Guidelines for Standardizing and Simplifying Operational Record Extension

South Carolina Surface Water Quantity Models – Unimpaired Flow Development

CDM Smith, March 2015

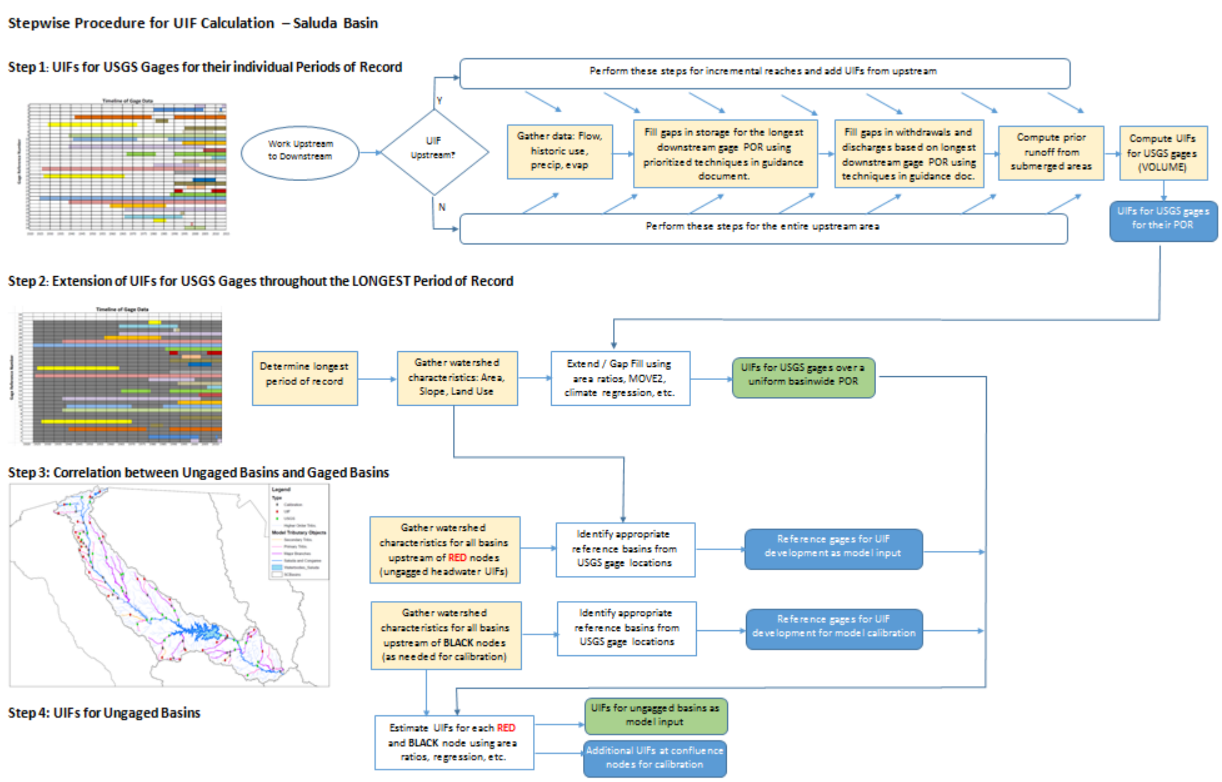
Objective:

This set of guidelines is intended to help simplify and standardize the process of extending and filling gaps in operational records of water **withdrawals, discharges,** and **storage impacts** as part of the process of developing Unimpaired Flows (UIFs) for the South Carolina water quantity models. It is based on the following principles of large-scale water planning:

a) De-emphasize the nuances of specific undocumented local issues (such as matching population trends with service area changes, etc.) and generalize water use trends regionally, and

b) Provide a consistent framework for filling data gaps and extending records

Summary text appears in blue. Note that the recommendations in this document apply only to the synthetic extension of operational records, and not to the extension of the UIFs themselves (the alternative procedures for which are described in the UIF Methodology TM). That is, the guidelines in this document apply to the gap-filling boxes in Step 1 of the overall UIF process below:



While the ultimate UIF data sets in any given basin are required to extend all the way back to the earliest USGS record in the basin, IT IS ONLY NECESSARY TO SYNTHESIZE OPERATIONAL DATA FOR EACH SPECIFIC USE BACK TO THE DATE OF THE EARLIEST **DOWNSTREAM** USGS GAGE RECORD, either on the tributary of use, or downstream on the mainstem. This is because the downstream gages will be the basis for UIFs using upstream impairments, but once each UIF is developed for the period of gaged record at each gage, the UIFs themselves will be statistically extended using other techniques that do not rely on historic use (Step 2 in the diagram above). In other words, if there are no streamflow records for which a given use would be used in unimpairment calculations, we do not need the use record.

