INFLOW DATASET DEVELOPMENT: STATISTICAL METHODOLOGY

PARR HYDROELECTRIC PROJECT

FERC NO. 1894

Prepared for:

South Carolina Electric & Gas Co.
Columbia, South Carolina

Prepared by:

Kleinschmidt

Lexington, South Carolina
www.KleinschmidtGroup.com

August 2014

INFLOW DATASET DEVELOPMENT: STATISTICAL METHODOLOGY

PARR HYDROELECTRIC PROJECT

FERC NO. 1894

Prepared for:

South Carolina Electric & Gas Co.
Columbia, South Carolina

Prepared by:

Kleinschmidt

Lexington, South Carolina
www.KleinschmidtGroup.com

August 2014

INFLOW DATASET DEVELOPMENT:
STATISTICAL METHODOLOGY

PARR HYDROELECTRIC PROJECT

FERC NO. 1894

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INFLOW DATASET DEVELOPMENT:
STATISTICAL METHODOLOGY

PARR HYDROELECTRIC PROJECT

FERC NO. 1894

1.0 PARR RESERVOIR INFLOW DATA DEVELOPMENT

1.1 INTRODUCTION

An inflow hydrology dataset is being developed in support of developing operations models and to satisfy the Final Parr Fairfield Operations Model Study Plan (Study Plan). As discussed in the Study Plan, the operation of the pumped storage development and lack of long-term operational records prevents the back-calculation of a sufficient inflow dataset. For this reason, the inflow to Parr Reservoir was calculated using upstream flow data adjusted by statistically-derived parameters. The inflow time series datasets for Parr Reservoir were developed using statistical algorithms based on flow data records from the USGS gages upstream and downstream of the Parr Dam.

The inflow dataset developed by this process will be used for two distinctly different simulation processes. The utilization of Parr Reservoir inflows for power generation by the Fairfield Pumped Storage development and the Parr Hydro development, and corresponding upper and lower reservoir fluctuations will be simulated using the USACE modeling package HEC-ResSim; this software's primary requirement is daily inflow values. The flows released from the Parr development will be used as upstream boundary conditions in the USACE model HEC-RAS, which will simulate the downstream flow and stage regimes. The HEC-RAS model requires flow values in increments of one-hour or less.

1.2 HYDROLOGIC DATA

Data used in the statistical analyses were obtained via the USGS web portal (http://waterdata.usgs.gov/nwis). The data were processed using spreadsheets and the USACE database program HEC-DSSVue. The USGS gage sites used in the analysis are listed in Table 1. Additional flow and stage data were obtained from the USGS server for use in other phases of this study, and will be fully cited and described in the applicable summary reports.

TABLE 1 USGS GAGE SITES

DATA SOURCE	USGS #	DRAINAGE AREA (SQ. MI.)	PERIOD OF RECORD	DATA TYPE
Enoree River at Whitmire	02160700	444	10-1-1973 to Current	Stage & Discharge
Enoree River near Woodruff	02160390	249	2-9-1993 to Current	Stage & Discharge
Tyger River near Delta	02160105	759	10-1-1973 to Current	Stage & Discharge
Broad River near Carlisle	02156500	2790	10-1-1938 to Current	Stage & Discharge
Broad River at Alston	02161000	4790	10-1-1896 to 12-1-1907, 10-1-1980 to Current	Stage & Discharge

1.3 PARR RESERVOIR INFLOW DATA SYNTHESIS

Prior to the statistical analyses, Kleinschmidt Associates performed a review of relevant hydrologic studies published by the USGS. These included:

- Low-Flow Frequency and Flow Duration of Selected South Carolina Streams in the Broad River Basin through 2008 (USGS Open-File Report 2010-1305);
- Magnitude and Frequency of Rural Floods in the Southeastern United States, 2006: Volume 3, South Carolina (USGS Scientific Investigations Report 2009-5156); and
- Techniques for Estimating the Magnitude and Frequency of Floods in Rural Basins of South Carolina, 1999 (Water-Resources Investigations Report 02-4140)

Although these studies included hydrologic analyses of the Parr watershed, their focus was primarily on the development of statistically-based estimates of extreme events as opposed to typical hydrology. These studies were reviewed as background information regarding the physiographic nature of the watershed, which could provide insight on the hydrologic behavior of the Broad River and its tributaries upstream and downstream of Parr Reservoir.

The synthesis of streamflow data using a proration of upstream gages typically uses a statistical regression technique based on drainage area ratios. Gages were selected for summing prorated inflows with the intention of maximizing the relevant, overlapping periods of record, as well as drainage area coverage. Periods of record that are relevant represent the current development of the waterway, which would be subsequent to the commissioning of the pumped storage project (December 1978) to current day. Three gages were selected that measure contributing flows for 84% of the project's total drainage area and compared with the corresponding period of record with the Alston gage downstream of the Parr dam[1].

In order to develop the inflow data set for Parr Reservoir, various statistical methods were assessed to determine the optimal estimate. These methods included statistical regressions to determine the weighting factors for scaling the measured upstream flows (see Figure 1) to estimate the inflow to Parr Reservoir. These methods are described in the following sections.

The statistical analyses will use monthly and annual flow data rather than daily average flows. The daily data are affected by reservoir operations, which introduce a significant degree of variability due to the cyclic transfer of up to 29,000 acre-feet between the upper and lower reservoirs. Flow releases from the project may be vastly different at any given hour from the inflows to the Parr reservoir. The monthly and annual flow data statistics are much less affected by day-to-day operations.

[1] It is worth noting that the Parr dam drainage area is 4,750 square miles compared to the slightly larger Alston gage drainage area of 4,790 square miles (about 0.8% less). However, the USGS cites the Alston gage as synonymous with reservoir outflow. No adjustment was made, as the difference is statistically insignificant.

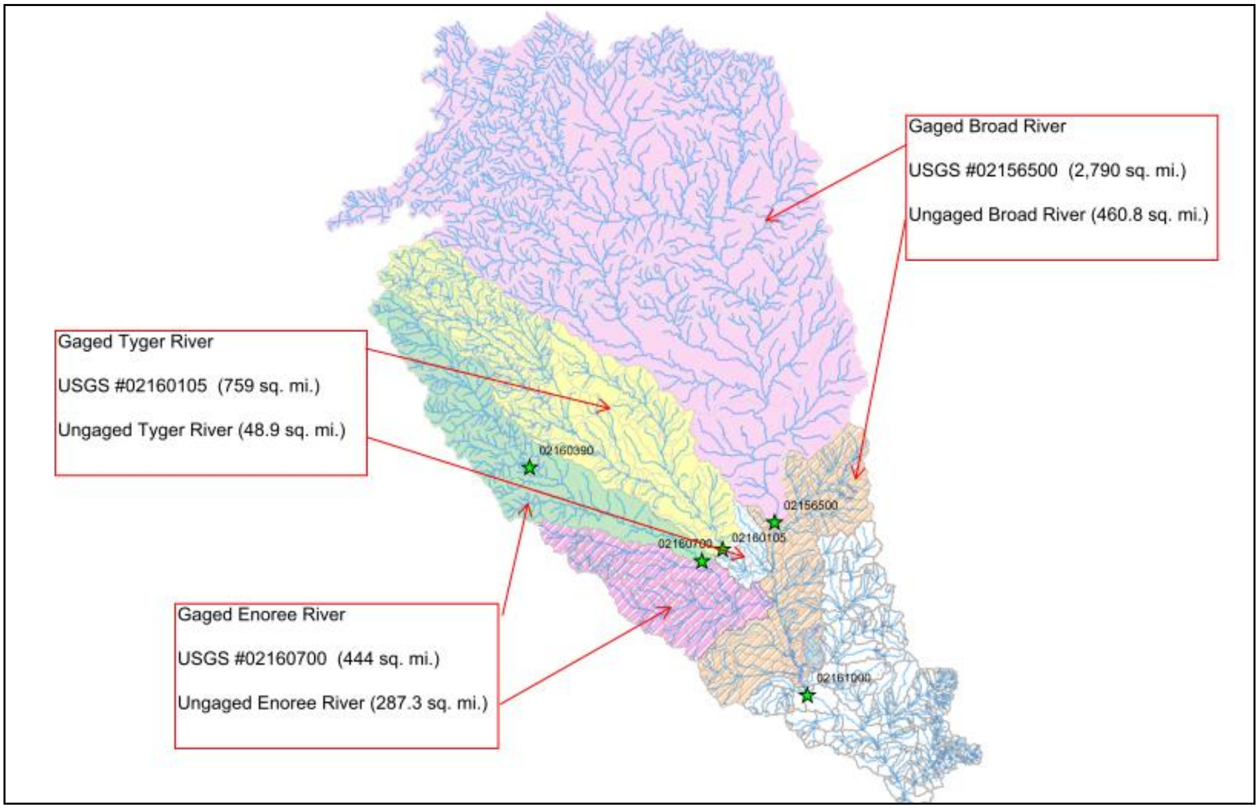


FIGURE 1 GAGED AND UNGAGED BROAD RIVER SUBWATERSHEDS

1.3.1 Preliminary Hydrologic Review

Prior to the statistical regression analyses, a cursory review was performed to assess the hydrologic response of the subwatersheds that contribute to the Parr Reservoir inflows. The review consisted of a comparison of a sampling of monthly average flows from the upstream gages on the Broad, Tyger, and Enoree rivers to the flows at the Alston gage (see Figure 2). The purpose of the review was to determine the degree of hydrologic similarity between the three contributing subwatersheds. A high degree of hydrologic similarity indicates that the soils, topography, and land use over the entire watershed are homogeneous. The subsequent analyses, which are predicated on this assumed homogeneity, provide a basis for developing a statistical relationship between the gaged and ungaged portions of the subwatersheds.

