miller (1990) found that there was an increasing trend in freshwater inflow to San Francisco Bay because precipitation increased faster than water use. Zhang and Schilling (2006) used trend analysis to demonstrate that an increase in Mississippi River streamflow since the 1940s was due to increases in baseflow caused by land use changes in the Mississippi River Basin over the last 60 years, primarily the expansion of soybean cultivation. Ye et. al (2003) used statistical analysis to document changes to the Lena River hydrology induced by human activities (reservoirs) and natural variations.

As an alternative to the development of a hydrologic model that utilizes many explanatory variables to simulate streamflow, statistical models of flow can be constructed using the best available explanatory variable, rainfall. If the developed statistical models demonstrate a change in the relationship between flow and rainfall during a pre-impacted period and a post-impacted period, then it can be shown that the relative contribution of rainfall to flow has changed over time. If the contribution has changed, then the relative contribution of other explanatory variables (water use and other anthropogenic factors) has also changed. The change in this relationship over time is the amount of influence that is due to non-rainfall factors.

## Methodology

Grubbs (2011) demonstrated that daily flow at the Ichetucknee Highway 27 station can be estimated using a regressive model with 24-month accumulated rainfall as the independent variable. A comparison of the LOWESS for the average 24-month rainfall time series ('Average24') and the Santa Fe River at Ft. White baseflow are shown in Figure 69. For this analysis, the rainfall datasets from the Lake City and Gainesville gauges were averaged in order to create a rainfall time series with average rainfall over the domain. The use of both gauges represents the best available long-term rainfall data for the analysis and characterization of flow. As shown, both time series exhibit several common characteristics: seasonal variability, an increasing trend until approximately 1970, and a decreasing trend from approximately 1970 through the end of the record. The stark difference between the two LOWESS analyses is the difference in the slope of the LOWESS line from 1970 through the end of the record. While the pre-1970 LOWESS data is comparable in steepness of slope, it is clear that the slope of the baseflow LOWESS post-1970 is steeper than the rainfall slope. Rainfall and baseflow are related, but it appears that the nature of the relationship changes over time. Using the exploratory data analysis as a basis for the flow adjustments, an adjusted time series was developed in order to maintain the early period relationship between rainfall and baseflow.

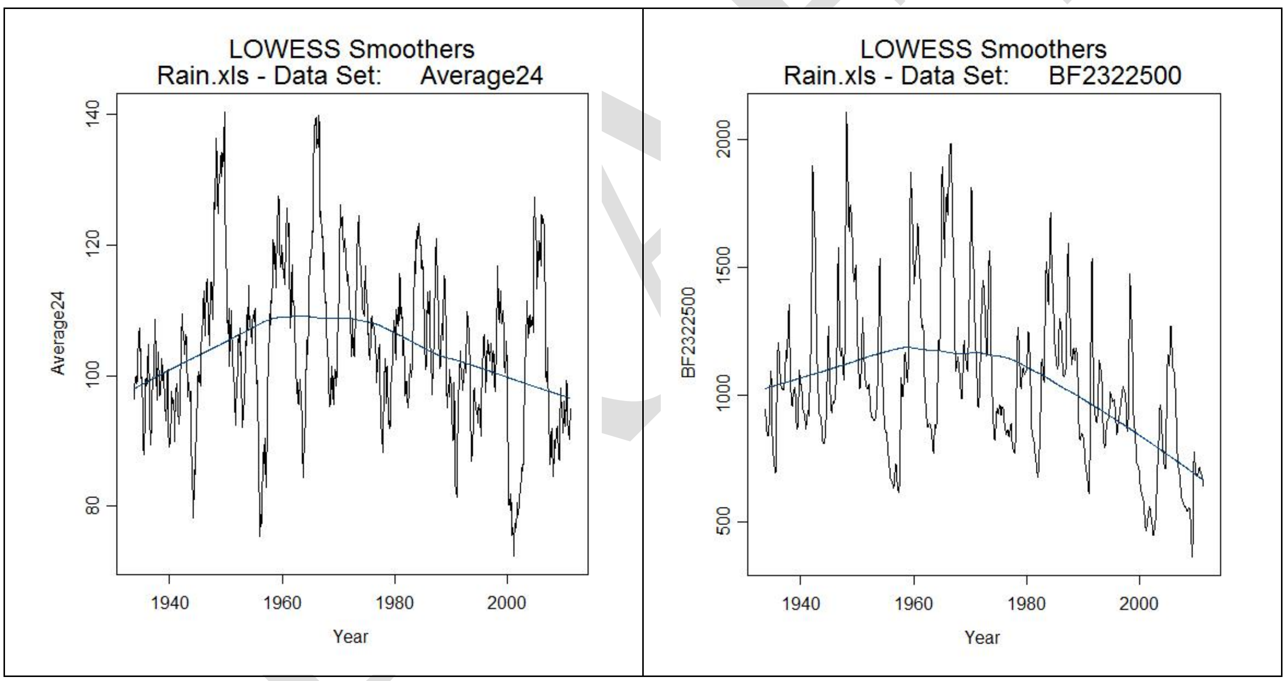



**Figure 69. LOWESS Comparison: Average 24-month Rainfall, inches (left), and Ft. White Baseflow, cfs (right)**

Regressive models can be developed for one of several purposes: to predict y given x, to estimate a variance for the prediction, to obtain the linear unbiased estimator of y, and to test hypotheses, estimate confidence intervals and prediction intervals. For the purposes of this analysis (to predict y given x), 2 assumptions must be true: the variables must be related in the correct form (y must be linearly related to x), and the data used to fit the model must be representative of the data of interest. For other intended uses of a regressive model, other assumptions must also be made regarding the data set, including homoscedasticity of residuals, independence of residuals, and normal distribution of residuals (Helsel and

Hirsch, 2002). It should be noted that the models developed for flow adjustments cannot be utilized to develop variances for prediction, for hypothesis testing, or for the estimation of prediction intervals due to the fact that the model residuals violate assumptions that are needed for these purposes. Additionally, the models developed for the adjustment of historical flows are not intended to be utilized to predict baseflow on a day-by-day basis, but rather to examine the long term behavior of baseflow. The flow adjustment models demonstrate valuable information regarding the relationship between rainfall and baseflow and indicate that this relationship changes over time, thus allowing for the calculation of flow adjustments.

In order to explain the changes in flow over the period of record for each gauge, simple linear regression (SLR) and multiple linear regression (MLR) models were constructed for each station to describe the relationship between flow and rainfall over the early time segment (pre-1970). Overall, the models provided a good fit for estimating base flow at each gauge using rainfall as the independent variables. Application of the flows generated by the pre-1970 models to the post-1970 rainfall data illustrates that there is systematic residual in the pre-1970 model which increases with time. This bias is due to the change in the influence of other factors in the dependent variable. These other factors include water use, land use changes, other anthropogenic factors, and other non-rainfall effects. In other words, as time increases, the early model becomes more and more erroneous because other factors are becoming more influential in describing the dependent variable (baseflow).

For each station for which the adjusted historical flow was to be determined, estimation of the flow adjustment involved the following steps:

- Baseflow was estimated as described (see *Exploratory Data Analysis and Trend Analysis*).
- Surface runoff was calculated by subtracting estimated baseflow time series from observed flow time series.
- A statistical model between rainfall and baseflow was developed for the period of streamflow record that was believed to be uninfluenced by human activities (before 1970s). Since the Lower Santa Fe and Ichetucknee Basins are highly groundwater driven, baseflow models were developed in order to isolate the impacts in the groundwater system. The buffering mechanism provided by the groundwater system filters the high frequency variability of flow and allows for the development of a more concise relationship between rainfall and baseflow.
- The statistical model was applied to rainfall time series for the entire period of record (through 2011).
- The residuals in the model performance were examined with respect to time. The trend of the residuals in the post-1970 period was examined. The trend in the residuals represents the degradation of the model fit due to the influence of factors other than rainfall.
- The trend in the residuals was used to alter the baseflow to generate an estimate of the adjusted baseflow by applying a daily adjustment to the historical flow which is calculated based on the slope of the trend in the model residuals.

- The adjusted historical total flow was calculated by adding surface runoff, from above, to estimated adjusted historical baseflow.

It should be noted that the models developed for flow adjustments are not intended to be utilized to predict discrete baseflow estimates for time series re-construction, but rather, are intended to serve as an indicator of general behavior at a given station, and assess the change in the relationship between rainfall and baseflow over time. That being said, it would be expected that the models would contain inherent error due to seasonality, which is acceptable given the overall intent of the models.

## Adjusted Historical Baseflow and Total Flow Time Series

The initial assessment of the long-term observed streamflow time series and the trend analysis revealed that 1970 was the break point; hence, the adjusted flow time series were developed from 1970 through 2011. Time series defining the adjusted historical flow conditions were developed for the Santa Fe River at Highway 441 near High Springs (2321975), Santa Fe River near Fort White (2322500) and Ichetucknee River at Highway 27 (2322700). The remaining stations on the Santa Fe River did not have a long-term flow record adequate for development of adjusted historical flows.

For each station, a statistical model was fit using average baseflow of that station and the previous 24 months of rainfall which was the average rainfall from 2 local gauges (Lake City and Gainesville). The average of the two available rainfall gauges was used in order to provide better spatial representation of rainfall throughout the Lower Santa Fe Basin.

### *Santa Fe River near Fort White (2322500)*

The MLR model between rainfall and baseflow for the pre-1970 Ft. White baseflow data is shown in Figure 70 and can be described by Equation (1). Model parameterization (coefficients and y-intercept) is shown in Table 34. All explanatory variables were significant in the regression. As shown in Equation (1), average baseflow in a given month is estimated using rainfall from the previous 24 months (the current month plus 23 prior months). The model fit exhibits a multiple $R^2$ of 0.7476, indicating that approximately 75% of the variance in baseflow can be explained by the MLR rainfall model.