> **GENERAL SIMPLIFICATION:** Only extend use data back to the date of the earliest downstream USGS flow record within the basin that would use the data in unimpairment calculations over its period of record.

Specific Guidelines for Water Withdrawals

Water withdrawals may need to be disaggregated into annual and then monthly values (monthly values would be spread evenly across the days in the month). To estimate undocumented water withdrawals on an **ANNUAL** basis (as an example, consider a documented withdrawal from 1990-2013, which requires extension back to 1950):

- **First Priority - Anecdotal Information:** If anectodal information about dates and volumes is available via direct communication from water users, this should be used and interpolated/extrapolated to the greatest extent possible. In the example above, if the water user informs us that the intake came on line in 1962 and started at 2mgd, linearly interpolate usage from 2 mgd in 1962 to the documented value in 1990. *Note: Do not synthesize water use prior to any known date of initiation (in this example, 1962).*

- **Second Priority – Regional Population Trends:** In the example above, if there is a correlation between population and withdrawals from 1990-2013, this correlation can be applied going back in time. Note that the correlation could be as simple as a per capita use rate. DO NOT attempt to fully reconcile local population, county population, and service area, as the relationship between all of these will change over time and would consume too much time to document in every case. Rather, use judgment on whether local, county, or service area estimates (based on availability of data and applicability to the case at hand) will serve as a reasonable indicator of trends in the service area. Note that correlation relationships should be simple – linear if possible, unless there are obvious nonlinearities in the observed trends. In no case should we use anything more than a second order polynomial (because these can exaggerate conditions at the ends of the time spectrum, and sometimes reverse directions inappropriately).

- **Short-Term Gap Filling**: For short-duration periods of missing information between documented periods (up to ~5 years), values may be linearly interpolated between dates of available data. Refer also to the guidelines for monthly estimation below.

To superimpose **SEASONAL OR MONTHLY** withdrawal patterns on these annual averages, compute average monthly multipliers for the documented period of record, and apply these for the period of record extension. Ensure that they average 100%. Do not adjust for the variability in the number of days per month.

Specific Guidelines for Water Discharges

To estimate undocumented discharges, first determine if there is a repeatable monthly pattern of discharge. If not, hindcast using annual values using the guidelines below and apply the discharge as a constant rate throughout the year per below. If there is an observable monthly pattern, refer to the monthly guidelines below the annual guidelines, and choose an option based on the data.

FOR ANNUAL AVERAGE DISCHARGE VALUES:

- **First Priority – Anecdotal Information:** If anectodal information about dates and volumes is available via direct communication from water users, this should be used and interpolated/extrapolated to the greatest extent possible.

- **Second Priority – Correlation with Withdrawal**: If documented discharges can be correlated with documented withdrawals, the correlation can be extended back in time. This actually matches the SWAM model construct, in which discharges are usually specified in terms of corresponding withdrawal percentages.

- **Third Priority – Permit Estimates**: In some cases, discharge permits estimate the discharge volume as a percentage of withdrawal. In such cases, this can be a simple approximation of the historical discharge volumes.

- **Fourth Priority – Regional Population Trends**: If there is a correlation between population and withdrawals during the documented period, this correlation can be applied going back in time. DO NOT attempt to reconcile local population, county population, and service area, as the relationship between all of these will change over time and would consume too much time to trace and document in every case. Rather, assume that either local or county level population (based on availability of data and applicability to the case at hand) will serve as a reasonable indicator of trends in the service area (especially if good correlation exists for the period of documented discharge). Note that correlation relationships should be simple – linear if possible, unless there are obvious nonlinearities in the observed trends. In no case should we use anything more than a second order polynomial (because these can exaggerate conditions at the ends of the time spectrum, and sometimes reverse directions inappropriately).

- **Short-Term Gap Filling**: For short-duration periods of missing information between documented periods (up to ~5 years), values may be linearly interpolated between dates of available data. Refer also to the guidelines for monthly estimation below.