The first comparison was the unadjusted monthly average flows from the upstream gages with the Alston gage. This comparison illustrates the relative contribution of the upstream gaged areas. For the given period, the monthly average flow at Carlisle was approximately 2/3 of the flow average at Alston.

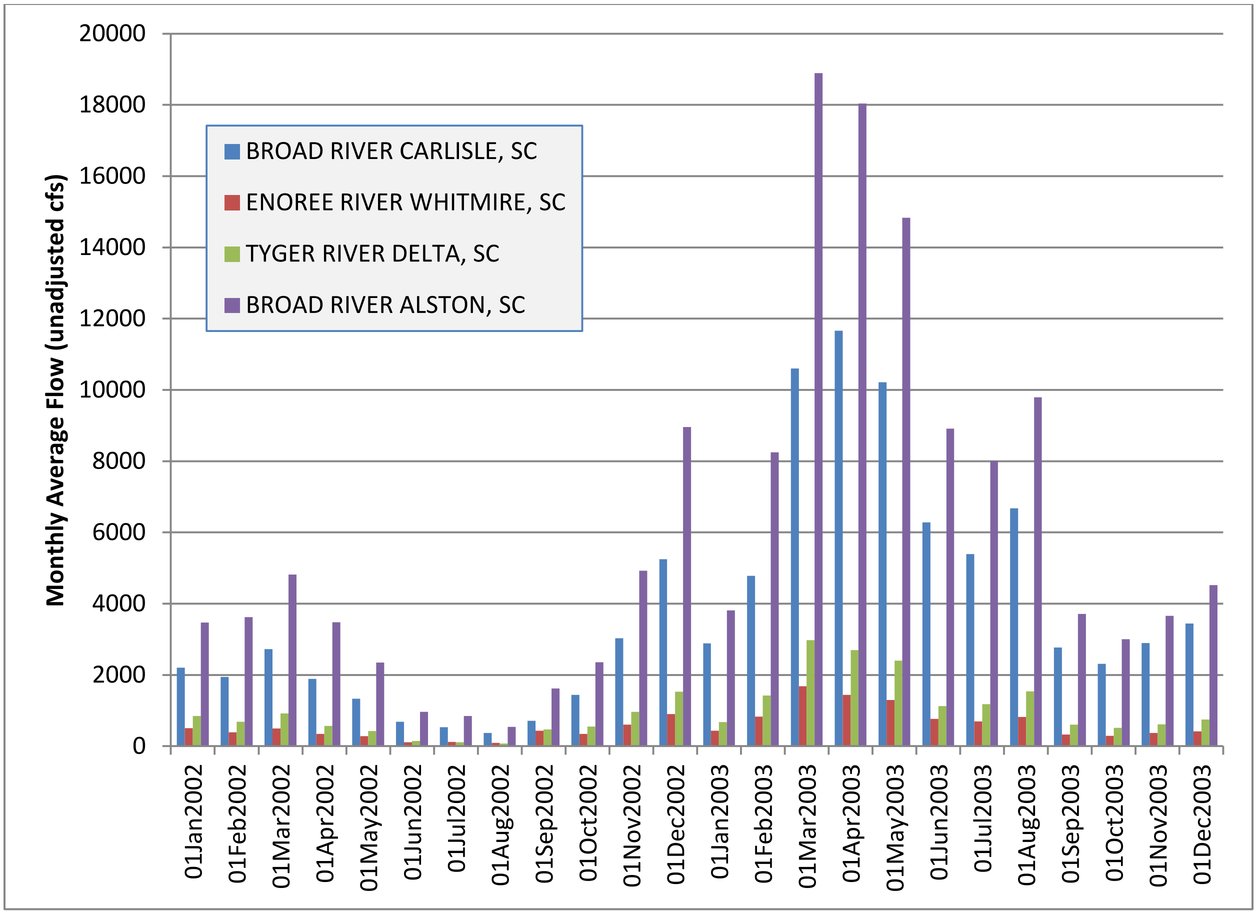


FIGURE 2 MONTHLY AVERAGE FLOWS, UNADJUSTED

The second portion of the review was a comparison of the runoff from the gaged upstream subwatersheds. The monthly average flows from the previous step were normalized by drainage area, resulting in the average flow per 100 square miles of drainage area. This comparison was performed to determine the similarity in runoff characteristics between the three gaged areas. The comparison (see Figure 3) illustrates that the range of the monthly averages (per 100 sq. mi.) was visually close to the aggregate average through a variety of flow ranges; this indicates the hydrologic similarity of the three subbasins.

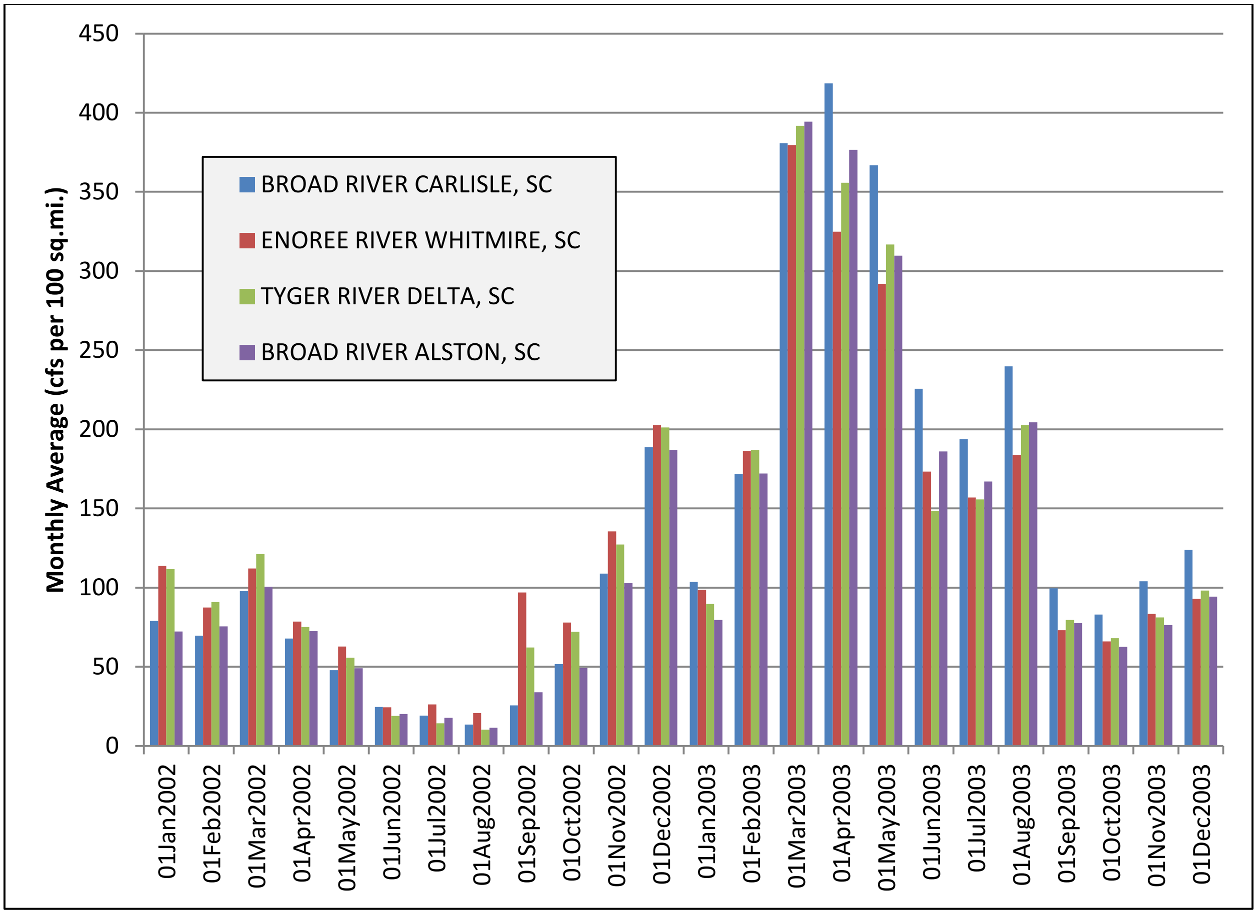


FIGURE 3 NORMALIZED MONTHLY AVERAGE FLOWS

1.3.2 MULTIVARIATE REGRESSION ANALYSIS

A multivariate regression was performed to determine the parameters of a generalized equation for estimating the inflow to Parr Reservoir. The flow estimate is based on the flows measured at three gage sites upstream of the impoundment. The two parameters include a fitted regional exponent (γ), and a fitted regional coefficient (α). The equation, shown below, is a summation of the three upstream flow values multiplied by scaling factors, which include the ratio of the total drainage area represented by each to that gage's actual drainage area.

Equation 1: $ParrInflow = \langle \alpha * BRC \left(\frac{3250.8}{2790}\right)^{\gamma} \rangle + \langle \alpha * TRD \left(\frac{807.9}{759}\right)^{\gamma} \rangle + \langle \alpha * ERW \left(\frac{731.3}{444}\right)^{\gamma} \rangle$

where,

BRC – Broad River at Carlisle
TRD – Tyger River near Delta
ERW – Enoree River at Whitmire
α – Fitted Regional Coefficient
γ – Fitted Regional Exponent

The regional exponent was developed by quantifying the relationship between monthly streamflow averages and drainage area using two unregulated stream gages on the same river with overlapping records. The only gages that meet this in the immediate Parr Dam watershed are on the Enoree River. The regional exponent was developed by performing a regression on monthly flow averages from the Woodruff gage (drainage area = 249 sq. mi.) and the Whitmire gage (drainage area = 444 sq. mi.). These two gages were selected because they have the longest overlapping (current) periods of record. The result of this regression produced the drainage area regional exponent (γ) of 0.599.