$$BF = b_0 + \sum_{i=0}^{23} m_i a_i \qquad (1)$$

Where BF= average monthly baseflow, cfs

$b_0$ = y-intercept

$m_i$ = coefficient for month i

$a_i$ = total average rainfall for month i, inches

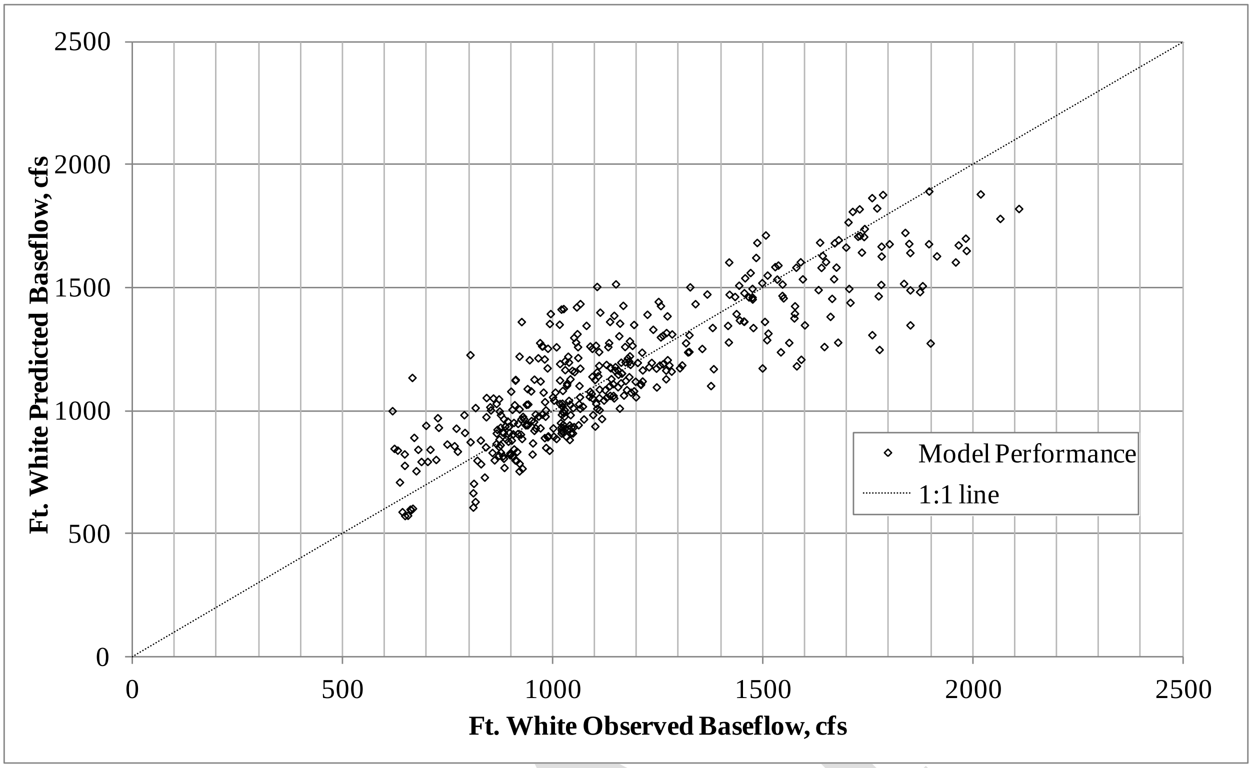



**Figure 70. 2322500: Baseflow Model Fit**

**Table 34. Ft. White MLR Model Fit**

| | m | p-value |
|---|---|---|
| Intercept, $b_o$ | -912.86 | 0.00 |
| a-0 | 14.27 | 0.00 |
| a-1 | 22.37 | 0.00 |
| a-2 | 29.33 | 0.00 |
| a-3 | 33.31 | 0.00 |
| a-4 | 33.00 | 0.00 |
| a-5 | 31.91 | 0.00 |
| a-6 | 28.29 | 0.00 |
| a-7 | 27.47 | 0.00 |
| a-8 | 24.72 | 0.00 |
| a-9 | 24.05 | 0.00 |
| a-10 | 23.14 | 0.00 |
| a-11 | 22.79 | 0.00 |
| a-12 | 21.13 | 0.00 |
| a-13 | 17.44 | 0.00 |
| a-14 | 13.32 | 0.00 |
| a-15 | 9.23 | 0.00 |
| a-16 | 9.19 | 0.01 |
| a-17 | 9.00 | 0.01 |
| a-18 | 12.00 | 0.00 |
| a-19 | 12.10 | 0.00 |
| a-20 | 13.07 | 0.00 |
| a-21 | 12.87 | 0.00 |
| a-22 | 12.36 | 0.00 |
| a-23 | 12.41 | 0.00 |

The model was applied to the period of record rainfall, and the model residuals were determined. The residual is defined as the difference between the model predicted baseflow and the observed (estimated) baseflow. Extrapolation of this model to the post-1970 data allows for the estimation of what the baseflow would have been had the relationship between rainfall and baseflow remained the same as it was pre-1970. The difference between this predicted baseflow and the estimated observed baseflow is the model residual. The daily residuals were examined for trends with time. Residuals with time should typically take the form of random noise. Structure in the residuals can indicate that seasonality and/or long-term trends were not taken into account (Helsel and Hirsch, 2002). As shown in Figure 71, there is seasonality present in the residuals throughout the analysis period. Since this analysis focused on assessing the average relationship between rainfall and baseflow on a long-terms basis, short-term seasonality was not considered. A long-term trend in residuals with time can suggest the need for additional terms in the regression equation. As shown, there is an increasing trend in the model residuals with time when the model is applied to the post-1970 data. This indicates that, on average, the model is predicting baseflows higher than observed, and as time progresses, the magnitude of the over-prediction increases. The systematic error in the model performance with the post-1970 rainfall indicates that the model is failing to describe a portion of the dependent variable. This, in effect, indicates that in the early portion of the record, a linear model between rainfall and baseflow provides an adequate representation of baseflow, but in the latter portion of the record, a linear model with rainfall as an explanatory variable cannot adequately describe baseflow due to the influence of anthropogenic factors. Using the trend in the model residuals, an adjustment was made to the time series, resulting in the adjusted historical baseflow shown in Figure 72. The total adjustment applied to the Ft. White record was 256.9 cfs at the end of the record (March 3, 2011) and the average annual adjustment was 6.2 cfs per year. This equates to a decrease in average flow of 0.41% per year based on an average historical flow of approximately 1500 cfs at Ft. White.

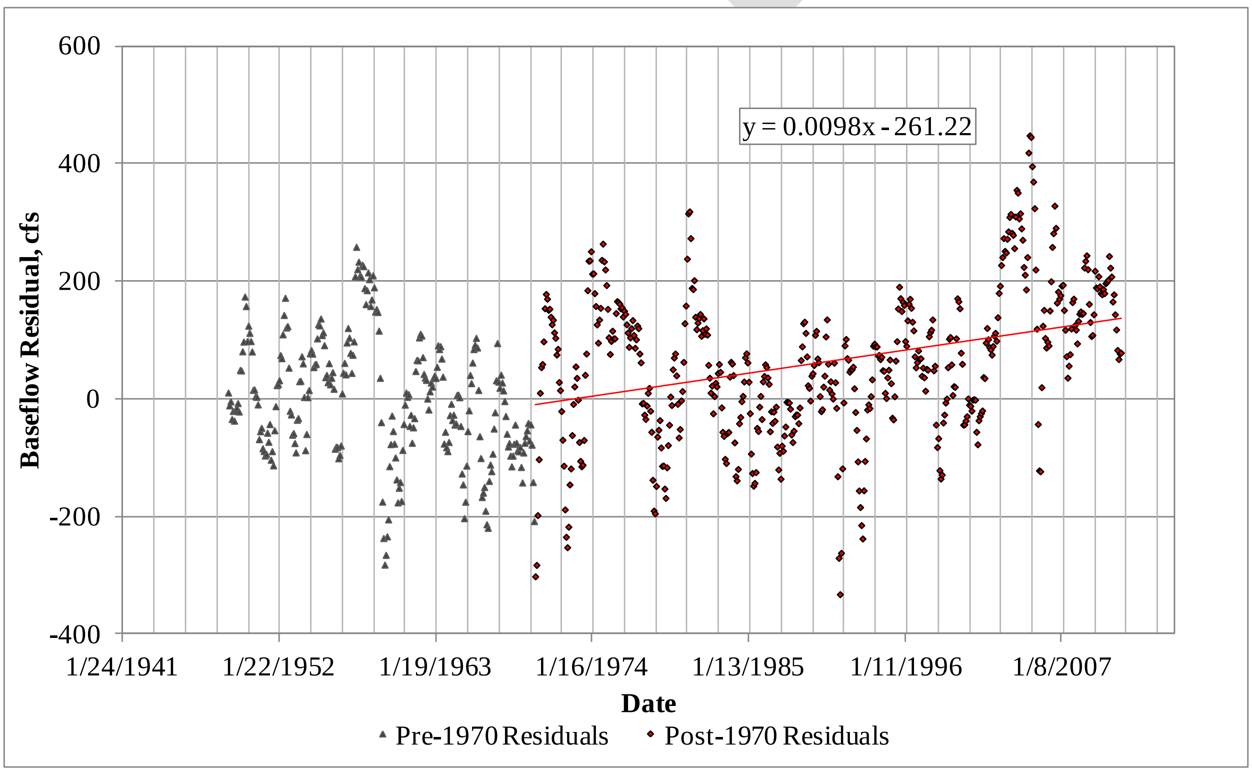
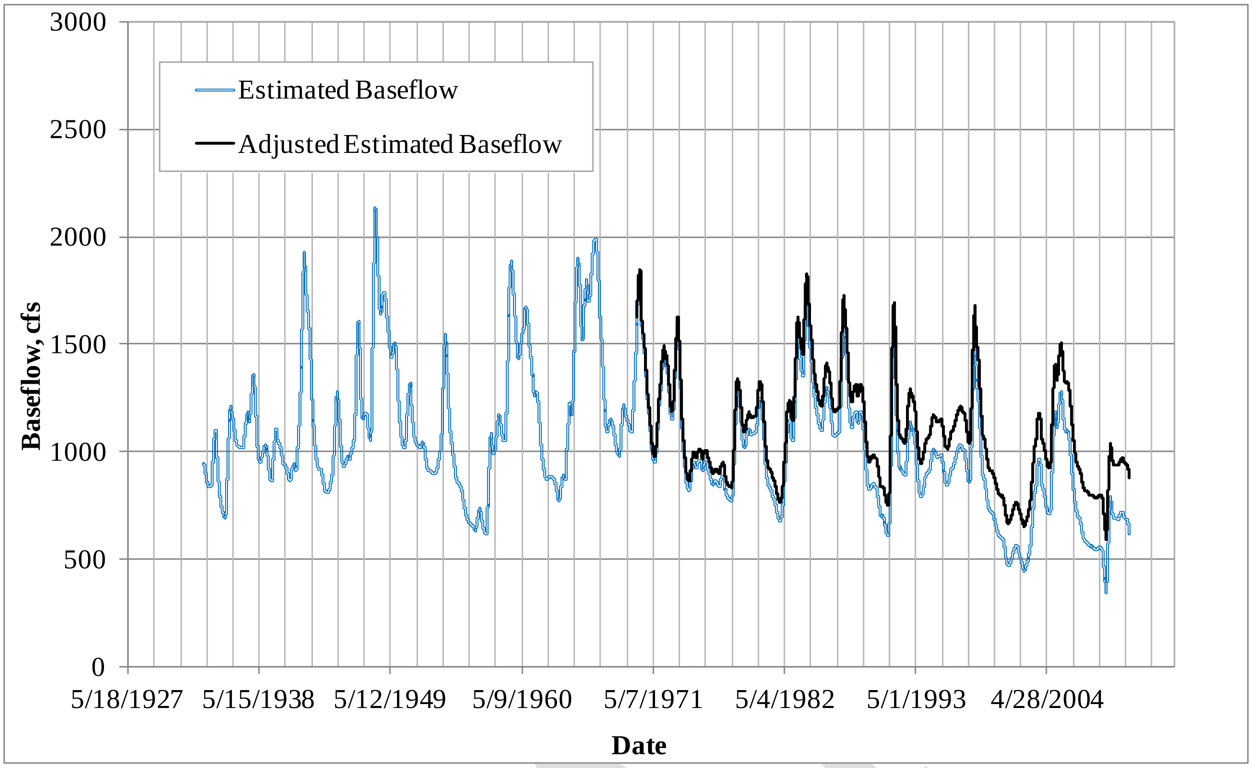



**Figure 71. 2322500: Baseflow Residuals**



3000
2500
2000
1500
1000
500
0
Baseflow, cfs
Estimated Baseflow
Adjusted Estimated Baseflow
5/18/1927
5/15/1938
5/12/1949
5/9/1960
5/7/1971
5/4/1982
5/1/1993
4/28/2004
Date

### *Santa Fe River at Highway 441 near High Springs (2321975)*

A MLR model between rainfall and baseflow for the pre-1970 Santa Fe Highway 441 data utilized the same form as the Ft. White model, as described by Equation (1). Model fit is shown in Figure 72, with parameterization (coefficients and y-intercept) as shown in Table 35. All explanatory variables were significant in the regression. The model fit exhibits a multiple $R^2$ of 0.7797, indicating that approximately 78% of the variance in baseflow can be explained by the MLR rainfall model. The model was applied to the period of record rainfall, and the model residuals were determined. As shown in Figure 73, there is a positive trend in the residuals after 1970. Using the trend in the model residuals, an adjustment based on the trend in the residuals was made to the time series, resulting in the adjusted historical baseflow shown in Figure 74. The total adjustment applied to the 441 record was 137.0 cfs at the end of the record (March 3, 2011) and the average annual adjustment was 3.6 cfs per year. This equates to a decrease in average flow of 0.51% per year based on an average historical flow of approximately 708 cfs at 441.