If there is an observable monthly pattern to withdrawals, then use the following guidelines and choose the approach that best matches the situation or available data:

FOR MONTHLY DISCHARGE VALUES (if observed patterns exist):

- **Option 1 – Correlate with Monthly Withdrawal**: If monthly discharge can be well correlated to monthly withdrawal, then it may not be necessary to estimate annual discharge. Rather, develop ratios between observed monthly withdrawal and observed monthly discharge for a period over which records overlap. The ratios would most likely be average values for each month, provided there is not too much scatter. Then apply these ratios to the full (possibly extended) record of withdrawals. Note: Do not use synthesized withdrawal data to establish the ratios – use only documented values. However, it is acceptable to use synthesized withdrawals as the basis for extending the discharge by applying the ratios from the documented values.

- **Option 2** – **Apply observed trends to annual discharge estimates**: If the periods of observed withdrawals and observed discharges do not overlap, or there is poor correlation between withdrawal and discharge, then annual average values will need to be determined per the above procedures, and monthly multipliers applied. Determine average monthly multipliers of discharge, using documented (not extended) annual average as a basis. Ensure that the multipliers average 100%. Then, apply these multipliers to annual average discharge estimates from the procedures above.

FOR INDUSTRIAL DISCHARGES:

For industrial discharges with no withdrawal (groundwater use, for example), simply extrapolate observed data back to the known or estimated date at which operations commenced. This would apply on an annual and/or monthly basis, as deemed appropriate based on the available data.

Specific Guidelines for Storage Impacts

There will be cases in which we need to synthesize the impacts of reservoirs in the absence of documented fluctuations in storage and/or elevation. The presence of reservoirs affects both the timing of flow and the volume of water in the river system. The following guidelines may be applied:

- **Surface Evaporation (volume impact):** Assume full reservoir area for computing surface evaporation in the absence of records of reservoir fluctuations.

- **Surface Precipitation (volume impact):** Assume full reservoir area for computing surface precipitation in the absence of records of reservoir fluctuations.

- **Change in Storage (timing impact)**: Knowing the historic fluctuation in storage is useful because by impounding water, drawing down, and recovering, the timing of when water is released can be affected. Impoundment does not, however, affect the total volume of water in the system, only the distribution of that water as flow over time. To estimate historical water level fluctuations accurately, a calibrated hydrologic and operations model would be needed. This is not always practical, so several alternatives are offered for hind-casting historical reservoir elevation/storage:

    - **First Priority – Published Estimates from Other Modeling Studies**: Many of the basins in South Carolina have been simulated with reservoir operations models (CHEOPS, for example, or HEC-ResSim). As available (without re-running the models), published values from these models can be used to help extend or fill reservoir records.

- **Second Priority – Extrapolation and Correlation with Precipitation**: There are three proposed approaches that can be applied in various conditions. The decision of which method to use should account for the availability and credibility of data, as well as the overall dynamics of the reservoir, per the guidelines below. The 2nd and 3rd methods are described in more detail on the pages that follow, but summarized here. Note that in many cases, it may simply be best to see which of these methods reproduces observed data the best, and rely upon that method purely on its predictive basis. It should be emphasized, though, that hindcasting reservoir storage *does not* account for detailed operational practices, but rather the observed patterns of drawdown, and the apparent dependence the drawdown may have on prior rainfall levels. The graphs that follow the detailed descriptions of the two regression methods illustrate how the two methods may be appropriate for different types of reservoir response patterns. Additionally, following the graphs, a procedure is outlined for adjusting the hindcast timeseries for the potential impacts of variable historical withdrawal rates (if such data are available).

    a) **METHOD 1: Simplest: Monthly Averages*: [To be used only if there is a clear and consistent pattern of drawdown and refill that does not vary significantly from year-to-year].*** Monthly average elevation/storage can be computed for the period of documented record, and these can be applied as estimated hindcasts. Daily values can be interpolated between monthly values. It should be noted with our UIF records that if this method is employed for reservoirs with a great deal of year-to-year variability in water levels, that this is a very approximate technique.