This proration exponent was used to normalize the monthly flow averages, prior to performing the second regression to develop the drainage area coefficient (α). The regression used monthly flow averages for the period 1/1/1981 through 12/31/2013, a total of 396 months. The target data used in the regression is the monthly average flow at the Alston gage, which was adjusted by adding the estimated evaporation from both the Monticello and Parr reservoirs. Evaporation estimates were based upon monthly losses in inches[2] applied to the average surface area of both reservoirs, plus estimated increased evaporation caused by the V.C. Summer Nuclear Station thermal plume in Monticello Reservoir. This adjustment ranged in value from 37.5 cfs in January to 103.5 cfs for July.

The results of this regression, using all 396 months, produced a value of $\alpha = 1.041$, an R^2 of 0.9828, and a standard error of 495.4. The scatter plot of Alston monthly flow vs. predicted flow, including a 1:1 reference line, is shown in Figure 4. The modeling residuals were also calculated and are shown graphically in Figure 5. The modeling residual values are the difference between the target value and the predicted value. In this case, a negative modeling residual indicates that the predicted value is greater than the target value. The plot of the modeling residuals indicates that the statistical model tends to overpredict flows during months for which the average flow was less than 7,700 cfs (the y-intercept shown on Figure 5) and tends to underpredict during months with flow averages greater than 7,700 cfs.

[2] Evaporative rates from "Pan Evaporation Records for the South Carolina Area," John C. Purvis, SC State Climatology Office, with FWS evaporation taken as 75% based on Discussions in "NOAA Technical Report NWS 33: Evaporation Atlas for the 48 Contiguous States," June 1982.

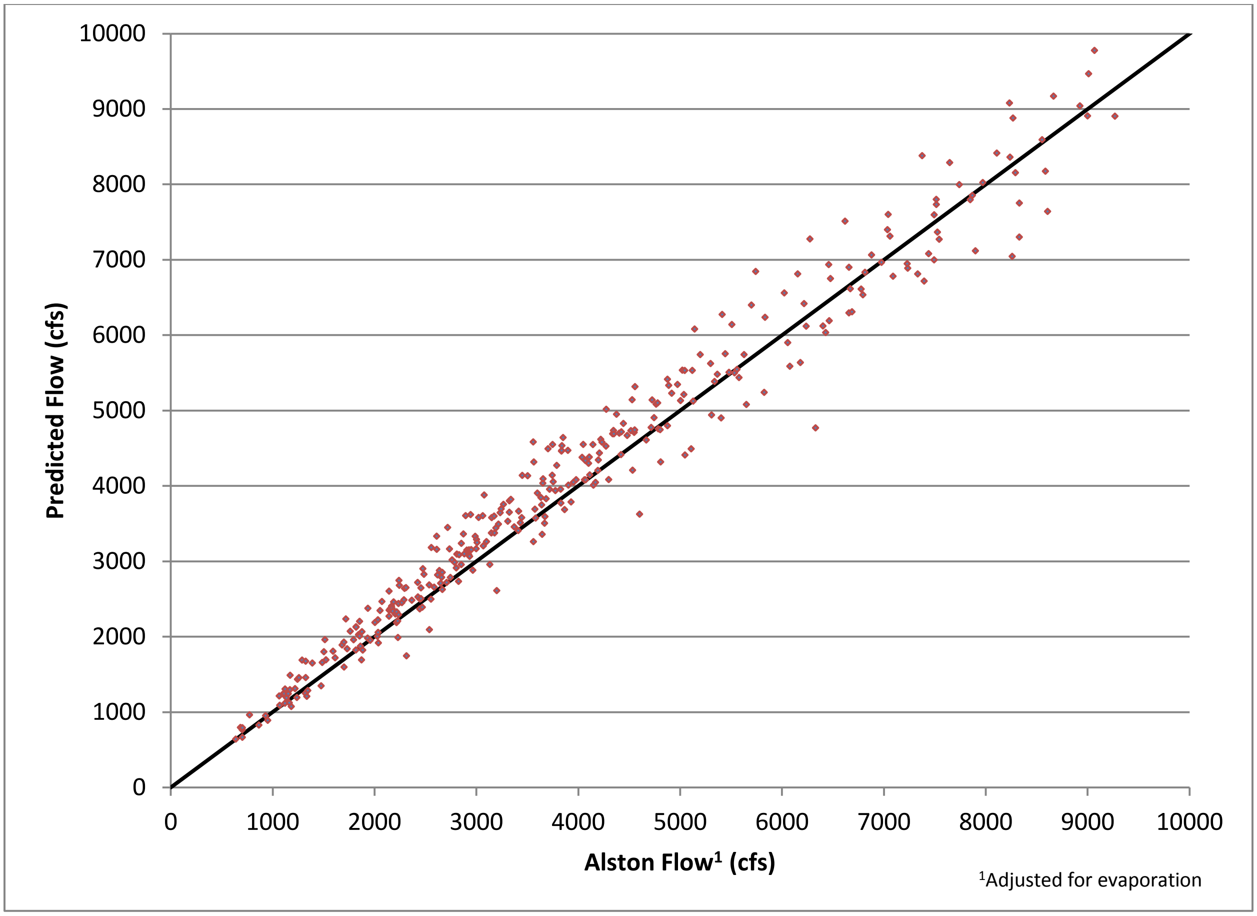


FIGURE 4 **ALSTON FLOW VS. PREDICTED MONTHLY AVERAGES (33 YEARS) – REGRESSION BASED ON ALL MONTHS**

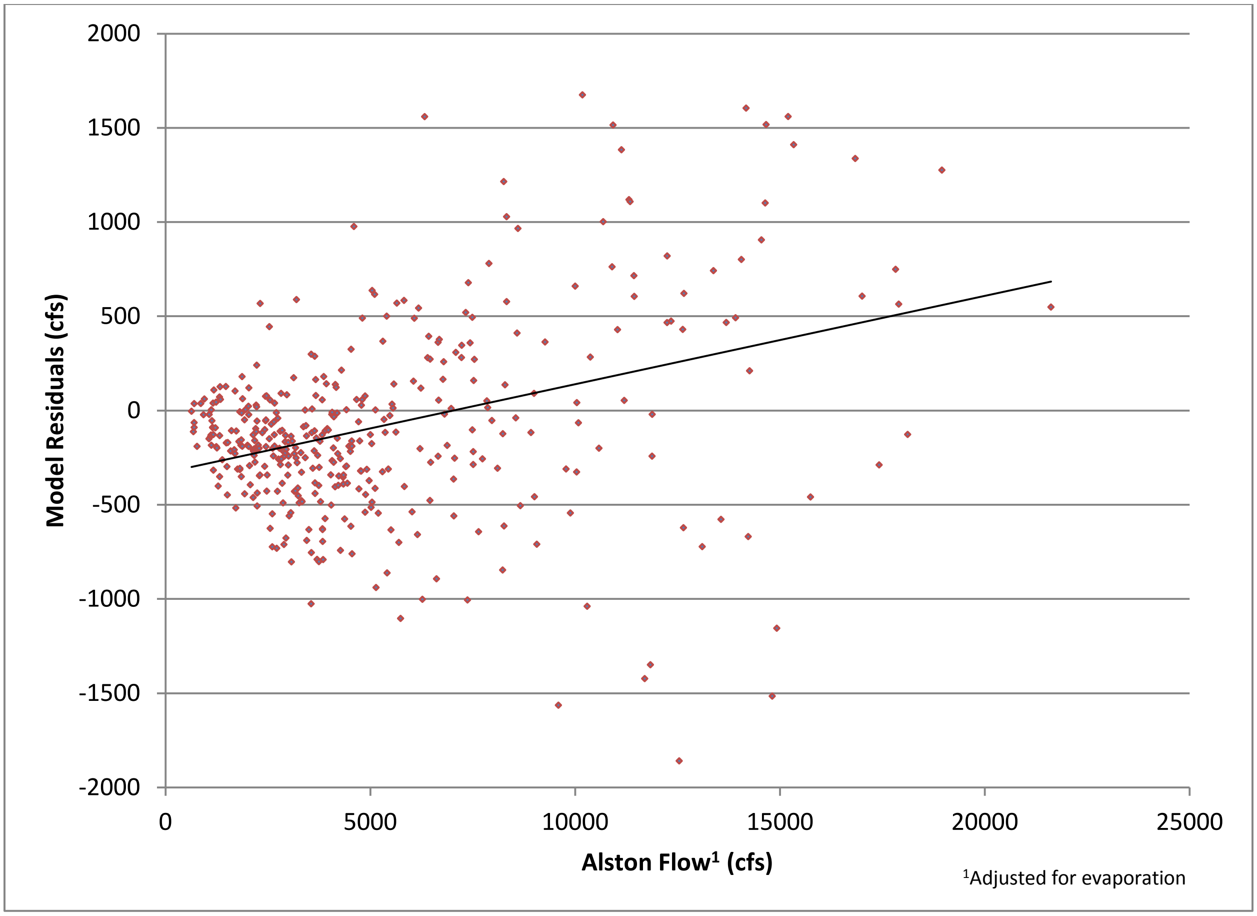


FIGURE 5 MODEL RESIDUALS – REGRESSION BASED ON CONCURRENT PERIOD OF RECORD

1.3.3 MODIFIED REGRESSION (ADJUSTED FLOW RANGE)

Due to the results of the first regression attempt, which indicated a tendency to overpredict during months with less than 7,700 cfs average flow, a second regression was developed. Because balancing the hydrologic resource is imperative during lower inflow conditions, this modified regression was performed to more accurately predict flows in the lower range. The second analysis used the lowest 75% of monthly average flows (289 out of 396 months) as a basis for the regression and then applied the resulting coefficients on the entire dataset to quantify the statistical performance.