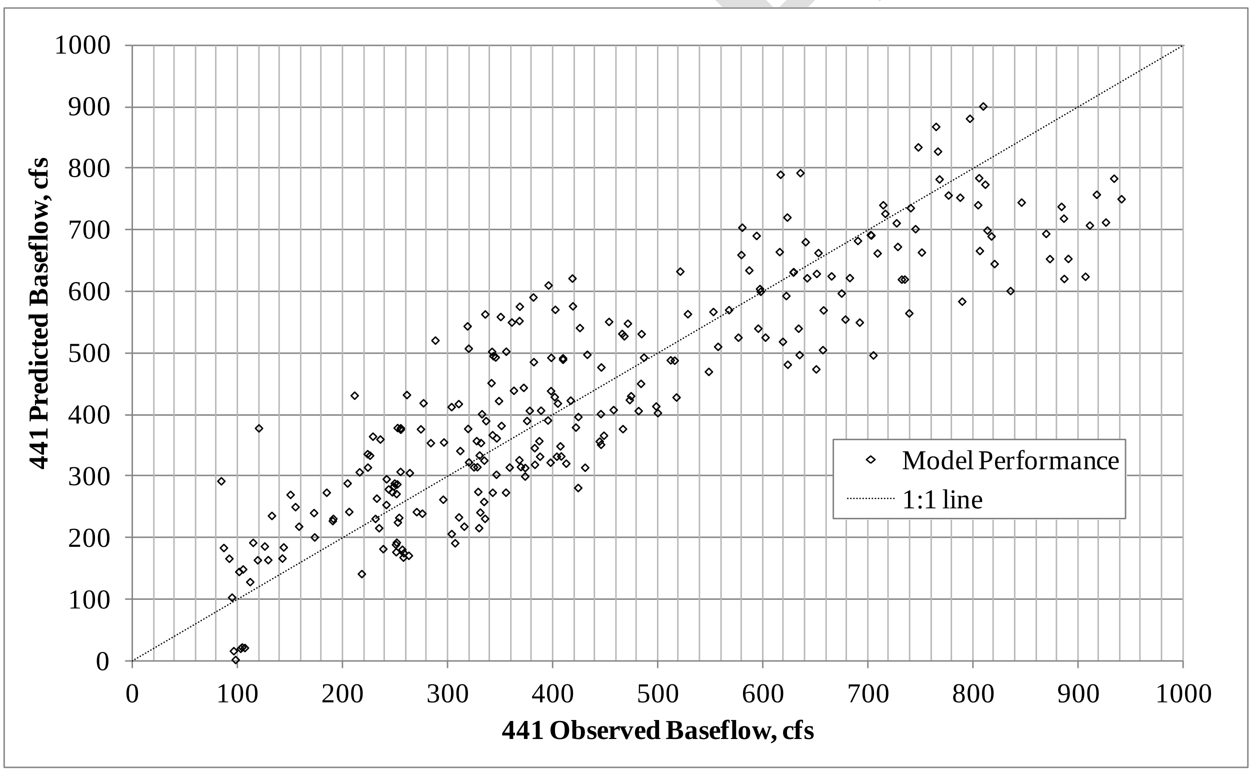



**Figure 72. 2321975: Baseflow Model Fit**

**Table 35. 441 MLR Model Fit**

| | b | p-value |
|---|---|---|
| Intercept | -1004.24 | 0.00 |
| a-0 | 10.17 | 0.00 |
| a-1 | 15.06 | 0.00 |
| a-2 | 18.76 | 0.00 |
| a-3 | 20.89 | 0.00 |
| a-4 | 20.38 | 0.00 |
| a-5 | 19.86 | 0.00 |
| a-6 | 18.22 | 0.00 |
| a-7 | 17.67 | 0.00 |
| a-8 | 15.28 | 0.00 |
| a-9 | 14.65 | 0.00 |
| a-10 | 14.19 | 0.00 |
| a-11 | 14.23 | 0.00 |
| a-12 | 14.05 | 0.00 |
| a-13 | 12.63 | 0.00 |
| a-14 | 10.65 | 0.00 |
| a-15 | 8.45 | 0.00 |
| a-16 | 8.59 | 0.00 |
| a-17 | 7.42 | 0.01 |
| a-18 | 8.49 | 0.00 |
| a-19 | 8.65 | 0.00 |
| a-20 | 9.70 | 0.00 |
| a-21 | 10.45 | 0.00 |
| a-22 | 10.36 | 0.00 |
| a-23 | 9.29 | 0.00 |

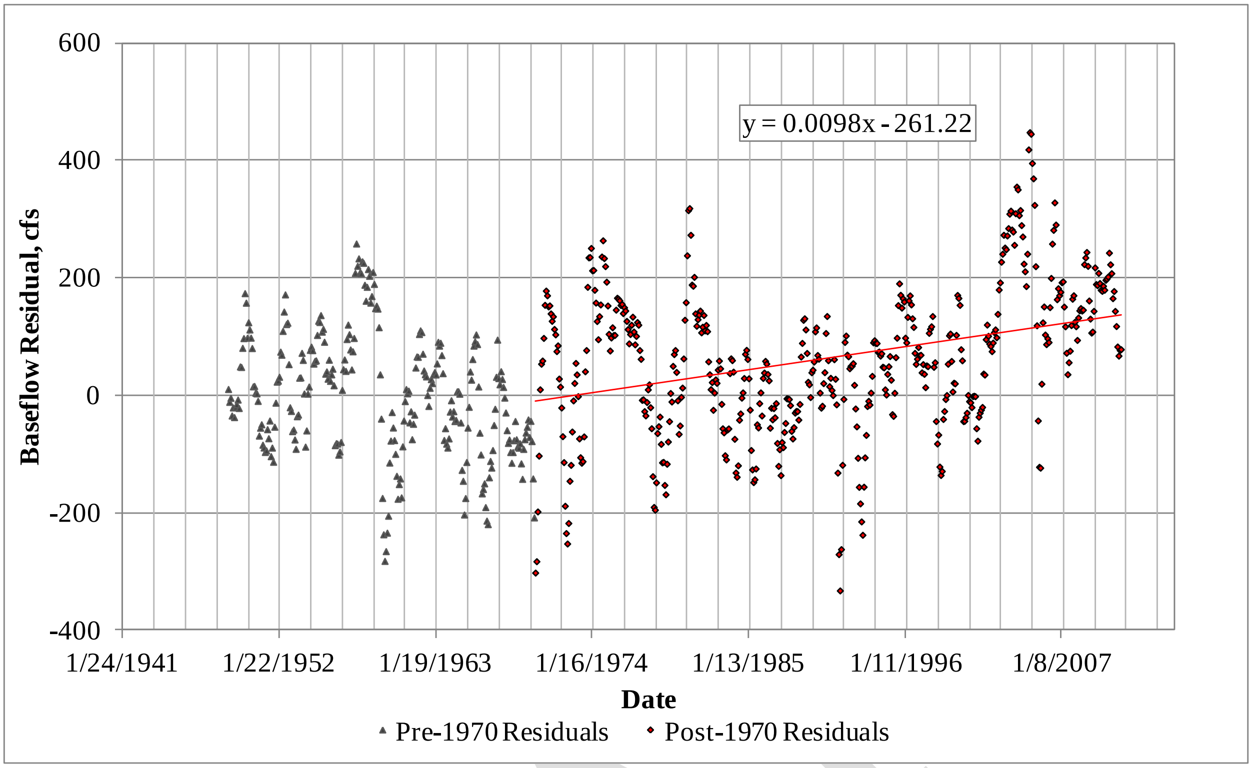



**Figure 73. 2321975: Baseflow Residuals**

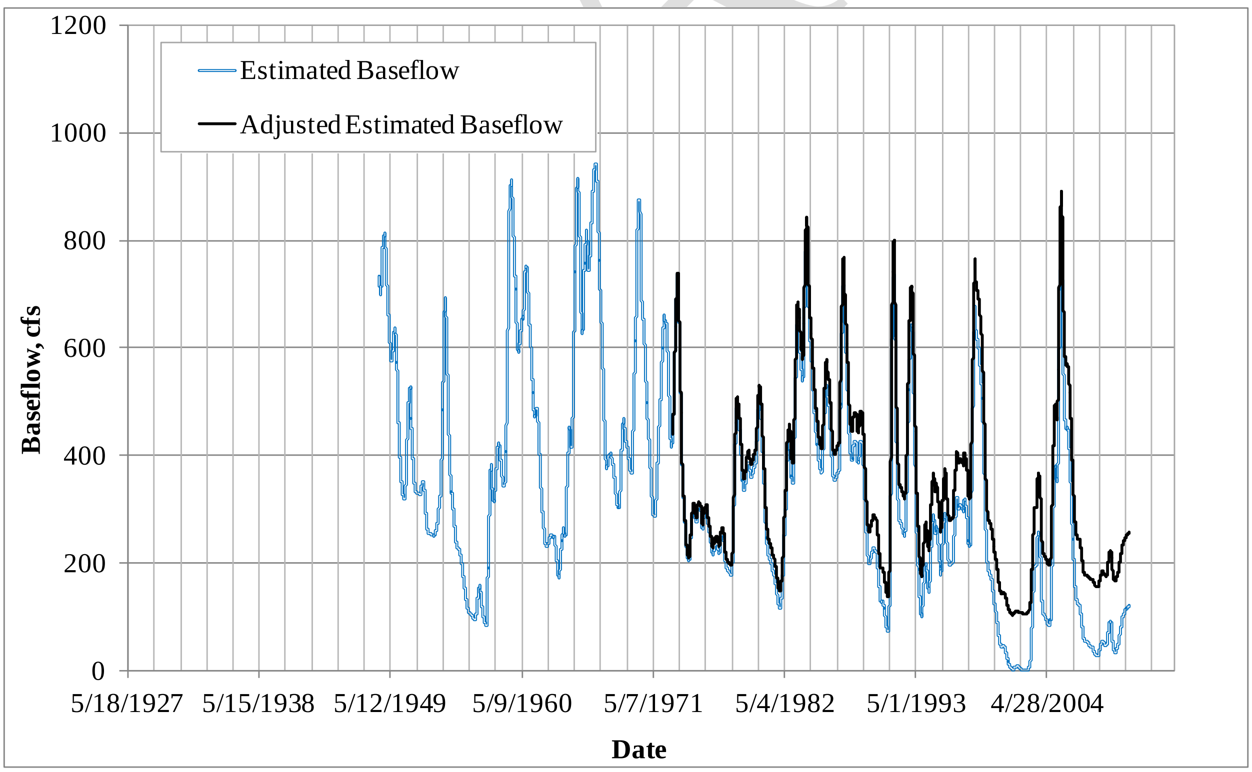



**Figure 74. 2321975: Estimated Baseflow and Adjusted Baseflow**

### *Ichetucknee River at Highway 27 (2322700)*

A SLR model was constructed for the prediction of baseflow at Highway 27. The model utilized the total 24-month rainfall accumulation to estimate annual baseflow. Model fit is shown in Figure 75. As shown, the SLR model describes approximately 81% of the variance in baseflow. Using the trend in the model residuals (Figure 76), an adjustment was made to the time series, resulting in the adjusted historical baseflow shown in Figure 77. The total adjustment applied to the Highway 27 record was 35.1 cfs at the end of the record (March 3, 2011) and the average annual adjustment was 1.1 cfs per year. This equates to a decrease in average flow of 0.3% per year based on an average historical flow of approximately 350 cfs at Highway 27.