    b) **METHOD 2: Next Simplest: (REGRESSION METHOD A) Correlation Between Daily Elevation and Cumulative Historic Precipitation**: ***[To be used if the reservoir is frequently full, but exhibits irregular drawdown during droughts] – SEE FULL PROCEDURAL DESCRIPTION BELOW FOR REGRESSION METHOD A.***

    c) **METHOD 3: More Complex: (REGRESSION METHOD B) Scaling the Monthly/Daily Averages from (a) above to expected min annual elevation based on historic precip**: ***[To be used if the reservoir experiences significant multi-year or irregular drawdowns during droughts, and is not frequently observed to be full.] - SEE FULL PROCEDURAL DESCRIPTION BELOW FOR REGRESSION METHOD B.***

- **Third Priority – Iteration:** If either of the two methods above are employed for the UIFs, they can be validated or refined once the SWAM models are constructed. This would be a time-consuming process, likely involving iteration between UIFs and model runs, so it should be employed with discretion, and only if truly needed for reservoirs that have pronounced impacts in a basin or a great deal of uncertainty in the hind-casting.

Full Procedure – METHOD 2 - REGRESSION METHOD A:
Hindcasting Reservoir Elevation Using Daily Precipitation Sums
Note: Example spreadsheets are available to assist as reference or templates for this procedure.

This method for developing a historical time series of elevation data for a specific reservoir uses available observed reservoir elevations and daily precipitation records. The precipitation records must cover the entire period of hindcasting and/or gap filling, as they will serve as the independent variable in a regression model. The observed reservoir elevations are needed to develop the regression model, and should cover a multi-year period. The observed reservoir elevations do not need to be continuous, but they must cover an overlapping period with available precipitation data. This procedure may be modified if only average monthly reservoir elevations are available, but will then only be able to hindcast average monthly elevations (or weekly, etc.). The following procedure assumes that daily precipitation data are available for the full hindcast period, and that there is a sufficient multi-year overlap between observed daily reservoir elevations and daily precipitation data.

Step 1: Compile daily observed data. The suggested format for the daily observed data is a continuous time series of dates that span from the 3 years before the earliest reservoir elevation observation to the latest daily reservoir elevation observation, with column headings: Date, Observed Elevation, Daily Precipitation. For example, if the reservoir elevations start on 1/1/2000 and end on 12/31/2010, the time series should span 1/1/1998 to 12/31/2010, and the first 2 years of reservoir observations will be blank.

Step 2: Check linear correlation between preceding daily precipitation sums and reservoir elevation. This step involves calculating the sum of precipitation for the previous X number of days, for each day in the observed data time series. The resulting time series of X-days previous precipitation sum should then be checked for correlation with the reservoir elevation using the RSQ()[1] function in Excel (or similar function to find the linear R-squared correlation in another software). If the table includes precipitation data for 3 years prior to the first reservoir observation, the precipitation sums can go up to the preceding 1,095 days (3 years). The process of computing the preceding X-day precipitation sum and linear correlation value may need to be repeated multiple times to find the best fit precipitation time series. The suggested procedure is to start with the 30-day sum and repeat in 30-day increments until a maximum linear R-squared value is found. For example, the table described in Step 1 is expanded to include the time series of preceding 30-day precipitation total, preceding 60-day precipitation total, preceding 90-day precipitation total, and so on.

Step 3: Use the best-correlated precipitation sums to develop regression equation. The ideal R-squared value is 1.0. If the best linear correlation of all incremental 30-day precipitation sums going back 3 years is not greater than 0.5, this may not be the best method to use to hindcast reservoir elevations. Once the best-linear-fit precipitation sums time series is established, additional regression functions should be explored that relate precipitation sums to reservoir elevation. For example, a logarithmic regression relationship between the 240-day precipitation and observed reservoir elevation may provide a slightly higher R-squared value than the linear regression. Generally, the function types should be limited to linear, logarithmic, exponential, and power. The final hindcast model formula, which uses the X-day preceding precipitation sum to estimate the reservoir elevation, will take the following form:

Elev = min(Max, F(Psum))
Psum: Sum of daily precipitation totals for the X-day period discovered in Step 2
Max: Maximum possible reservoir elevation
Elev: Calculated reservoir elevation
F(Psum): Regression function that produces highest R-squared correlation between Psum and Elev
An example of this model function is:
Elev = min(1230, 32*LN(Psum)+1078)

---

[1] If the precipitation sum time series is in column A, and the reservoir elevation time series is column B, the format for this formula is: RSQ(column B, column A); or more generally: RSQ(known Ys, known Xs)

Where:
Max = 1230, and
F(Psum) = 32*LN(Psum)+1078

Step 4: Check the agreement between observed and modeled reservoir elevations. This step is qualitative. Does the model capture the times when the reservoir is full? Does the model adequately reproduce significant drawdowns? Is the model biased high or low throughout the overlap time period? This step will determine if this method is appropriate for hindcasting elevations for this reservoir. For example, if significant annual drawdowns are not represented by the modeled elevations, another method for hindcasting should be explored.