The results of the second regression, using 289 of the 396 months, produced a value of α = 0.988, an R^2 of 0.9828, and a standard error of 469.6. Compared to the first regression, the reduced α-value did not change the R^2 value, but reduced the standard error. The most significant change was the modeling residuals. The y-intercept for the residual plot for the second regression is approximately 3,900 cfs. This indicates that the second regression has a lower statistical bias in the range of the most typical flows than the first regression. The scatter plot of Alston monthly flow vs. predicted flow is shown in Figure 6, and the modeling residuals are shown in Figure7.

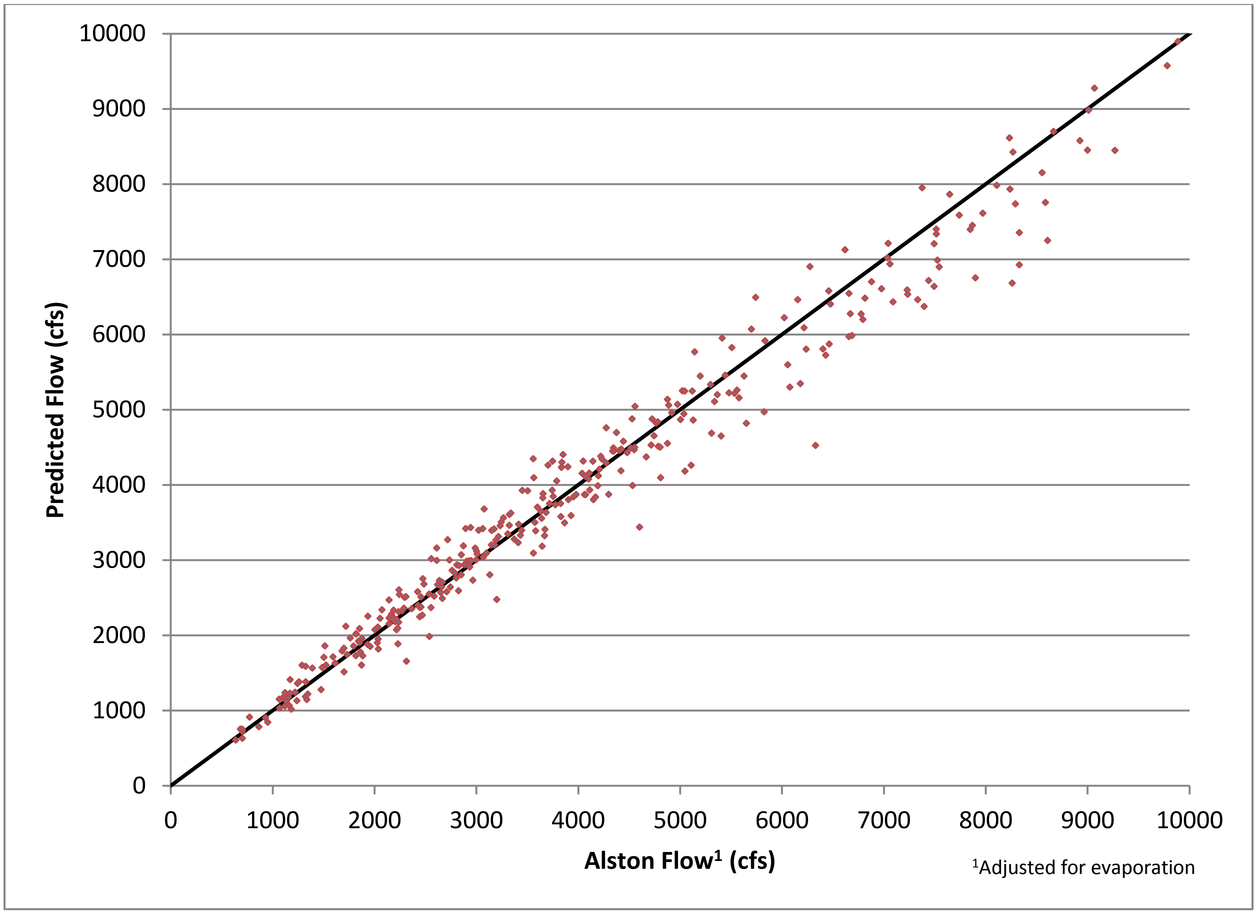


FIGURE 6 ALSTON FLOW (ADJUSTED) VS. PREDICTED MONTHLY AVERAGES (33 YEARS) - REGRESSION BASED ON DRIEST 75% MONTHS

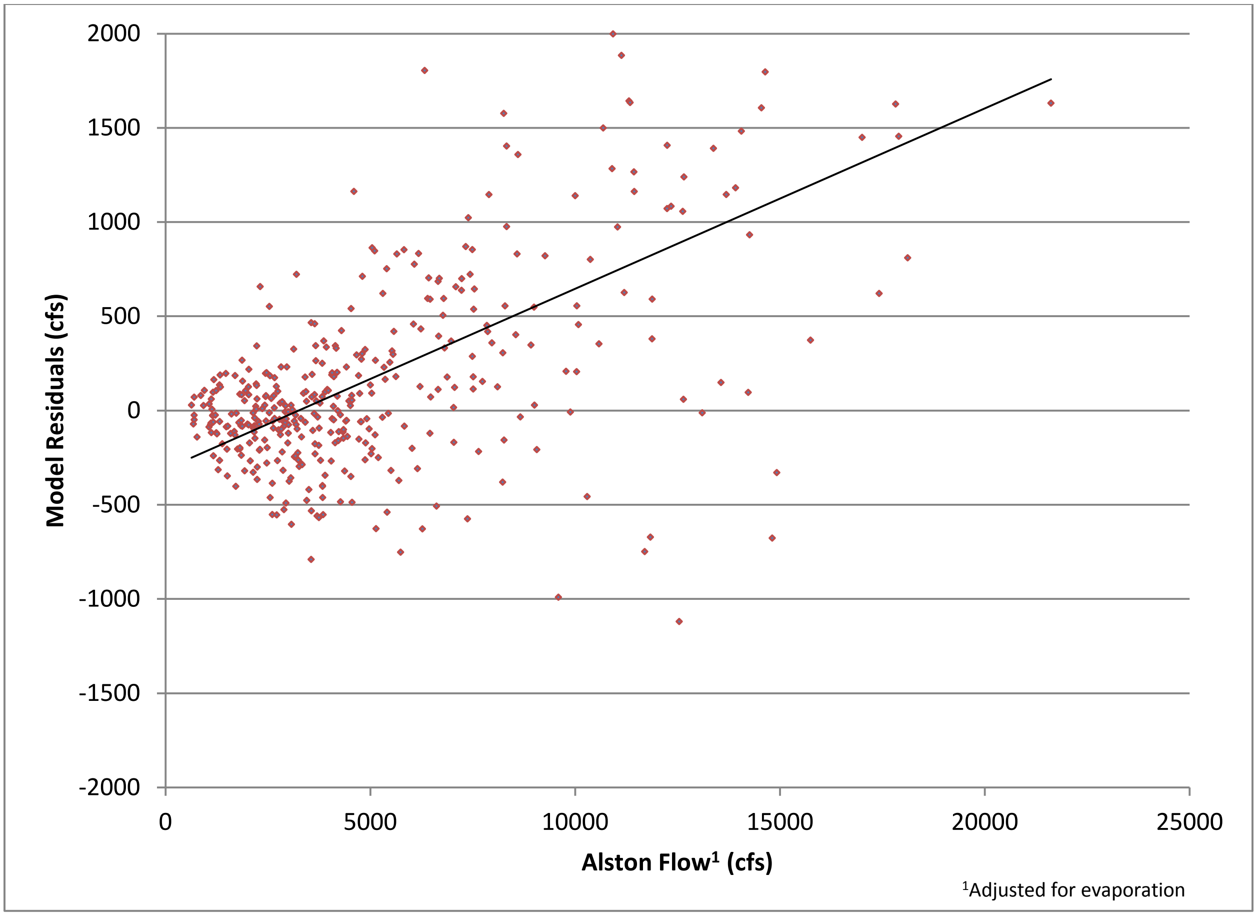


FIGURE 7 **MODEL RESIDUALS - REGRESSION BASED ON 75% LOWEST FLOW AVERAGE MONTHS**

1.3.4 Model Verification

The verification of the model results was performed by comparing the predicted flows vs. the target flows for three year periods, including statistically wet and dry periods (see Figures 8 and 9). The dry period was from January 2006 to December 2008, inclusive. The wet period was from January 1993 to December 1995, inclusive. These periods were selected on the basis of the average flow of the three years and of the 33-year period for which there was a complete flow dataset for the gages, which spanned January 1981 to December 2013.

These comparisons indicate that the estimated values have a slight overprediction bias during prolonged low-flow periods. During higher flow periods, such as 1993 - 1995, there is very little bias on the lower flows and a slight underprediction bias on the higher flows.