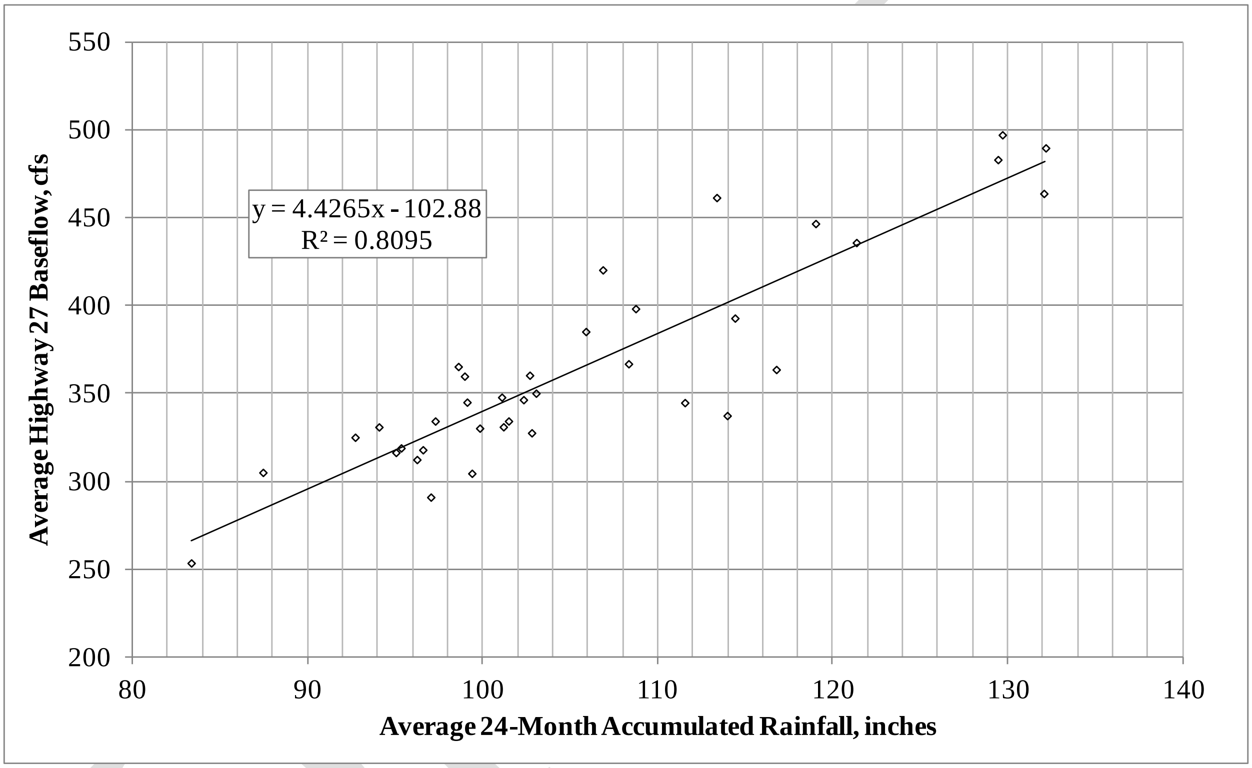



**Figure 75. 2322700 Baseflow Model Fit**

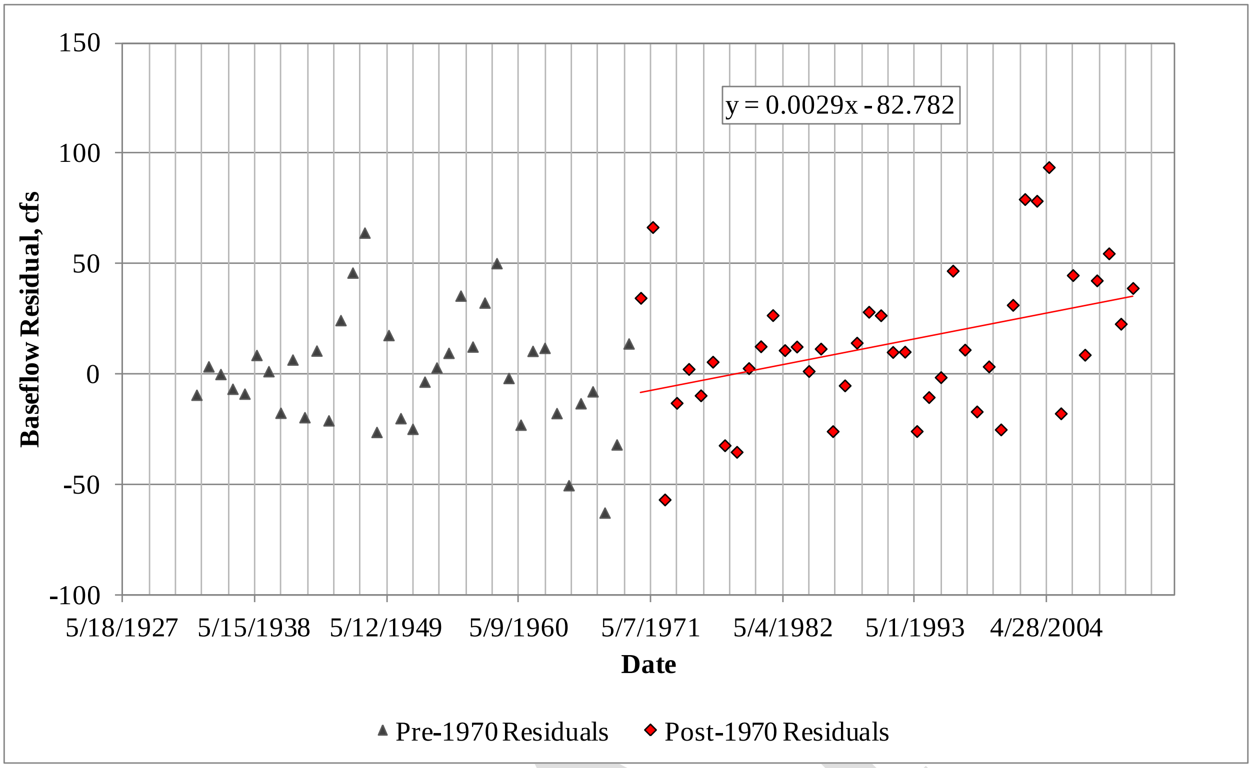



**Figure 76. 2322700: Baseflow Residuals**

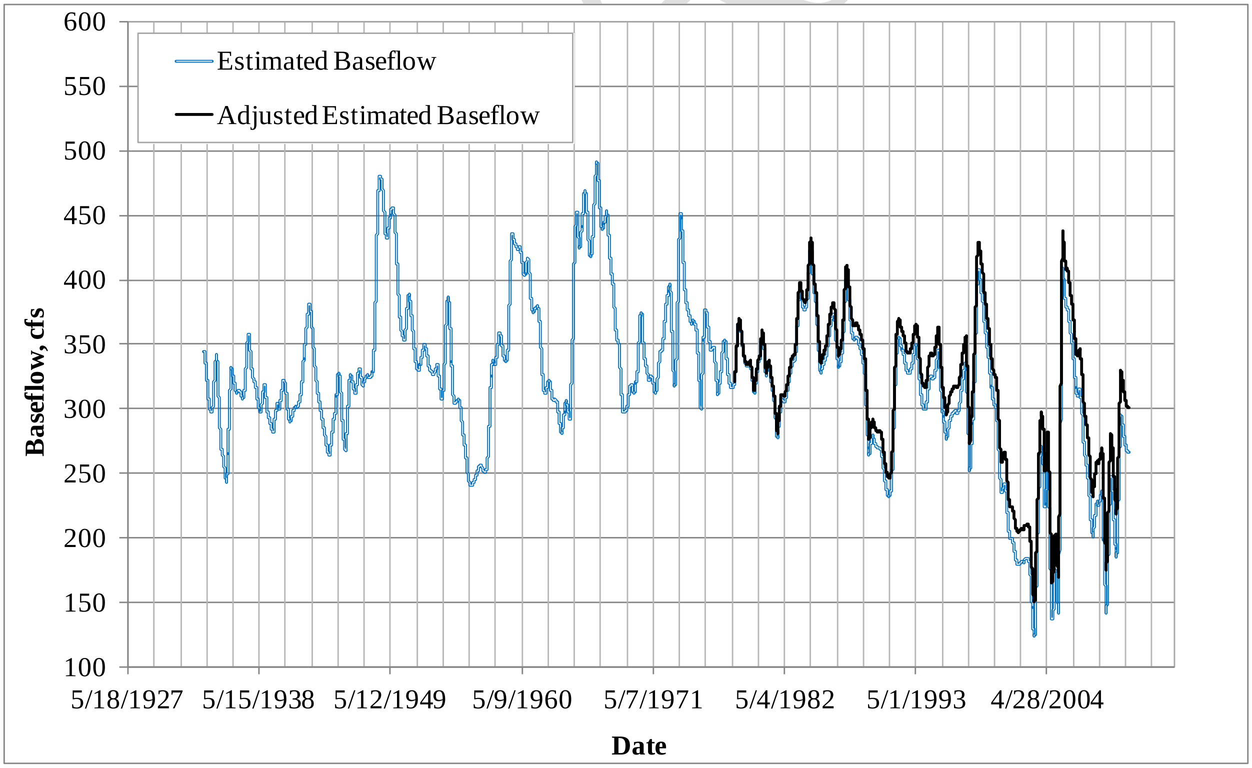



**Figure 77. Estimated Baseflow and Adjusted Baseflow**

### *Santa Fe River at Worthington Springs (2321500)*

The fit of the SLR model at Worthington Springs, shown in Figure 78, was poor. Unlike the other stations that were analyzed, the Worthington Springs station is dominated by surface runoff processes. The hydrologic response from surface water dominated basins is difficult to represent with linear models due to the non-linearity of the processes as well as the complexity of inter-related processes including rainfall intensity, evaporation, interception storage, depression storage, and inter-event time. In addition, the Worthington Springs gauge, unlike other gauging stations, is influenced by significant storage in Santa Fe Lake and Santa Fe Swamp. Despite the poor fit of the SLR model at the Worthington Springs station, the model was applied to the post-1970 data, resulting in a small (relative to the observed baseflow) trend in the model residuals in the post 1970 time period. The lack of a trend in the residuals combined with the fact that the basin upstream of Worthington Springs is well confined from the Upper Floridan aquifer and therefore isolated from groundwater impacts indicates the station shows minimal impacts due to changes in the groundwater system. Figures 78 and 79 show the fit of the SLR model and baseflow residuals at Worthington Springs. Although calculated, these adjusted baseflow time series were not used since the relative magnitude of computed impact was so small.