Step 5: Hindcast the reservoir elevations using the regression model and historic precipitation data. The final step is to calculate estimated reservoir elevation for each day in the full hindcast time series for which there are no observations. This will be done using the X-day precipitation sum time series for the full period, and the model equation developed in Step 3. The suggested format for this step is a daily time series table covering the full hindcast period (e.g. 1/1/1925 to 12/31/2013) with the following columns: Date, Observed Precipitation, X-day precipitation sum, Observed Elevation, Modeled Elevation. The Observed Elevation rows will be blank for days with no reservoir observations. The modeled Elevation rows will be blank for days with reservoir observations. The combination of these time series will be used for the unimpaired flow development.

Full Procedure – METHOD 3 - REGRESSION METHOD B:
Scaling Monthly/Daily Average Elevation to Expected Minimum Annual Elevation Based on Historic Precipitation
Note: Example spreadsheets are available to assist as reference or templates for this procedure. See "Reservoir Hindcasting – Method 2 Example.xlsx"

Like Method 2 above, this method for synthesizing a historical time series of elevation data for a specific reservoir uses available observed daily or monthly reservoir elevations and annual precipitation records. The precipitation records must cover the entire period of hindcasting and/or gap filling, as they will serve as the independent variable in a regression model. The observed reservoir elevations are needed to develop the regression model, and should cover a multi-year period. The observed reservoir elevations do not need to be continuous, but they must cover an overlapping period with available precipitation data. At a minimum, the data should cover a significant drawdown and full recovery of the reservoir to a full condition. This procedure may be applied with either daily or monthly reservoir elevation data, and any form of precipitation data that can be aggregated into annual totals. The following procedure assumes that there is a sufficient multi-year overlap between observed reservoir elevations and precipitation data.

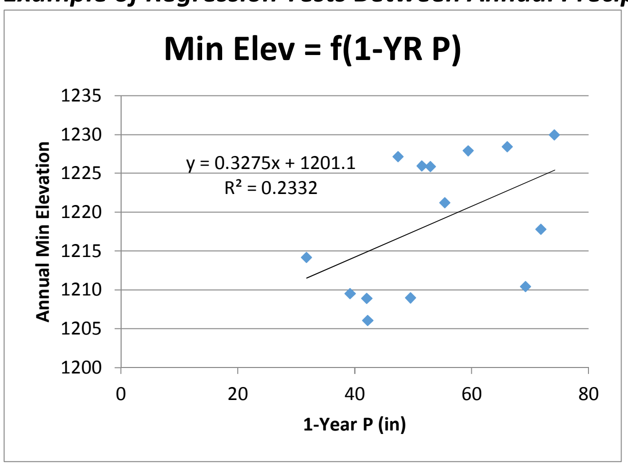
Step 1 - Collect Data: Gather all available information on precipitation and reservoir elevation. Precipitation may be daily, monthly, or annual. Reservoir elevation may be daily or monthly.

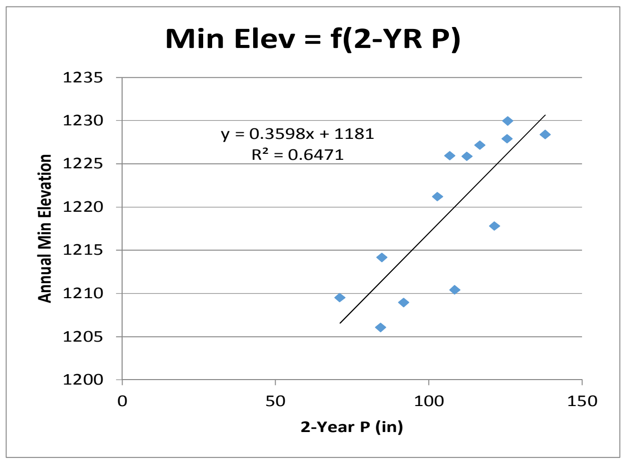
Step 2 - Compute Daily Average Elevation: Over the reservoir period of record, compute a *one-year timeseries* of daily average elevation for each day of the year. For example, the elevation for January 1 would be the average values of all records from January 1 in the period of record. If reservoir elevation is reported monthly, interpolate linearly to approximate daily values. (This is the same as Method 1, above, but it will serve as an interim step in Method 3, here).