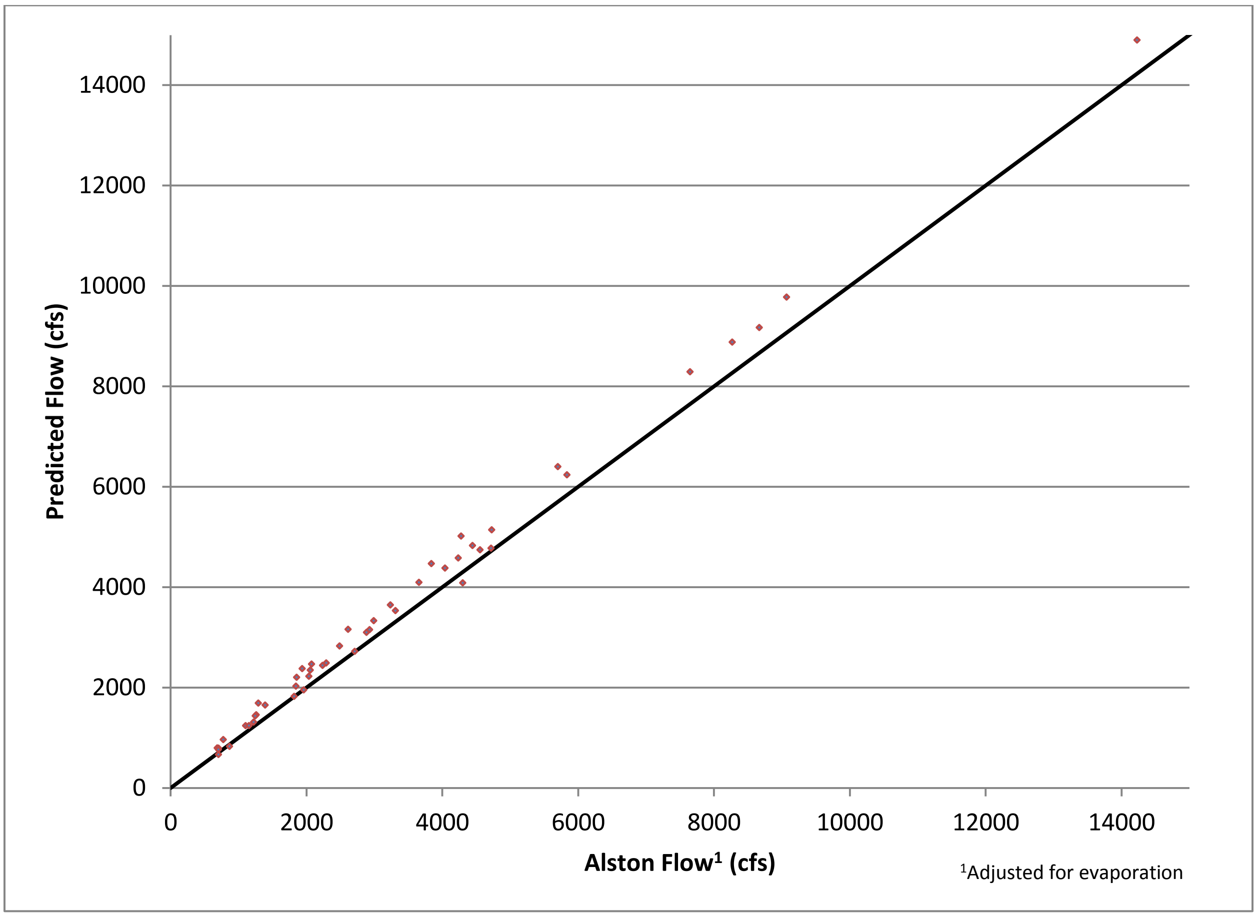


Figure 8 Alston Flow (adjusted) vs. Predicted Monthly Averages (dry 3-year period) - Regression Based on Driest 75% Months

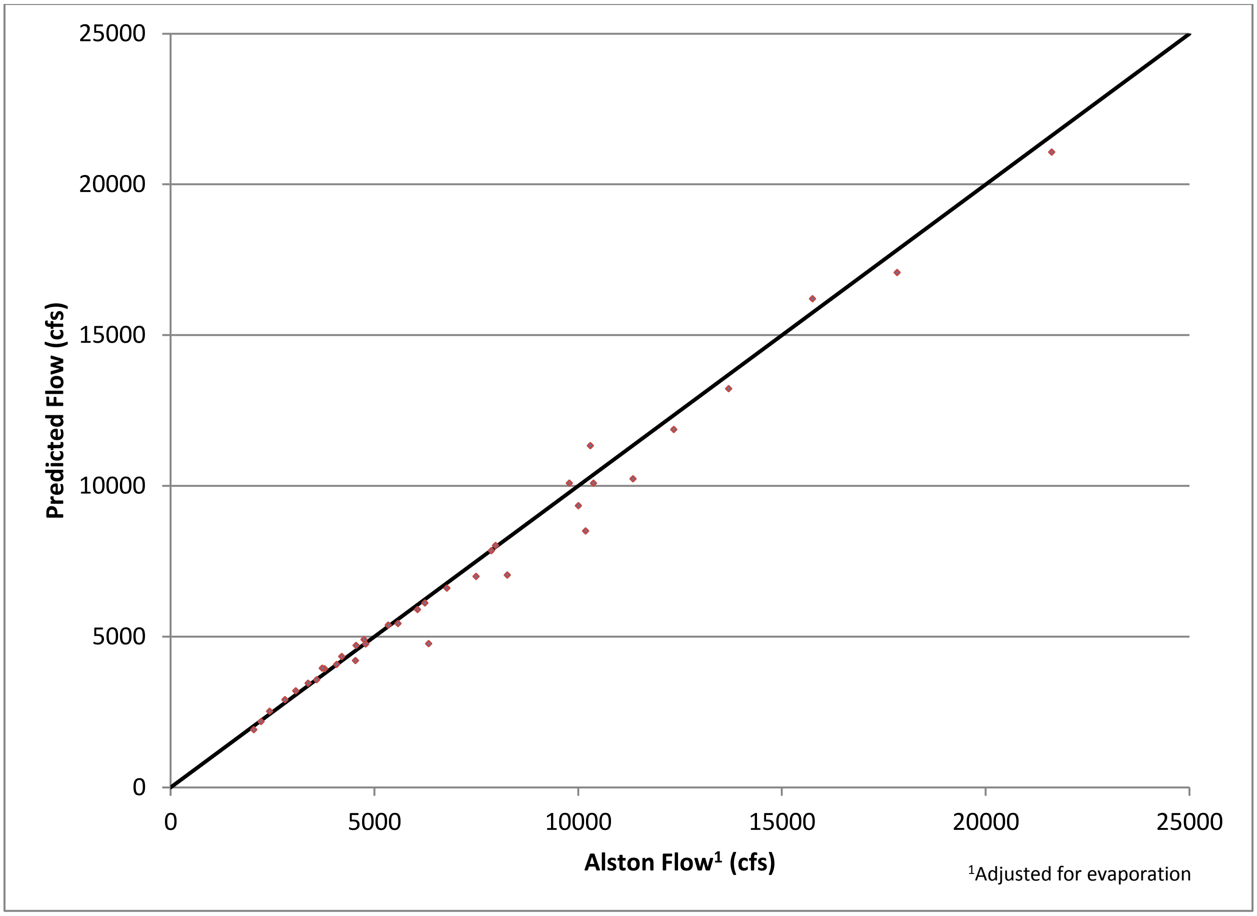


FIGURE 9 ALSTON FLOW (ADJUSTED) VS. PREDICTED MONTHLY AVERAGES (WET 3-YEAR PERIOD) - REGRESSION BASED ON DRIEST 75% MONTHS

1.4 SUMMARY

Two statistical regressions were performed to develop the coefficients used in Equation 1 (see Section 1.3.2). The first regression, using all of the monthly flow averages, resulted in a trend of negative modeling residuals (overprediction) for months with flow averages less than 7,700 cfs. A subsequent regression, using monthly flow averages less than 6,000 cfs (approximately 75% of the data values) produced a better balance between negative and positive modeling residuals. This regression performed statistically better in the range of the most frequent values of monthly average flows, with flows nearest 3,900 cfs predicted most accurately. As this lower flow range is of greater importance than the entire historic range for balancing the hydrologic resource, the coefficient and exponent determined through the second regression are preferred for the development of the inflow dataset (see Table 2).

TABLE 2 STATISTICAL MODEL RESULTS SUMMARY

MODEL NAME	REGRESSION DATASET OF ALL MONTHLY AVERAGES (396 VALUES)	REGRESSION DATASET OF LOWEST 75% MONTHLY AVERAGES (289 VALUES)
α – Coefficient	1.041	0.988
γ – Exponent	0.599	0.599
Standard Error	495.0	469.6
R^2	0.9828	0.9828

APPENDIX A

INFLOW DATASET MEMO: OPERATIONS RCG QUESTIONS AND ANSWERS

Scott Harder

Hydrologist, LWC Division, SCDNR

5/30/14

Comments regarding Kleinschmidt's "Inflow Dataset Development: Statistical Methodology" for the Parr Hydroelectric project (FERC No. 1894).

1. The methodology pertaining to how the monthly statistical analysis will used to develop daily (or hourly) Parr inflow dataset needs to be clarified in the report. Also, will time of travel be factored in when moving to a daily or hourly time step?

We propose to edit the report during the meeting so the clarifications are agreed to and understood by the RCG. Preliminary clarification follows: The statistical analyses were performed on data points that were monthly average flow values for each of the gages, for the common gaged periods of record (1981 – 2013). The regional coefficients derived from these analyses will be applied to recorded data for each of the three upstream gages. The resulting sum of these inflows will serve as the dataset input to the HEC reservoir and downstream river models. The reservoir and downstream models will use hourly (or longer) time steps for evaluating operations. The downstream river model will include travel time on an hourly basis.

Hourly inflows can use mean daily data as a substitution, or they can be calculated from hourly gage data. If done on an hourly basis, the flows will be routed from the upstream gages using one of several routing algorithms (such as Muskingum, Muskingum-Cunge and Modified Puls), the selection of which will be based on the stream hydraulics. The routing of hourly data would include travel time, whereas mean daily data would not be adjusted for travel time because the gages are only hours away from the project.