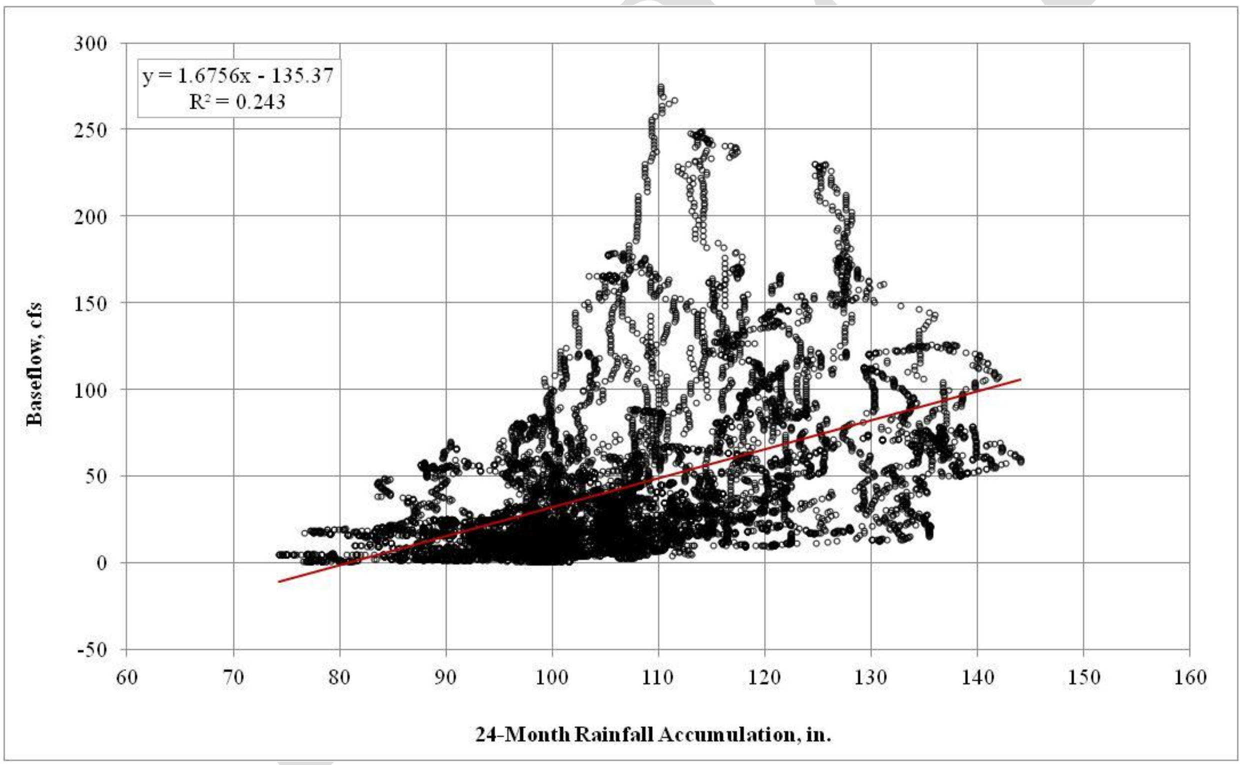



**Figure 78. 2321500: Baseflow Model Fit**

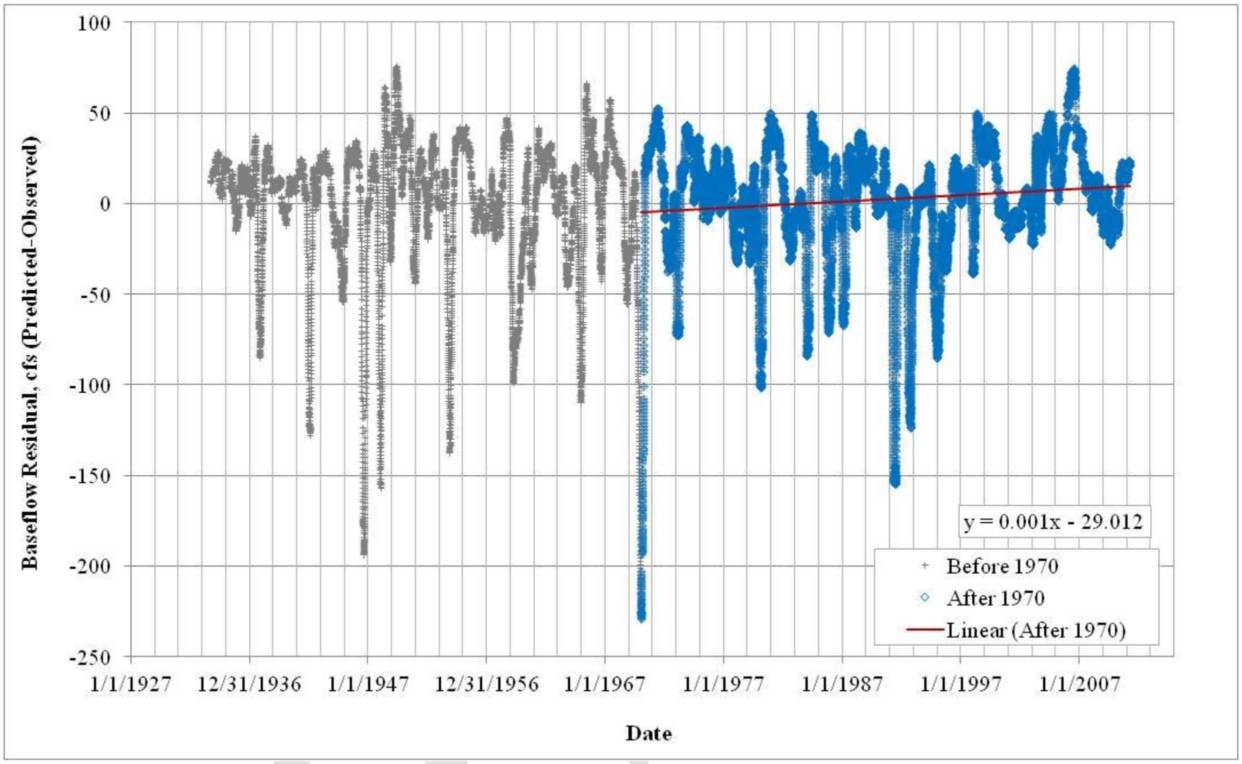



**Figure 79. 2321500: Baseflow Residuals**

### *Multivariate Models*

The inclusion of additional explanatory variables other than rainfall was also investigated. An extended time series of temperature (pre-1900 through present) was available at the Lake City 2e station. Since there was no evaporation data available at this station, the inclusion of temperature in the MLR models for the Santa Fe River stations was investigated for use as a surrogate for evaporation. Both average temperature and maximum temperature were included in the rainfall MLRs (separately) with the following results:

- Average temperature was not significant in the MLR when included with 24 months of rainfall at 441.
- Maximum temperature was not significant in the MLR when included with 24 months of rainfall at 441.
- Average temperature was not significant in the MLR when included with 24 months of rainfall at Ft. White.
- Maximum temperature **was** significant in the MLR when included with 24 months of rainfall at Ft. White.

In summary, temperature was only significant in the MLR model for Ft. White when maximum temperature was utilized. The application of the Ft. White temperature/rainfall model was compared to the application of the Ft. White rainfall model as shown below (Figure 80). As shown, differences between the models were minimal (and not visible when plotted). The average difference between the monthly residuals was 3.5cfs over the period of record.

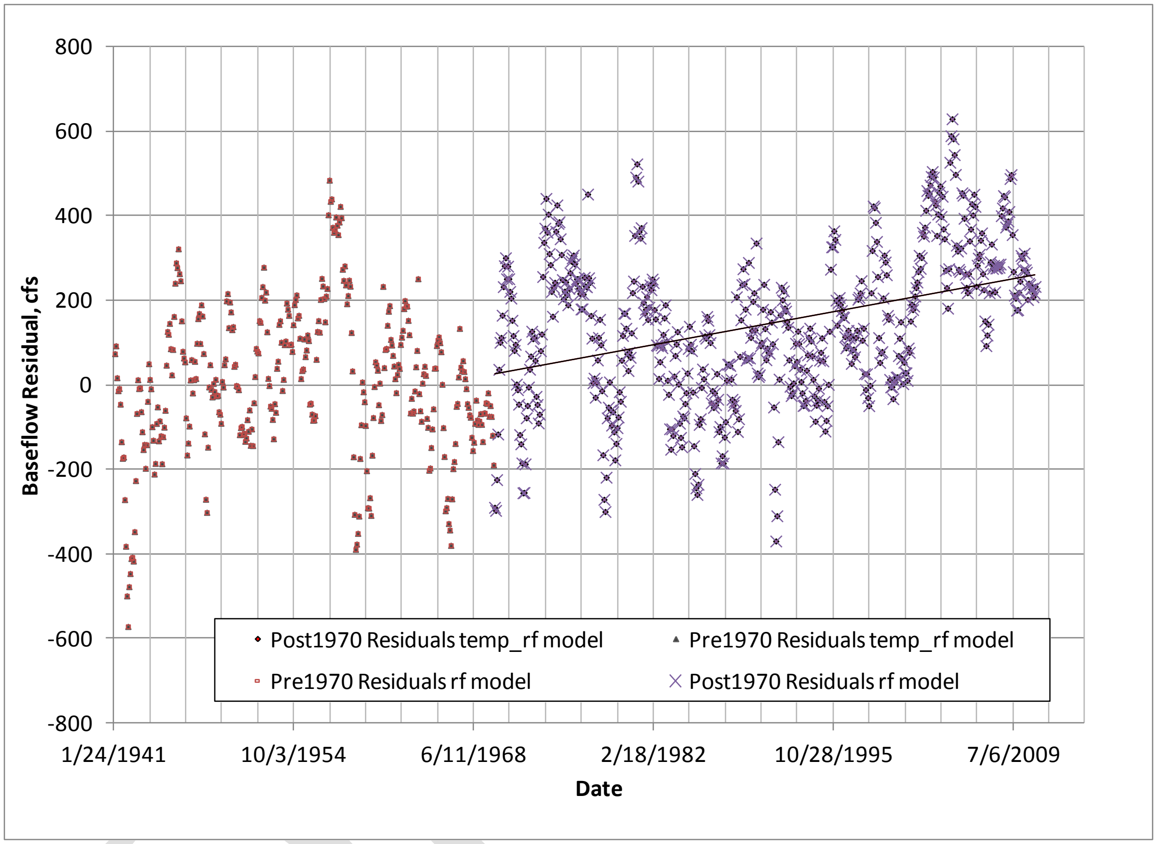



**Figure 80. Ft. White, Model Comparison, Max Temperature Included**

### *Adjusted Historical Flow Summary*

The historical and adjusted historical discharge time series were aggregated to annual means in order to examine the differences between the mean annual flows, shown in Figures 81 through 83. Both time series exhibit declining trends in streamflow. The declining linear trend in the observed flow time series (shown in black) is due to climatic variations, anthropogenic changes, and other non-rainfall effects. The declining trend in the adjusted flow time series (shown in red) is due to rainfall effects only. Since the declining trend in rainfall influences the calculated flow, this trend line slope is decreasing at a shallower rate than the observed data (which shows the effects of both rainfall and the groundwater system).