Step 3 – Annualize Data from Step 1: Using pivot tables or other means, summarize the recorded data from Step 1 in the form of **Total** Annual Precipitation (summation) and **Minimum** Annual Elevation. For each year in the reservoir's period of record, then, there will be a value of annual precipitation that can be correlated in the next step with the minimum elevation (maximum drawdown) for that year.

Step 4 – Regression Relationship Between Annual Precipitation and Annual Minimum Elevation: Develop a relationship (preferably linear) between Annual Precipitation and Annual Minimum Elevation. In some cases, a relationship may not develop until the past 2 or 3 years of precipitation are added together, so multiple regression tests may be needed to find a good relationship between antecedent rainfall totals and minimum reservoir elevation in a given year. *If a good relationship cannot be clearly developed for the period of record, or if the record does not include a good example of significant drawdown and full recovery, this method may not be appropriate.* The example below shows poor correlation using 1-year total rainfall, but reasonably good correlation using 2-year total rainfall:

Example of Regression Tests Between Annual Precipitation and Annual Minimum Elevation





Step 5 – Extend Minimum Annual Elevation Record: Using the regression relationship from Step 4, extend the annual timeseries of minimum annual elevation over the entire period of record for the basin (defined by the earliest recorded USGS streamflow) using the precipitation statistics as the predictive variable. Also validate the relationship over the period of record for reservoir elevation.

Step 6 – Develop Annual Scaling Factors: For each year in the period for which no reservoir elevation data exist, develop a single annual scaling factor that relates the estimated minimum annual elevation (from Step 5) with the minimum elevation of the Average Year pattern from Step 2. However, before computing these values, convert the minimum elevation into Maximum Drawdown in order to properly scale the relativity of the two values (Full Reservoir Elevation – Minimum Elevation). For example, for a reservoir with a maximum elevation of 1230 feet, if the estimated minimum elevation from Step 5 for year X is 1210 feet, and the minimum elevation of the average year pattern from Step 2 is 1225 feet, the scaling factor would be:

$$Scale\ Factor_{Year\ X} = \frac{Max\ Drawdown_{Year\ X}}{Max\ Drawdown_{Avg\ Year}} \frac{(Full\ Elev - Min\ Elev_{Year\ X})}{(Full\ Elev - Min\ Elev_{Avg\ Year})} = \frac{(1230 - 1210)}{(1230 - 1225)} = \frac{20}{5} = 4$$

The end product of this step will be a timeseries of ANNUAL scaling factors for each year in which no reservoir records exist. It is conceivable that some scale factors could be negative, depending on the regression relationship from Step 4. Consider these carefully, and possibly apply a lower bound of 0 for the scaling factors.

Step 7 – Develop Synthetic Timeseries of Reservoir Drawdown: This is the final step in this procedure, and will result in a DAILY timeseries of estimated reservoir elevation for the entire period of record for the basin.

7a) First, convert the average daily elevations from Step 2 into daily drawdown by subtracting each value from the full reservoir elevation.

7b) Then, copy this annual pattern for every year for which the reservoir record is to be extended or filled.

7c) Next, multiply each value of daily drawdown by the scale factor computed for the corresponding year. *Caution: Do not multiply the actual elevation by the scale factor – rather, multiply the DRAWDOWN (Full Elevation – Daily Elevation) by the scale factor, and then recompute the resulting elevation in 7d.*

7d) Lastly, convert the drawdown values into reservoir elevation values by subtracting them from the full reservoir elevation.