Hourly inflows are not expected to have noticeable affects on the project model runs due to the magnitude of the usable storage, except during high inflow hydrographs. The RCG should consider the benefit of developing hourly inflow data versus capturing a longer period of record with daily data. If the daily data is used, hourly model runs will assume the mean daily inflow is occurring for that 24-hour period. If the hourly data is used, the gages are limited to October 1, 1987; daily data is available back to October 1, 1980 (although monthly values used to determine the regional coefficients were truncated for complete calendar years, 1981-2013).

2. Regarding the technique to compare the hydrologic similarity between the three gages area (Tyger, Enoree and Broad in section 1.3.1:

a. Only two years were used for comparison (2002 and 2003) in Figure 3. Was there an attempt to include more years? These two years represent extremes, or close to it, for dry and wet years back to back and the comparison would be more robust if it included more normal periods as well or if a comparison was made for a longer period of time (see below also).

The comparison of normalized flows for evaluating hydrologic similarity was performed using the monthly average flows for the period 1/1/1981 to 12/31/2013, a thirty-two year period. Only two years were charted for the document for visibility, selected to illustrate consistent gaged contributions across a range of hydrologic conditions: extreme drought conditions during the summer of 2002, and high inflows the following spring. We can present additional years for comparison, and propose to include them in appendices. Our conclusions apply to the entire period of record and range of flows.

The statistical regressions were performed using several variations of inflow subsets including the entire 32-year period, as well as using an abridged dataset that included only the lowest 75% of the flow values. The abridged version used an equivalent of 24 years of monthly average flows.

b. Please rewrite or elaborate on the following statement at the end of page 6: "The comparison (see Figure 3) illustrates that the range of the monthly averages (per 100 sq. mi.) was visually close to the aggregate average through a variety of flow ranges; this indicates the hydrologic similarity of the three subbasins." Please consider summarizing the point you are trying to make here quantitatively in a table and not just visually from a plot. In Figure 3, normalized monthly average runoff is consistently higher for the Broad basin in 2003 than for the Tyger and Enoree, which maybe isn't surprising given that the Broad is a much larger basin that extends up into the North Carolina mountains. It would be instructive to see if this was observed for other years besides 2003 (my own preliminary analysis shows that it does). The higher runoff suggests that the assumption of homogeneity for the gaged portion of Broad basin (as a whole) at Carlisle as compared to the Enoree and Tyger basins may not be valid. As a result, it may be problematic to use the Broad River gage at Carlisle to develop a regional coefficient. However, I think that the assumption that the *ungaged* parts of the three basins (Tyger, Enoree, and Broad) are very nearly homogeneous is likely valid, but the question remains on how to best account for the additional flow from these ungaged areas (but see 4 below).

Visual examination of the normalized flows was done to check for consistent, significant discrepancies between gaged areas under a range of hydrologic conditions. The comparison of any single normalized gage with the aggregate average was visibly within the same order of magnitude for all months across a large range of inflow conditions, and was the basis for concluding the similarity. The Carlisle gage does appear to contribute more flow more often, but to a nominal degree compared to the aggregate. In the interest of simplicity, consistent regional coefficients were used for the analysis.

The desired end product is a dataset that consists of six time series of flow data, three of which are USGS flow records measured at the gage sites for the three rivers, and the other three time series are estimates of ungaged flows from the three rivers. Several statistical models were evaluated in an attempt to determine the most effective regression, using statistical metrics such as r-square and standard error values. The selected statistical model produced r-squared values above 95%, suggesting a strong correlation using consistent fitted regional coefficients.

Although not documented in the report, the initial screening of statistical models included many variations of regressions that were attempted in order to determine if the ungaged flows appeared to be more similar to one or two of the upstream gages as opposed to all three. A regression model was evaluated, using 1) all data, 2) three consecutive dry years, and 3) three consecutive wet years. This regression model included alpha values for each of the streamgages. The statistical regression results indicated that the ungaged flows were more similar to the Tyger River than the Broad or Enoree, but the relationship shifted between wet and dry periods. The statistical model used in this initial screening was dropped from consideration and not documented in the report.

3. In section 1.3.2, please make sure that the x and y axes scales are set to display all data points in Figures 4 and 5. For example, in figure 4, average flows at Alston extend well beyond 10,000 cfs for some months, but the maximum flow is cutoff somewhere between 9000-9500 cfs.

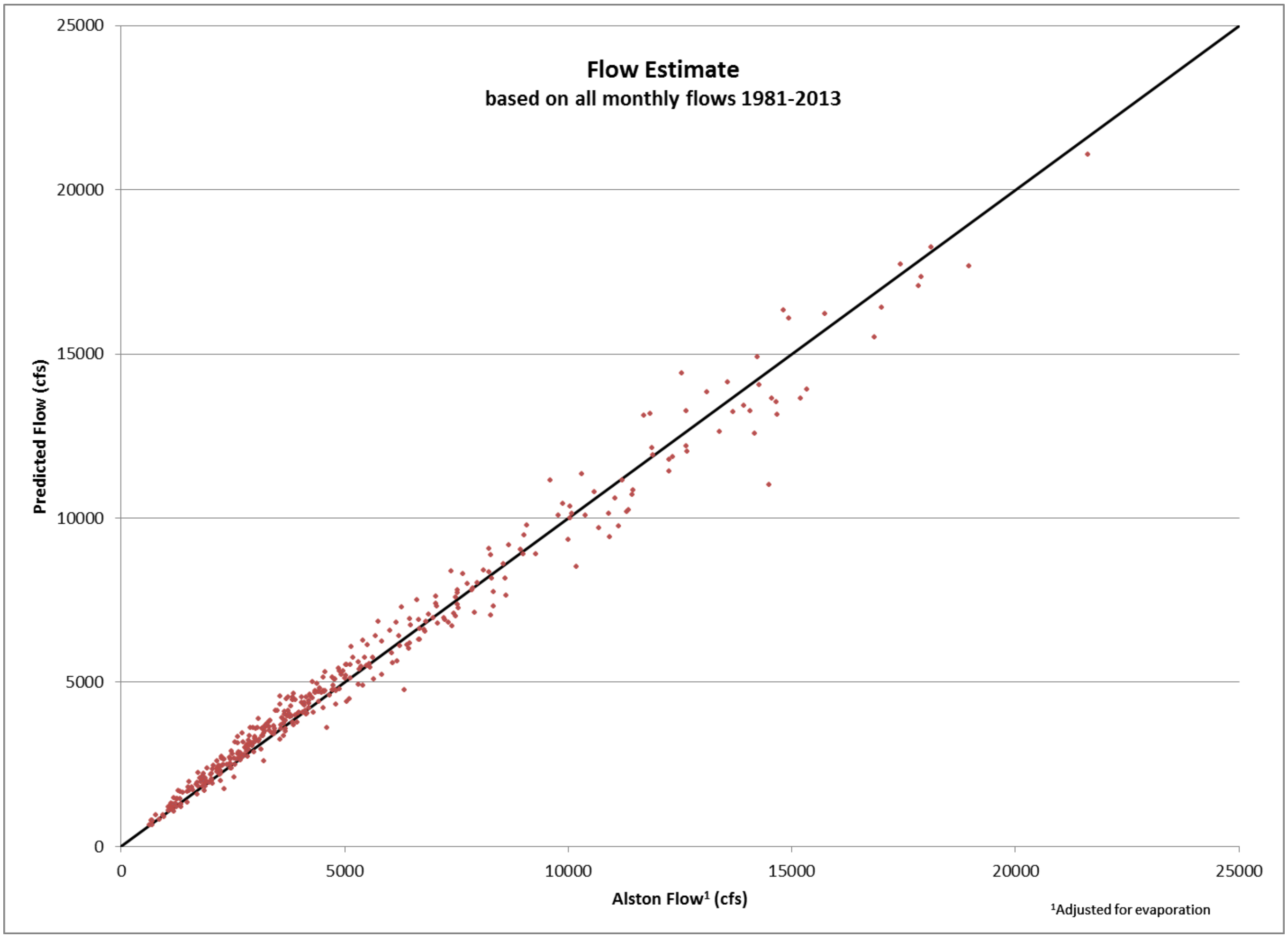


FIGURE 1 (EXPANDED) ALSTON FLOW VS. PREDICTED MONTHLY AVERAGES (33 YEARS) – REGRESSION BASED ON ALL MONTHS

4. I initially had some strong reservations with applying a regression using monthly average flows at the Alston gage as a driver for computing daily inflows to Parr. Part of the reason (maybe the whole reason) for using an alternative method for estimating daily inflow is that the straight area proration method likely overestimates daily inflow during low inflow periods. I at first was not convinced that the method presented here would provide the best estimate of low flows on daily to weekly time scales due to the reliance on statistics from monthly averages which tends to smooth out the daily variations. After comparing hydrographs for several low flow years (2002, 2007, etc.) using the method presented in this report with a hydrograph developed using the area proration method (and with a hydrograph using just the sum of the 3 gages) the resulting daily inflow dataset seems reasonable (and thus, the concern over

homogeneity above may not be an issue) for low to moderate flows. I did not look at high flows in detail since I am not too concerned at that end.

Daily data evaluation for the development of the regional coefficients is a noted concern due to the potential short-term mass balance impacts associated with the significant usable storage. Even under low flow conditions, a mass balance approach for determining the regional coefficients should have good correlation. Using the entire range of flows for developing the regional coefficients has more effect on the accuracy at the upper and lower ends, as prorating coefficients are widely acknowledged to vary with flows. Observation of the initial regression results, with coefficients derived using the entire range of flows, indicated a tendency for the model to over-predict lower flows. This inflection was noted in section 1.3.2 to be around 7,700 cfs, above which the model tended to under-predict flows. Concern for low-end accuracy led to the regression based upon flows at or below the Parr Hydro capacity, which was approximately 75% of the inflow months. This reduced the tendency of the model to over-predict lower flows, at the expense of higher flow predicted accuracy.