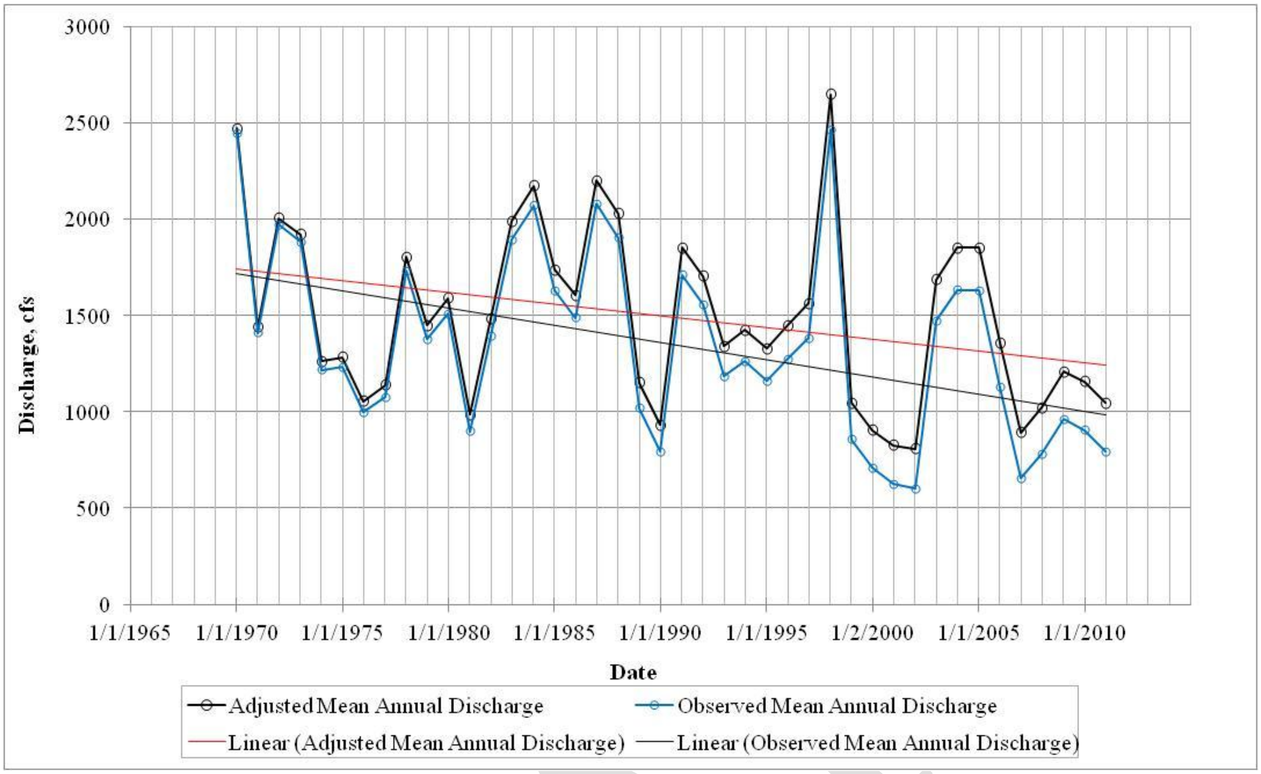



**Figure 81. Santa Fe River near Fort White: Mean Annual Adjusted and Observed Discharge**

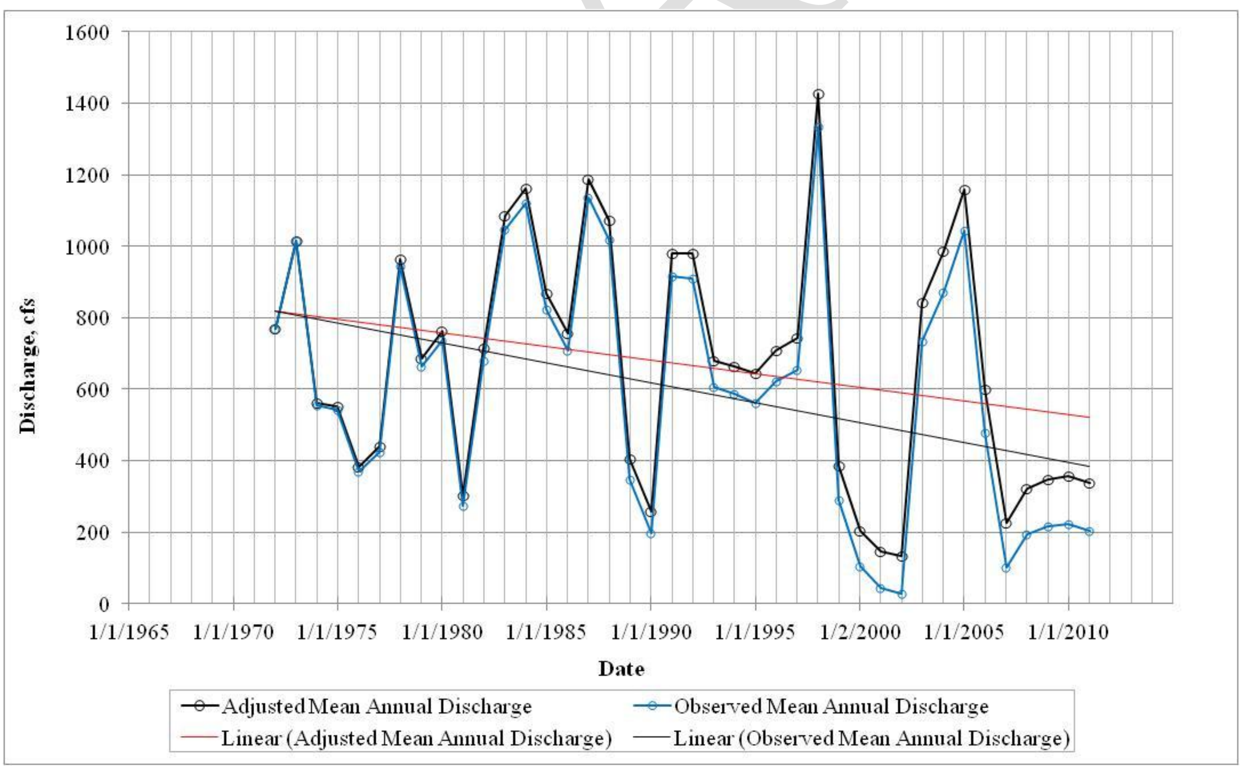



**Figure 82. Santa Fe River at Highway 441 near High Springs: Mean Annual Adjusted and Observed Discharge**

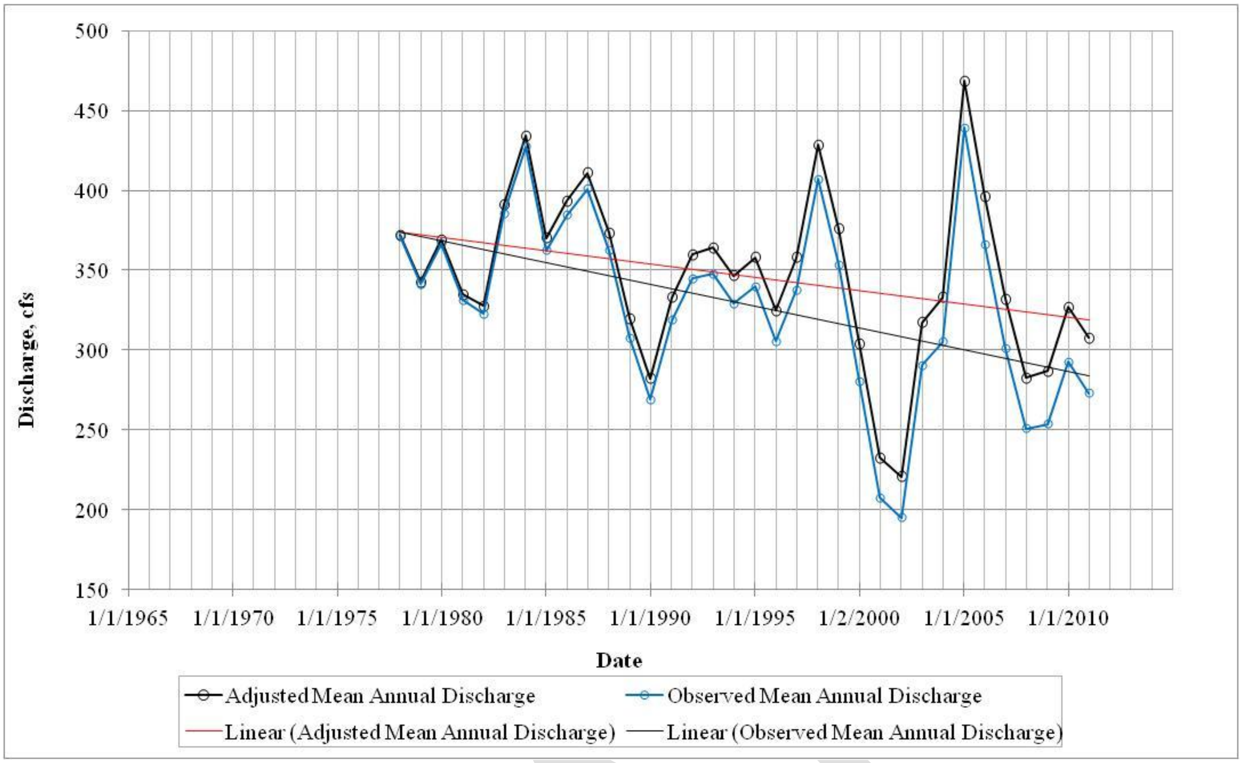



**Figure 83. Ichetucknee River at Highway 27: Mean Annual Adjusted and Observed Discharge**

A summary of flow adjustments and flow reduction percentages is shown in Table 36. As shown, based on the average historical flow, flow reductions range between 10% at the Ichetucknee River at Highway 27 to 19.4% at the Santa Fe River at 441. The Santa Fe at Ft. White station had the highest total adjustment of 256.9 cfs, which equated to approximately 17% of the historical flow. This is expected since it is the most downstream station on the Lower Santa Fe and experiences the highest flows. It should be noted that the adjustments were applied to the daily time series, at a rate of 1.1 cfs per year, 3.6 cfs per year, and 6.2 cfs per year for stations 2322700, 2321975, and 2322500, respectively. This equates to no more than 0.51% of the average flow per year, as shown.

**Table 36. Adjustment Summary**

| | Ichetucknee Highway 27 (2322700) | Santa Fe 441 (2321975) | Santa Fe Ft. White (2322500) |
|---|---|---|---|
| Total Adjustment, cfs (Adjustment at end of record) | 35.1 | 137.0 | 256.9 |
| Annual Average Adjustment, cfs/year (Total Adjustment/Years of Adjustment) | 1.1 | 3.6 | 6.2 |
| Average Historical Flow, cfs | 351.5 | 707.5 | 1501.9 |
| Average *Adjusted* Historical Flow, cfs | 358.8 | 749.2 | 1572.3 |
| Average Historical Baseflow, cfs | 322.9 | 347.4 | 1048.7 |
| Average *Adjusted* Historical Baseflow, cfs | 330.1 | 389.1 | 1122.4 |
| Percent Historical Flow Reduction | 10.0% | 19.4% | 17.1% |
| Percent *Adjusted* Historical Flow Reduction | 9.8% | 18.3% | 16.3% |
| Percent Baseflow Reduction | 10.9% | 39.4% | 24.5% |
| Percent *Adjusted* Baseflow Reduction | 10.6% | 35.2% | 22.9% |
| Percent Historical Flow Adjustment per Year | 0.30% | 0.51% | 0.41% |

### *Adjusted Historical and Historical Flow Comparison*

A trend analysis was conducted on the adjusted historical flow record for comparison to the observed flow record and rainfall trends. As demonstrated through the LOWESS analysis and as shown in Figure 84, there are 2 periods of monotonic trends in present in rainfall which were divided in 1970. The division of rainfall into wetting and drying periods in 1970 is also consistent with analysis done by others (Kelly, 2004) where rainfall data was divided into 2 data sets based on the warming and cooling phases of the Atlantic Multidecadal Oscillation (AMO). Similar trend analysis was conducted for all flow stations with adjusted flows. Figures 85 through 87 show the trend analysis of the original flow time series (left) and the adjusted flow time series (right). As shown, the adjustments made to the flow records maintain the downward trend in period 2 of the original flow time series, but the slope of the trend has decreased. A summary of the changes in Sen Slope (slope of the trend) is shown in Table 37.