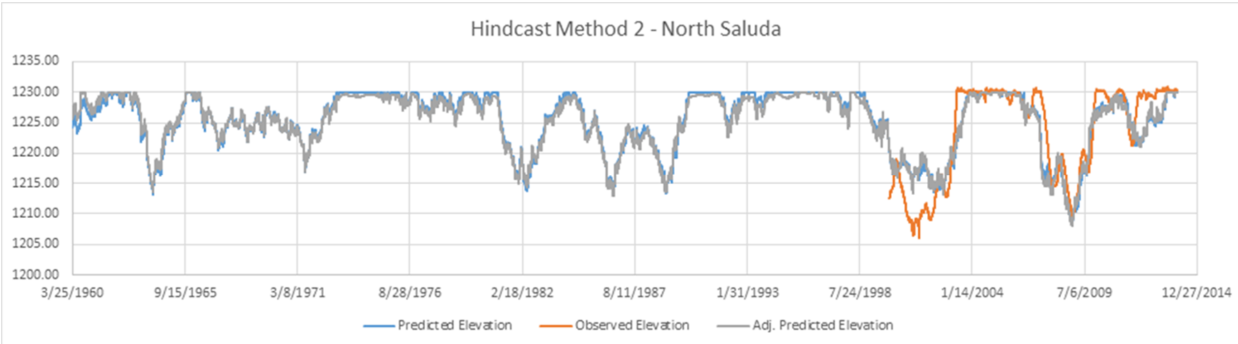
7e) Validate the approach by comparing estimated daily elevation with observed daily or monthly elevation for the period in which the reservoir records exist.

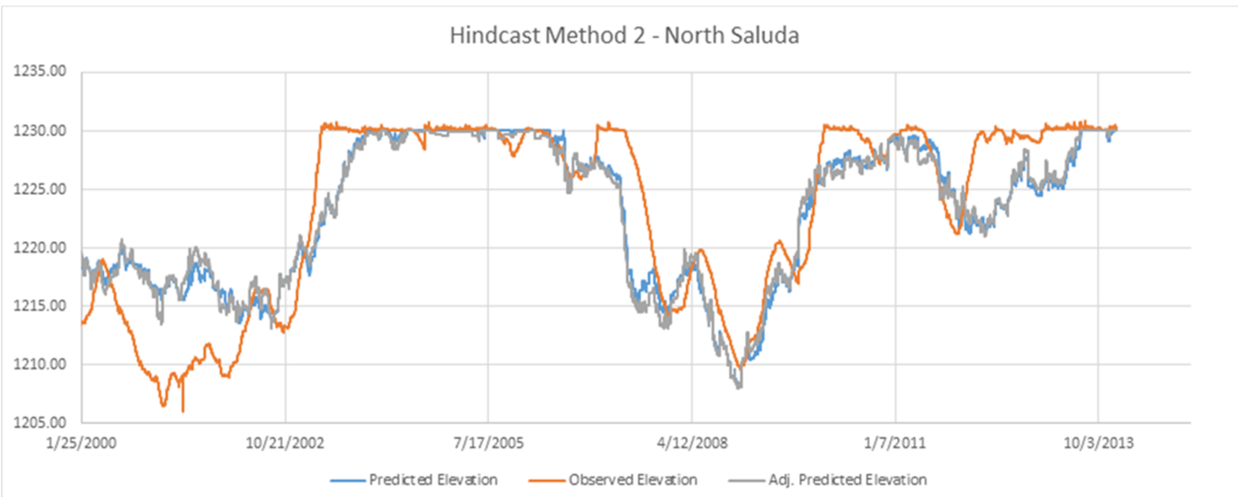
Examples of the Regression Methods:

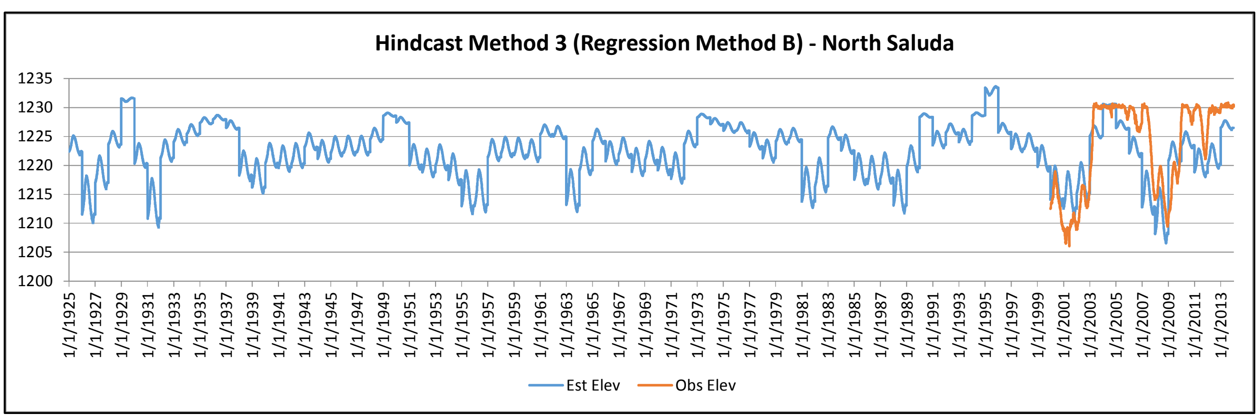
Examples of using these two regression techniques: The two techniques are applied to two reservoirs in the Saluda Basin, and demonstrated below. As noted, this example demonstrates that the best approach may simply be the one with the most obvious predictive ability, but there are some distinguishing features about these two reservoirs that may be important.