5. As has been suggested by others, a meeting is probably necessary to further discuss and clarify the inflow methodology.

Responses to Byron Hamstead, USFWS Fish and Wildlife Biologist

Email:

Hi Kelly,

Please see attached for the USFWS's comments/questions in track changes regarding the Parr inflow dataset statistical methodology.

Thank you,

Byron

Requested edit: "As discussed in the Study Plan, the ~~existence~~ operation of the pumped storage development and lack of long-term operational records prevents the back-calculation of a sufficient inflow dataset." [Replace existence with operation].

Answer: Agreed, edit incorporated.

Comment: *Y axis label = unadjusted Q* (regarding the Figure 2 Monthly Average Flows column chart)

Answer: Agreed, Label Added to Chart in final version.

Comment:

"The comparison (see Figure 3) illustrates that the range of the monthly averages (per 100 sq. mi.) was visually close to the aggregate average through a variety of flow ranges; this indicates the hydrologic similarity of the three subbasins."

BH: Is there a benefit of normalizing discharge by 100 sq. mi. versus normalizing by 1 sq. mi.?

Answer: The scale for normalizing was selected to match the order of magnitude of the contributing (smallest) drainage area. Examining the three gages on a cfs per unit square mile would not change the results or the relative contribution of any gage area, but only the scale. The lower flows would change from around 10 cfs/100 square miles to 0.1 cfs/square mile, while the higher 420 cfs/100 square miles would reduce to 4.2 cfs/square mile.

BH: I think it is necessary to quantify statistical differences between gages in terms of Q/square mile since subbasin hydrologic homogeneity is an important assumption included in the model. Accounting for these differences might further reduce the variance in the model, making it more accurate at lower flows.

Answer: Visual examination of the normalized flows was done to check for consistent, significant discrepancies between gaged areas under a range of hydrologic conditions. The comparison of any single normalized gage with the aggregate average was visibly within the same order of magnitude for all months across a large range of inflow conditions, and was the basis for concluding the similarity. While any given month may show one gaged area has a noticeably higher contribution, no general trend indicates a consistent bias across the range of hydrologic conditions. Significant differences in runoff characteristics would be indicated by one or more normalized areas consistently contributing more or less than the aggregate average. In the absence of significant consistent contribution by any single gage, consistent fitted regional coefficients (alpha and lambda) were selected for all three gaged areas. Variances observed for individual months, where one gaged area contributes more or less than others, is attributable to precipitation that was inconsistent for the entire drainage area, rather than differences in runoff characteristics.

BH: Was this the sole period of record [referring to Figure 3, Normalized Monthly Average Flows, which shows 2002 – 2003 calendar years] used to infer similarity of runoff characteristics among subwatersheds? According to table 1 there are overlapping discharge data for all of these gages since 1973.

There appear to be potentially significant differences in mean monthly discharge between gages even when the data is normalized by drainage area.

Answer: The period of record used to infer similarity was 1981 – 2013, the longest concurrent period for the four gages available (in complete calendar years); the Alston Gage period of record has a gap in the dataset from 1907 through 1980. We will correct the current period or record in Table 1 in the final version. Only two years were charted for the document for visibility, selected to illustrate consistent gaged contributions across a range of hydrologic conditions: extreme drought conditions during the summer of 2002, and high inflows the following spring.

Comment:

"These two gages [Woodruff and Whitmire gages on the Enoree River] were selected because they have the longest overlapping (current) periods of record."

BH: What is the period of record for discharge here?

The proposed Riverdale Project (formerly Inman Mills) was licensed in 1982, but became inoperable 12-years ago. Since this calculation assumes that the hydrologic characteristics of the Enoree River apply throughout the Broad River subwatershed, I want to make sure that the regional exponent/model is not confounded by a period of record that includes river regulation activity.

Answer: The overlapping period of record for the Whitmire and Woodruff gages is indicated in Table 1 as 2-9-1993 to present, limited by the Woodruff gage. The use of monthly flow averages to establish the pro-rating coefficient would eliminate any effects of short-term regulation upstream of the Parr dam. FERC documentation (correspondence from project licensee) indicates the Riverdale project has not operated since August 2001.

With respect to daily average flows that will be prorated to create the dataset, the project has insignificant storage and re-regulating capacity with respect to the Parr Reservoir (9 acre pond with a gross storage of 22 gross acre-feet, compared to 4,400 acres and 32,000 acre-feet).

Comment:

TABLE 1 STATISTICAL MODEL RESULTS SUMMARY

MODEL NAME	REGRESSION DATASET OF ALL MONTHLY AVERAGES (396 VALUES)	REGRESSION DATASET OF LOWEST 75% MONTHLY AVERAGES (289 VALUES)
α – Coefficient	1.041	0.988
γ – Exponent	0.599	0.599
Standard Error	495.0	469.6
R^2	0.9828	0.9828

BH: The standard error [469.6] for this model may be too high considering that annual daily flows are often below 3,000, and approach 2,000 cfs in late Summer/ early Fall.

Figure 6 shows a few stray data points that may be driving up SE. Were any statistical outliers omitted from analysis?

Answer:

The Standard Error represents the standard deviation across the entire range of flows. The Standard Error on the left and right columns are based on the associated regional coefficient and exponent, which were established according to the conditions of the headings (all flows vs. lower 75% flows, approximately 6,000 cfs limit). The Standard Error for only low-flow scenarios would have lower values. The Standard Error calculated for flows up to 6,000 cfs is 321 for the left column, and 304 for the right column. The Relative Standard Error of the entire dataset more accurately explains the error versus the total range of flows. For both regressions, the RSE is calculated at 9.3%.

No statistical outliers were omitted from the analysis, as the good correlation between the predicted and measured flows across the range of data did not suggest that data points needed eliminated.

Responses to Gerrit Jobsis, American Rivers Sr. Director:

Email:

Kelly,

Please find attached American Rivers comments on the inflow data plan. It is intended to support the Final Parr Fairfield Operations Model Study Plan. That study plan says "The goal of this task is to create the best available historic inflow series, which will form the input to the operations models, energy models, and habit and recreational studies." As my comments in the document state, I do not agree that this inflow data set will be usable to evaluate the effects of project operations on habitat and recreation. Project operations via inflow alterations and reservoir fluctuations affect habitat and recreation values *on a real time basis (hourly or less) that cannot be estimated using monthly average inflow estimates. Smoothing the data with regression equations removes the hourly and sub-hourly variation that is essential to understanding project effects.*

I received USFWS comments which also raise some important questions. It would useful to convene a call among those interested to answer some of the questions raised in our respective comments.

Gerrit

> Answer:
>
> The inflow dataset is a model input that is independent of the project operations. This effort is to determine accurate coefficients for prorating the gaged inflows for summing the total dataset. They are being determined on a monthly basis because mass balance between the upstream gages and the Alston gage can be significantly affected by project operations. Daily analysis could be performed, but would introduce a significant level of inaccuracy in determining the coefficients. The inflow dataset will be developed as mean daily flows, using the coefficients determined through the mass balance effort. Hourly inflows are proposed to be the same as daily average, as the travel time between gages under varying flows would introduce high potential for inaccuracy. The model outputs will evaluate the hourly and daily impacts on the areas within the PBL and the reach downstream of the Parr Shoals dam.

Comment:

"The statistical analyses will use monthly and annual flow data rather than daily average flows."

GJ: I don't agree with this for evaluating a project effects on stream flow (inflow versus outflow) and reservoir fluctuations. Project effects occur on an hourly or shorter time frame. Analysis of project effects should be done similarly. The issue for habitat and recreation is not how Parr/Fairfield affects monthly or annually, but within the day and hour.

> Answer:
>
> Project effects will be evaluated via modeling efforts on time steps of an hourly basis, in addition to any longer periods requested.

Comment:

"Flow releases from the project may be vastly different at any given hour from the inflows to the Parr reservoir."

GJ: This is exactly what we need to understand

> Answer:
>
> This statement is alluding to the inherent error associated with calibrating the inflows with the Alston gage on a daily basis, due to the storage of the project. The model will facilitate the

understanding of these releases. The inflow dataset will not be affected by project operations, but is an independent input.

Comment:

" A multivariate regression was performed to determine the parameters of a generalized equation for estimating the inflow to Parr Reservoir."

GJ: Again, this may be good for the operations models and energy models but I don't understand how this will help answer the question of how the project affects streamflow and reservoir fluctuations. Smoothing things out with a regression takes away the variability of inflow that is essential to understanding project effects on habitat and recreation.

Answer:

This regression is performed only to determine the regional prorating coefficients. Project effects on streamflow and fluctuations are addressed in the Res and RAS models. The regression is not intended to smooth out the extreme high and low flows, but rather best establish the prorating coefficients to most accurately represent the inflow. Inflows will still be highly variable, based on mean daily records.