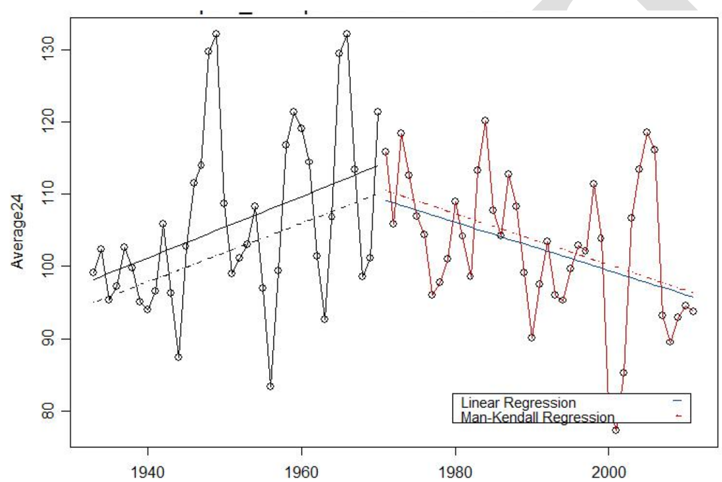



**Figure 84. Rainfall Trends**

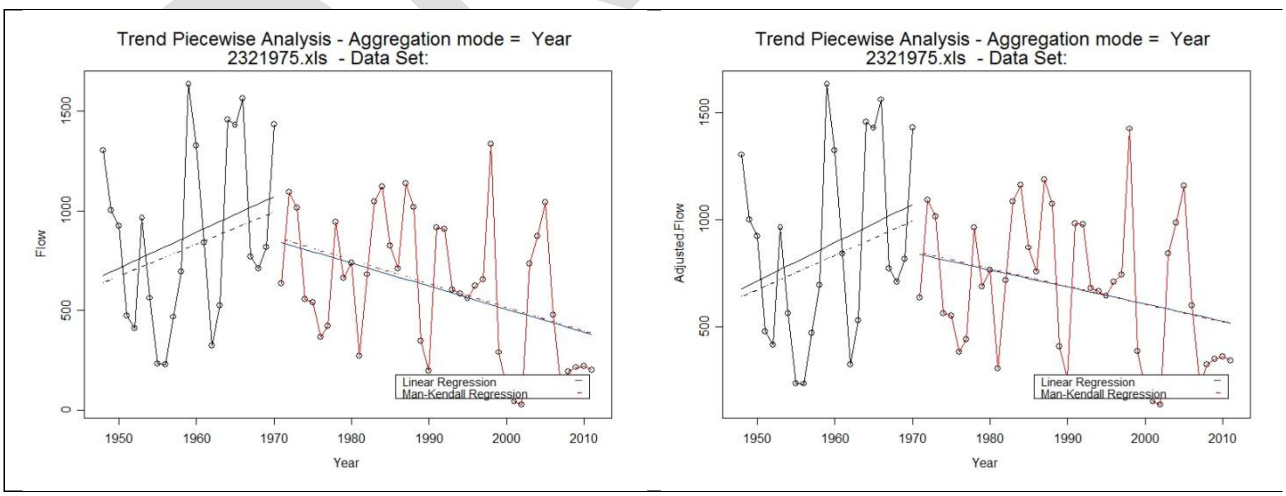



**Figure 85. Santa Fe near Highway 441 (2321975) Trends: Historical Flow (left), and Adjusted Historical Flow (right)**

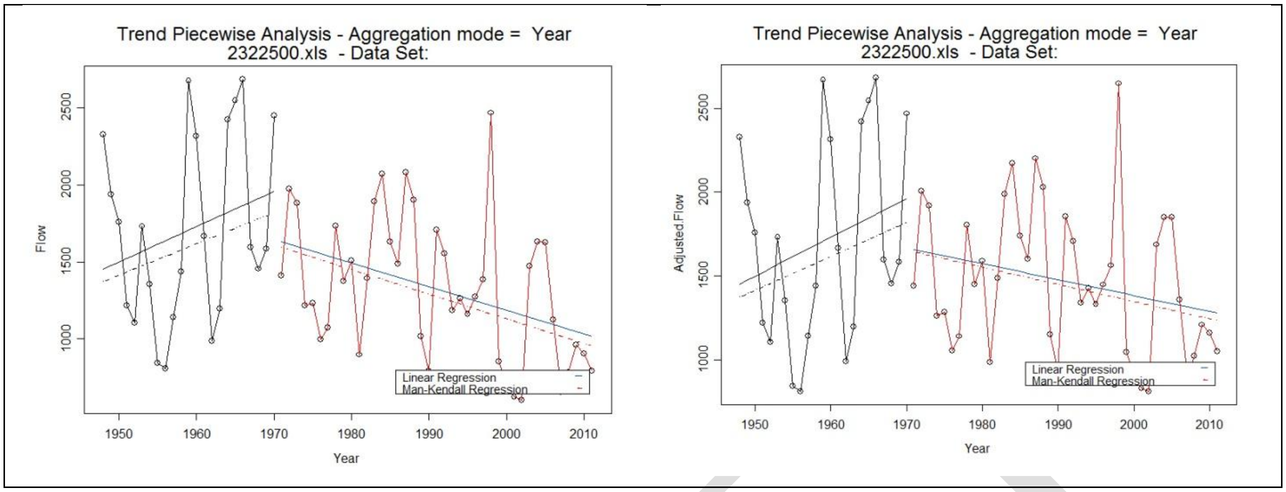



**Figure 86. Santa Fe near Ft. White (2322500) Trends: Historical Flow (left), and Adjusted Historical Flow (right)**

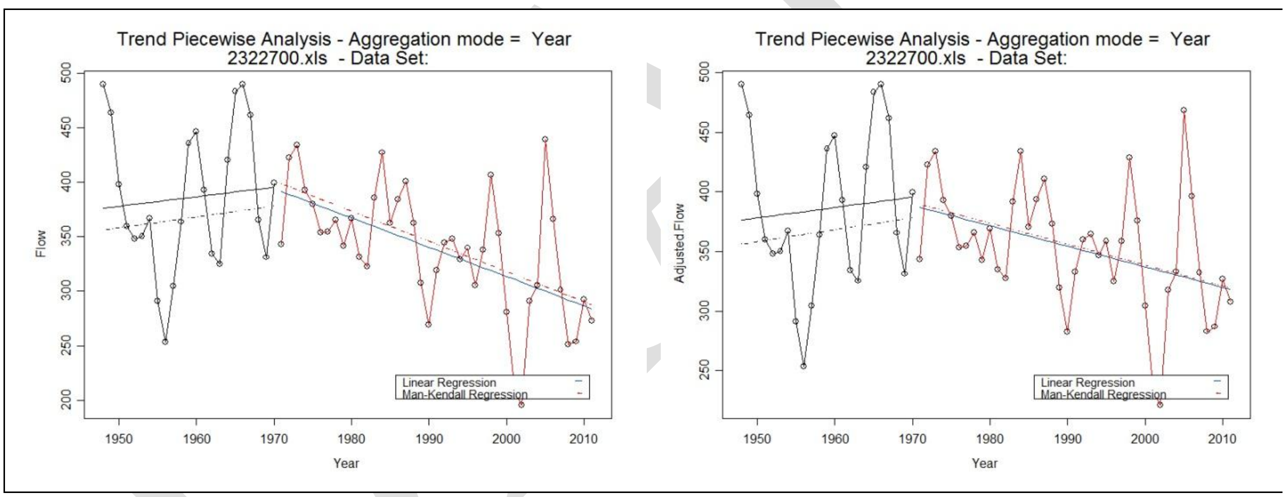



**Figure 87. Ichetucknee River Highway 27 (2322700) Trends: Historical Flow (left), and Adjusted Historical Flow (right)**

**Table 37. Trend Comparison: Historical and Adjusted Historical Flows**

| Station Number | Station Name | Period 1: Historical Slope cfs/year | Period 2: Historical Slope cfs/year | Period 2: Adjusted Historical Slope cfs/year |
|---|---|---|---|---|
| 2321975 | Santa Fe at 441 | 16.08 | -11.84 | -8.34 |
| 2322500 | Santa Fe at Ft. White | 20.27 | -15.97 | -10.24 |
| 2322700 | Ichetucknee at Highway 27 | 0.99 | -2.78 | -1.75 |

Exceedence plots for the adjusted historical flow and the historical flow time series were compared in order to determine the effect of flow adjustments on flow percentiles (Figures 88 through 90). A summary of the differences in the exceedence percentiles is shown in Table 38. As shown, at high percentiles (low exceedences) the differences between the historical flow and the adjusted historical flow are less than at low percentiles. All mean flows increased slightly, with Highway 27 mean flow increasing by approximately 7cfs over the period of record, and Ft. White increasing by approximately 71cfs over the period of record.

**Table 38. Excedence Percentiles**

| | Mean Flow, cfs | | 10th Percentile (90% Exceendence), cfs | | 50th Percentile (50% Exceedence), cfs | | 90th Percentile (10% Exceedence), cfs | |
|---|---|---|---|---|---|---|---|---|
| | Hist. | Adj. His | Hist. | Adj. His | Hist. | Adj. His | Hist. | Adj. His |
| Santa Fe 441 | 707.5 | 749.2 | 128.5 | 208.2 | 495.9 | 528.6 | 1478.0 | 1507.3 |
| Santa Fe Ft. White | 1501.9 | 1572.4 | 766.0 | 884.1 | 1240.0 | 1305.2 | 2500.0 | 2544.2 |
| Ichetucknee Hwy 27 | 351.5 | 358.8 | 271.0 | 285.9 | 346.1 | 352.0 | 442.5 | 448.3 |