Comment on graph:

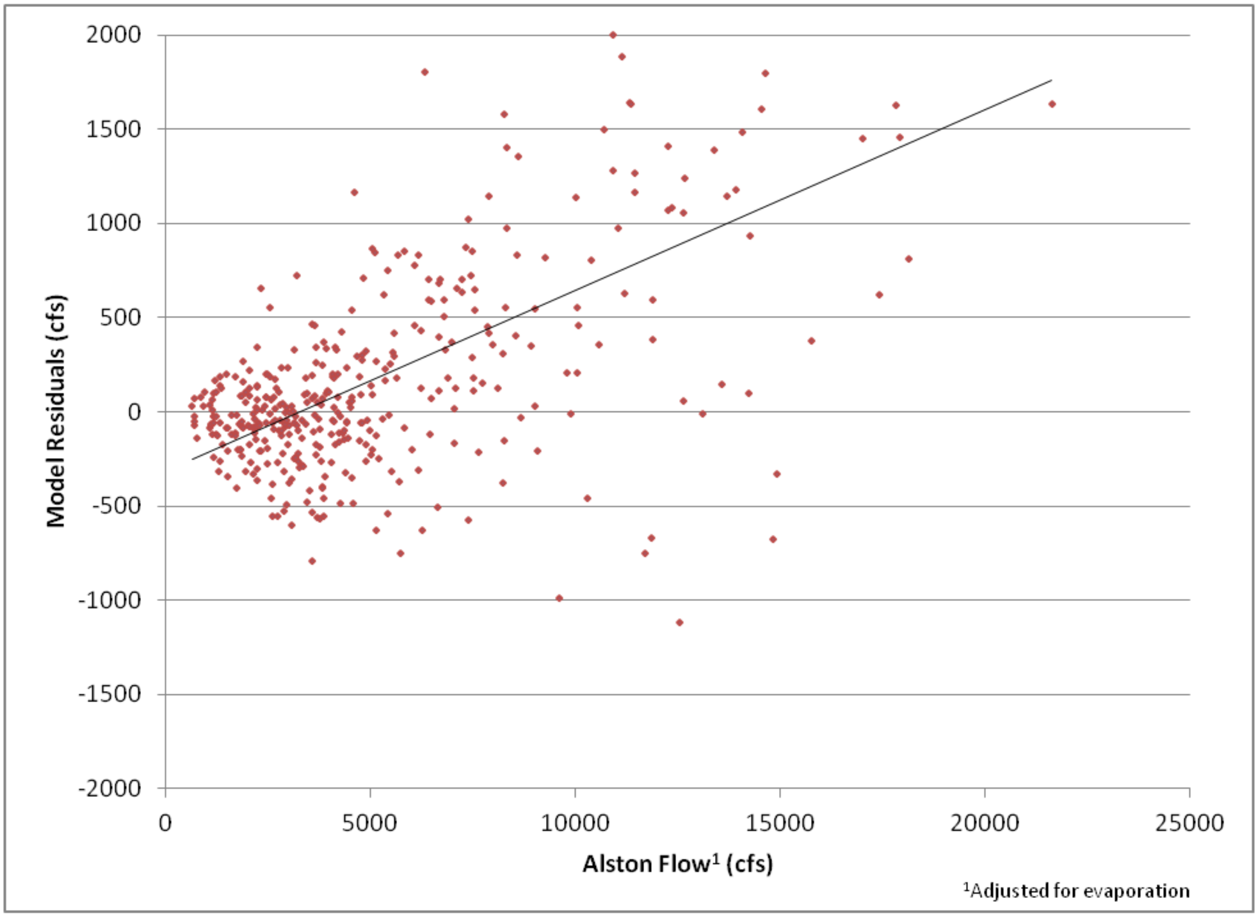
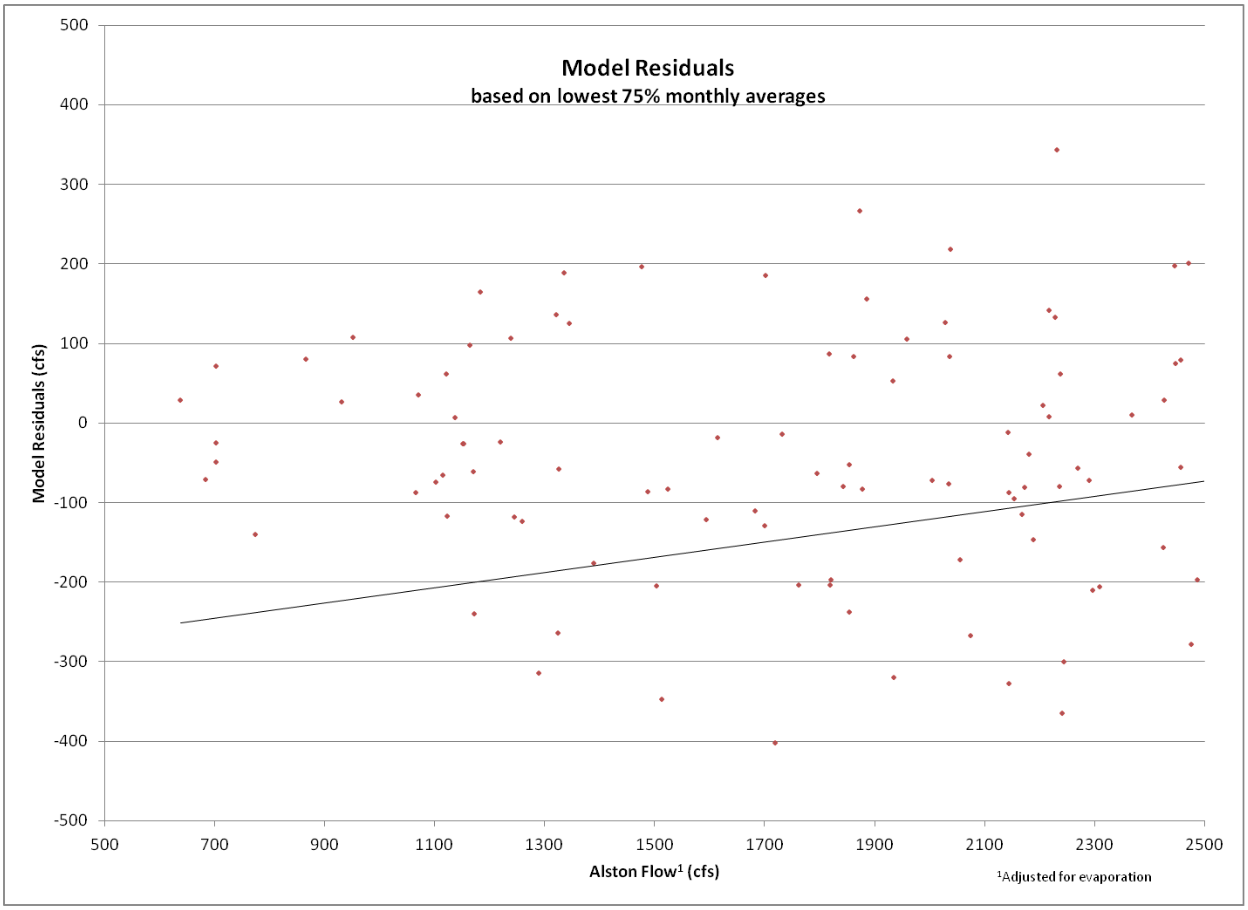


FIGURE 1 **MODEL RESIDUALS - REGRESSION BASED ON 75% LOWEST FLOW AVERAGE MONTHS**

GJ: Poor fit at lower end of flow range affects the reliability of the model

Answer:

The residuals diminish in magnitude as flows decrease, are appear evenly distributed about the zero value. While the inflow dataset will have calculated values both higher and lower than the Alston readings, no significant bias is evident under low flow conditions. A closer examination of the low-end flows can be made with the graph below, scaled to flows below 2500 cfs. (The trendline is a linear average across all flows for the 75% lower inflow months, and does not represent the trend of the lower flow residuals alone.)



Comment:

TABLE 2 STATISTICAL MODEL RESULTS SUMMARY

MODEL NAME	REGRESSION DATASET OF ALL MONTHLY AVERAGES (396 VALUES)	REGRESSION DATASET OF LOWEST 75% MONTHLY AVERAGES (289 VALUES)
α – Coefficient	1.041	0.988
γ – Exponent	0.599	0.599
Standard Error	495.0	469.6
R^2	0.9828	0.9828

GJ: This [referring to the 469.6 standard error value] seems significantly high when evaluating low flow periods and could represent 20% to 25% of the average flow

Answer:

The Standard Error represents the standard deviation across the entire range of monthly average flows (up to 20,000 cfs). The Standard Error on the left and right columns are based on the associated regional coefficient and exponent, which were established according to the conditions of the headings (all flows vs. lower 75% flows, approximately 6,000 cfs limit). The Standard Error calculated for low-flow conditions has lower values. For example, the calculated Standard Error for the two columns limited to flows up to 6,000 cfs are 320 and 304 (left and right respectively). For flows up to 2,000 cfs, they are 155 and 147. If considered from a percentage perspective, as the Relative Standard Error, it would more accurately explain the error versus the total range of flows. For both regressions, the RSE is calculated at 9.3%.

Response to Pace Wilber, NOAA National Marine Fisheries Service Atlantic Branch Supervisor

Hi Kelly. I agree with the comments from FWS and American Rivers that short-term variation important for assessing project effects on fishes and riverine habitat may be masked by using monthly average flows as model inputs. I also agree there are much better ways to judge the similarity of flows between subwatersheds than "eyeballing" the histograms in figures 2 and 3. A correlation matrix may be a more rigorous way to make the comparisons. Pace

Answer: Short-term variation will still be performed using daily mean inflows. Monthly average flows are only being used to determine regional pro-rating coefficients for daily inflow calculations, due to the mass balance errors associated with daily operations.

Visual examination of the normalized flows was done to check for consistent, significant discrepancies between gaged areas under a range of hydrologic conditions. The comparison of any single normalized gage with the aggregate average was visibly within the same order of magnitude for all months across a large range of inflow conditions, and was the basis for concluding the similarity. Due to the good overall correlation, it is unlikely that altering one set of regional coefficients to more accurately represent the contributing ungaged area will offer significant improvement to the model. Lower homogeneity in runoff characteristics may be inferred from metrics when the contributing factor is actual weather event(s) specific to a single subbasin within a given month.