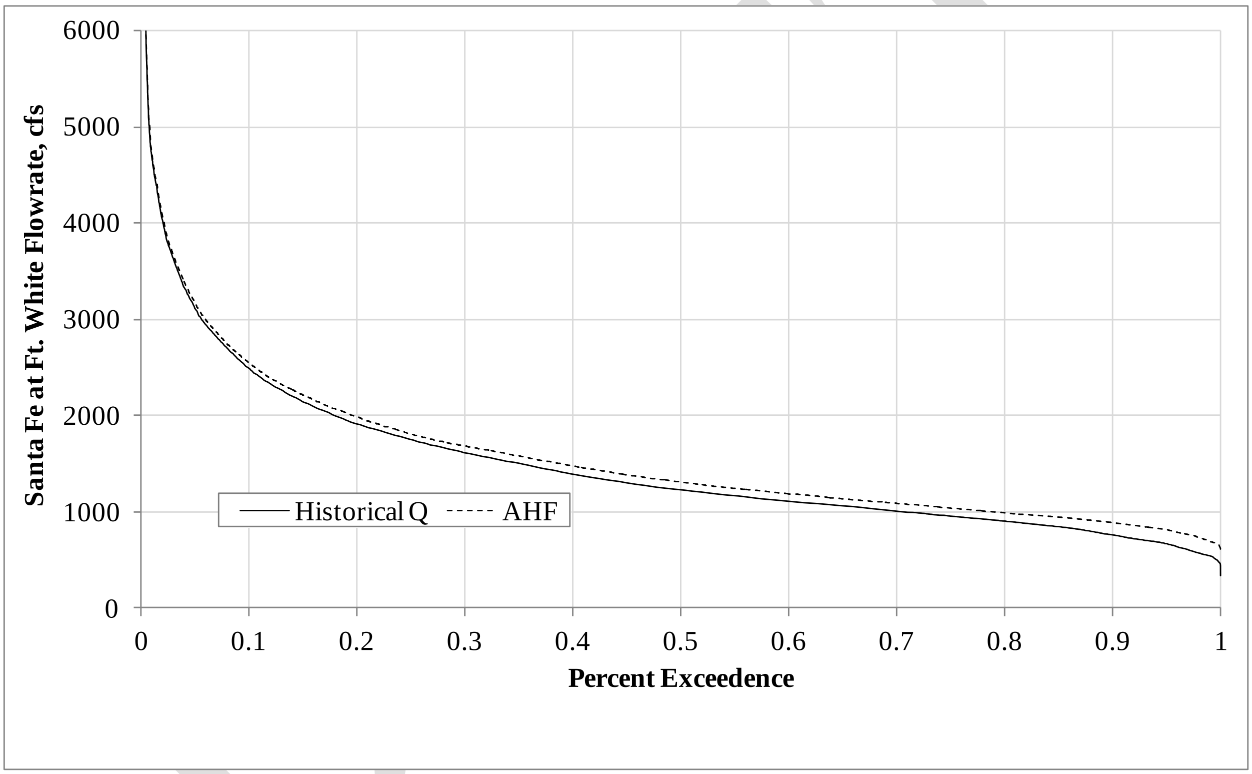



**Figure 88. Santa Fe at Ft. White Exceedence Plot**

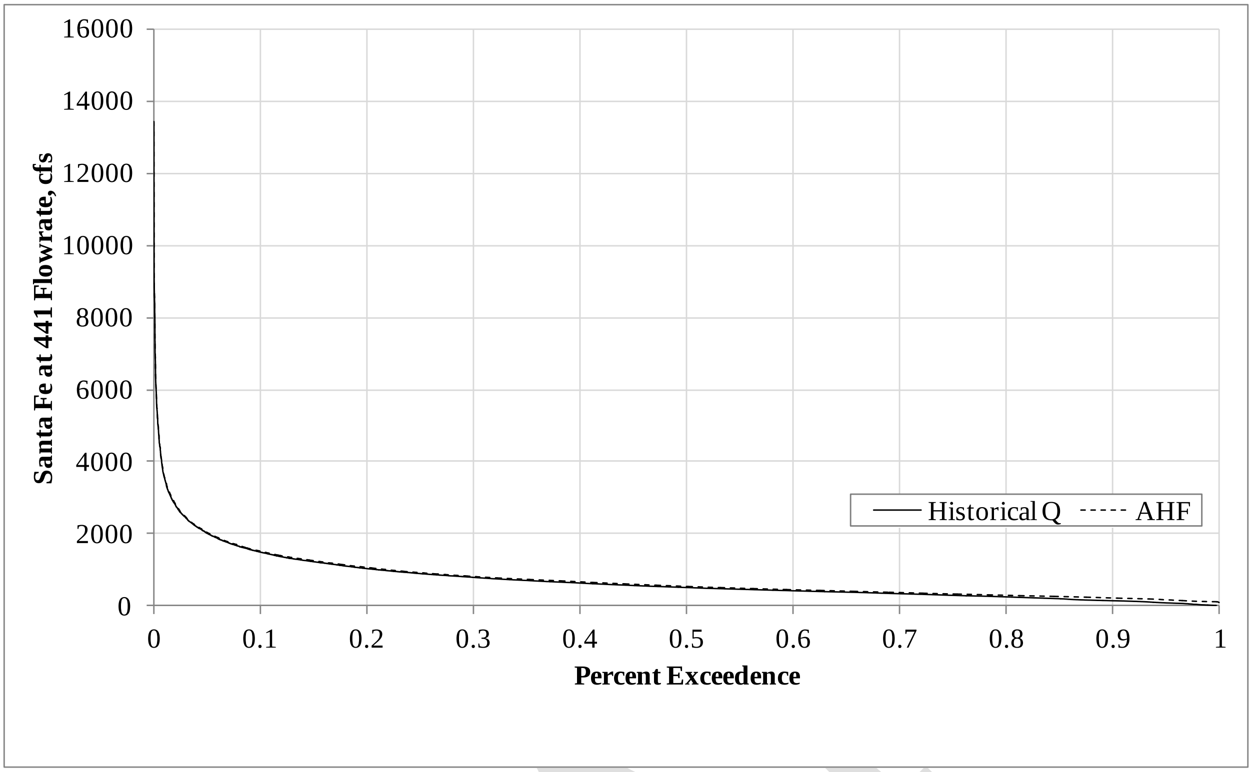



**Figure 89. Santa Fe at 441 Exceedence Plot**

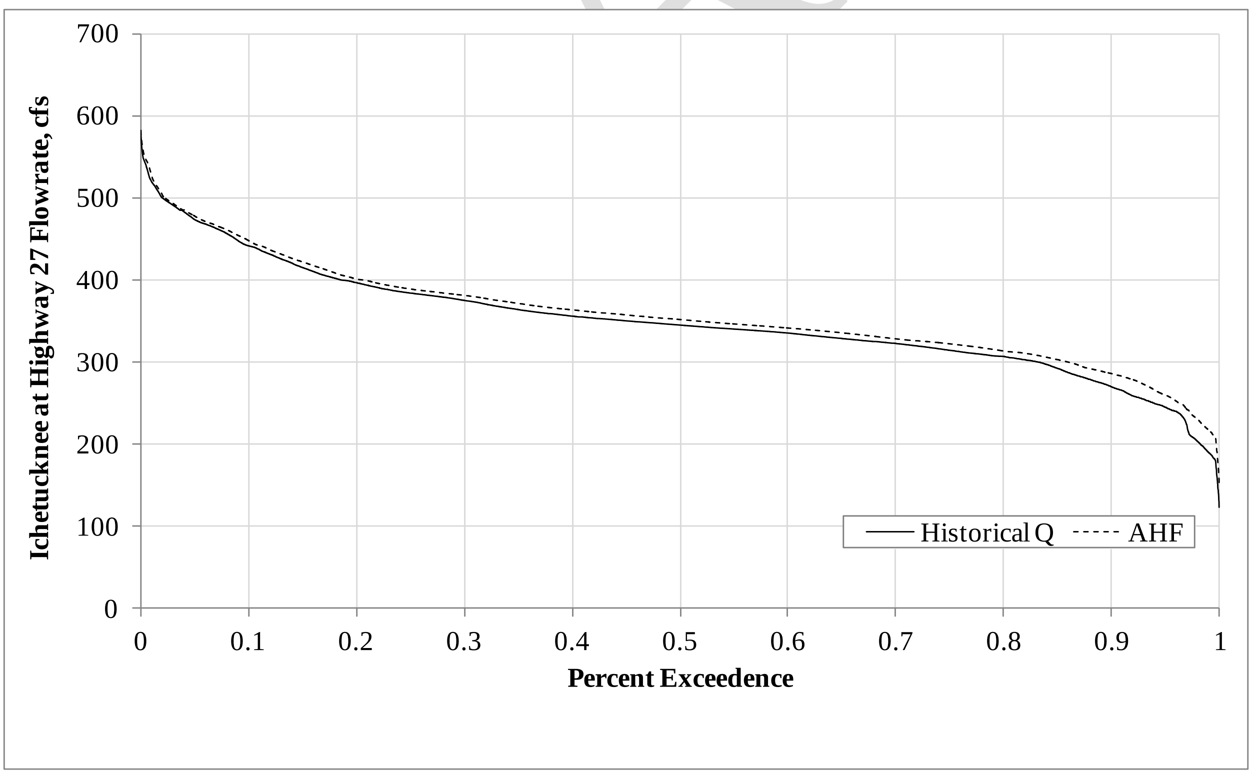



**Figure 90. Ichetucknee at Highway 27 Exceedence Plot**

## Conclusions

A comprehensive flow database of the Lower Santa Fe River was constructed for the District in support of MFL development. When necessary, gap filling was employed in order to develop complete time series for use in modeling efforts. For each station, the most appropriate method was selected based on the data availability and the length of the gap. Gap filling methods included linear interpolation, simple linear regression, multiple linear regression, and artificial neural networks.

Trend analysis on the resulting time series revealed the presence of statistically significant trends in the flow and baseflow time series over the period of record. Local rainfall also experienced a decreasing trend in the latter portion of the record. Although there were decreasing trends in both rainfall and flow, flow was decreasing at a steeper rate than rainfall.

A methodology was developed to filter anthropogenic impacts to river flows. The methodology employed linear regression analyses to develop a relationship between rainfall and baseflow. Based on the degradation of the model fit over time, flow adjustments to account for changes in groundwater pumping were estimated. Baseflow adjustments, shown in Table 39, range from 35.1 cfs on the Ichetucknee River to 256.9 cfs at the Santa Fe at Ft. White station. It should be noted that this adjustment was applied using the average annual adjustments shown, which range from 1.1 cfs per year (Highway 27) to 6.2 cfs per year (Ft. White). Over the post-1970 period, the average adjustments ranged from 17.6 cfs to 128.5 cfs. As a percent of total flow, the final adjustments ranged from 10% to 19.4%.

**Table 39. Flow Adjustment Summary**

| Station ID | Station Name | Adjustment at end of record, cfs | Average adjustment, cfs | Average annual adjustment, cfs/year | Percent Historical Flow Reduction |
|---|---|---|---|---|---|
| 2321975 | Santa Fe Highway 441 | 137.0 | 68.5 | 3.6 | 19.4% |
| 2322500 | Santa Fe Ft. White | 256.9 | 128.5 | 6.2 | 17.1% |
| 2322700 | Ichetucknee Highway 27 | 35.1 | 17.6 | 1.1 | 10% |

Based on the calculated baseflow adjustments, a total flow adjusted historical time series was developed for each station of interest. The adjusted historical flow time series exhibit decreasing trends similar to the precipitation record, indicating that the anthropogenic signature has been removed.

The above documented analysis is a statistically based estimate which is based on the best available hydrologic data. It is an estimate because a direct measurement of anthropogenic impacts to flow is impossible. The methodology was limited in scope therefore only simple linear regression models were employed. It is possible to utilize more statistically rigorous and complex models (multiple linear regressions, artificial neural networks, etc.) which may improve the fit of the models and therefore improve the estimate of impacts to river flows. Even more complex physically based numerical models could be used to define the impacts to river flows. The numerical model would have to be a complex groundwater and surface water integrated hydrologic model to capture all the processes that are found within the Santa Fe River watershed. Monte Carlo based model applications (either statistical or numerical) would be able to capture the distribution of estimated anthropogenic impacts to river flows and bound the result with confidence intervals. All possible techniques to estimate the anthropogenic impacts to river flows will have some degree of error and therefore uncertainty of the predictions of the adjusted historical record. The relative improvement from using more complex methods is unknown.

## References

Aly, A., and Biggs, T., (2010). *SPLUS Statistical Modules for the Analysis of Trends in Water Quality, Water Levels, and Spring Discharge, Version 2.0, User's Manual.* Prepared for the St. Johns River Water Management District.

ARCADIS (2010). *Unimpaired Flow Data Report: Surface Water Availability Modeling and Technical Analysis for Statewide Water Management Plan*. Prepared for Georgia Department of Natural Resources.

Birsan, M.-V., Molnar, P., Burlando, P., Pfaundler, M. (2005). *Streamflow Trends in Switzerland*. Journal of Hydrology 314: 312-329.

Bolger B.L, Park, Y.-J., Unger, A.J.A., Sudicky.E.A. (2011). *Simulating the Pre-development Hydrologic Conditions in the San Joaquin Valley, California.* Journal of Hydrology 411:322-330.

Christian, L.N., Banner, J.L., Mack, L.E. (2011). *Sr Isotopes as Tracers of Anthropogenic Influences on Stream Water in the Austin, Texas, Area.* Chemical Geology 282: 84-97.

Coarsey, C. (2012). *Ichetucknee River Discharge Measurement Discussion with USGS*. Inter-Office Memorandum to John C. Good, Suwannee River Water Management District.

Gao, P., Mu, X.-M., Wang, F., and Li, R. (2011). *Changes in Streamflow and Sediment Discharge and the Response to Human Activities in the Middle Reaches of the Yellow River*. Hydrology and Earth System Sciences 15: 1-10.

Grubbs, J.W. (2011). Analysis of Long-Term Trends in Flow from a Large Spring Complex in Northern Florida. U.S. Geological Survey Karst Interest Group Proceedings, Fayetteville, Arkansas, April 26-29, 2011. Scientific Investigations Report 2011-5031. 160-167.

Hao, X.-M., Chen, Y.-N., Li, W.-H. (2009). *Impact of Anthropogenic Activities on the Hydrologic Characters of the Mainstream of the Tarim River in Xinjiang during the Past 50 Years.* Environmental Geology 57 (2): 435. DOI: 10.1007/s00254-008-1314-0.

Helsel, D.R. and R. M. Hirsch, (2002). *Statistical Methods in Water Resources*. Techniques of Water Resources Investigations, Book 4, chapter A3. U.S. Geological Survey. 522 pages.

Johnston, C.A. and Shmagin, B.A. (2008). *Regionalization, Seasonality, and Trends of Streamflow in the US Great Lakes Basin*. Journal of Hydrology 362: 62-88.

Kahya, E. and Kalayci, S. (2004). *Trend Analysis of Streamflow in Turkey*. Journal of Hydrology 289: 128-144.

Kumar S., Merwade, V., Kam, J., Thurner, K. (2009). *Streamflow Trends in Indiana: Effects of Long Term Persistence, Precipitation and Subsurface Drains.* Journal of Hydrology 374: 171-183.

Li, L.-J., Wang, H., Wang, J., Yang, J.-W., Jiang, D.-J., Li, J.-Y., Qin, D.-Y. (2007). *Assessing the Impact of Climate Variability and Human Activities on Streamflow from the Wuding River Basin in China.* Hydrological Processes 21: 3485-3491. DOI: 10.10052/hyp.6485.

Lins, H. (2005). *Streamflow Trends in the United States*. U.S. Geological Survey Fact Sheet 2005-3017. http://pubs.usgs.gov/fs/2005/3017/pdf/FS2005_3017.pdf.

Milly P.C.D. and Dunne K.A., (2002). *Macroscale Water Fluxes 2. Water and Energy Supply Control of Their Inter-annual Variability.* Water Resources Research 38(10): 1206. DOI: 10.1029/2001WR000760.

Milly P.,C. D. and Wetherald, R.T. (2002). *Macroscale Water Fluxes 3. Effects of Land Processes on Variability of Monthly River Discharge.* Water Resources Research 38(11): 1235. DOI: 10.1029/2001WR000761.

Perry, R. (1995). *Regional Assessment of Land Use Loading to Unconfined Aquifers*. A dissertation submitted to the Department of Civil and Environmental Engineering, University of South Florida.

Peterson, H.M., Nieber, J.L., Kanivetsky, R. (2011). *Hydrologic Regionalization to Assess Anthropogenic Changes.* Journal of Hydrology 408: 212-225.

Rossi, A., Massei, N., Laignel, B., Sebag D., Copard, Y. (2009). *The Response of the Mississippi River to Climate Fluctuations and Reservoir Construction as Indicated by Wavelet Analysis of Streamflow and Suspended-sediment Load, 1950-1975*. Journal of Hydrology 377: 237-244.

Rumenik, R., and Grubbs, J. (1996). *Low-Flow Characteristics of Florida Streams.* Water-Resources Investigations Report 93-4165, Prepared in cooperation with the Florida Department of Environmental Protection, Tallahassee, Florida.

Schneider, J.W., Upchurch, S., Chen, J., and Cain, C. (2008). *Simulation of Groundwater Flow in North Florida and South-Central Georgia*. Prepared for the Suwannee River Water Management District.

Slack, J.R. and Landwehr, J.M. (1992). *Hydro-climatic Data Network: A U.S. Geological Survey Streamflow Data Set for the United States for the Study of Climate Variations, 1874-1988*. U.S. Geological Survey Open-File Report 92-129.

Tao, H., Gemmer, M., Bai, Y., Su, B.., Mao, W. (2011). *Trends of Streamflow in the Tarim River Basin during the Past 50 years: Human Impact or Climate Change?* Journal of Hydrology 400: 1-9.

Tarboton D.G. (1994). *The Source Hydrology of Severe Sustained Drought in the Southwestern United States.* Journal of Hydrology 161:31-39.

Tomer, M.D. and Schilling, K.E. (2009). *A Simple Approach to Distinguish Land-use and Climate-change Effects on Watershed Hydrology*. Journal of Hydrology 376: 24-33.

Wang, D. and Cai, X. (2009). *Detecting Human Interferences to Low Flows through Base Flow Recession Analysis*. Water Resources Research 45, W07426: 1-12. DOI: 10.1029/2009WR007819.

Ye, B., Yang, D., and Kane D.L. (2003). *Changes in Lena River Streamflow Hydrology: Human Impacts versus Natural Variations*. Water Resources Research 39 (7), 1200. DOI: 10.1029/2003WR001991.

Yu, S., and Pilon, P. (2004). *A Comparison of the Power of the t Test, Mann-Kendall and Bootstrap Tests for Trend Detection*. Hydrological Sciences Journal, 41(1), 21-37.

Zhang, Y.-K., Schilling, K.E. (2006). *Increasing Streamflow and Baseflow in Mississippi River since the 1940s: Effect of Land Use Change*. Journal of Hydrology 324: 412-422.