In the first example, the two methods are applied to the North Saluda Reservoir. The data suggest that there are extended periods of time over which the reservoir is full, or nearly full, but that it can draw down somewhat irregularly during droughts. **METHOD 2 (Regression Method A) is preferred in this example** because it appears to preserve the full condition more realistically than Method 3, and also simply because it provides a more credible reproduction of the historical drawdown pattern.

First Example: North Saluda Lake

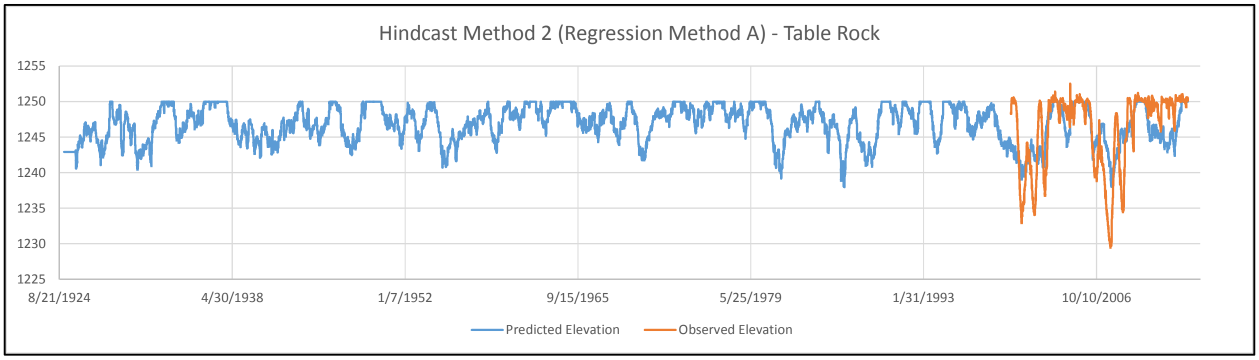


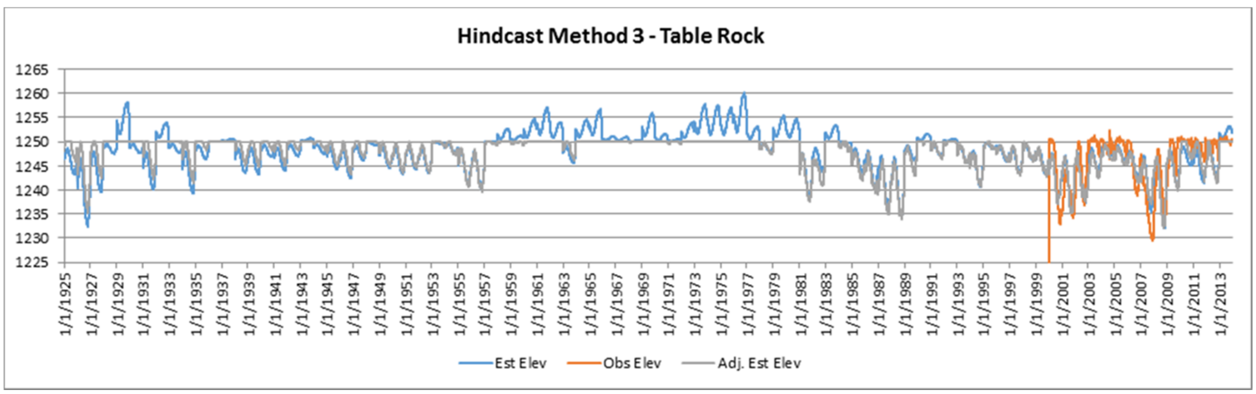


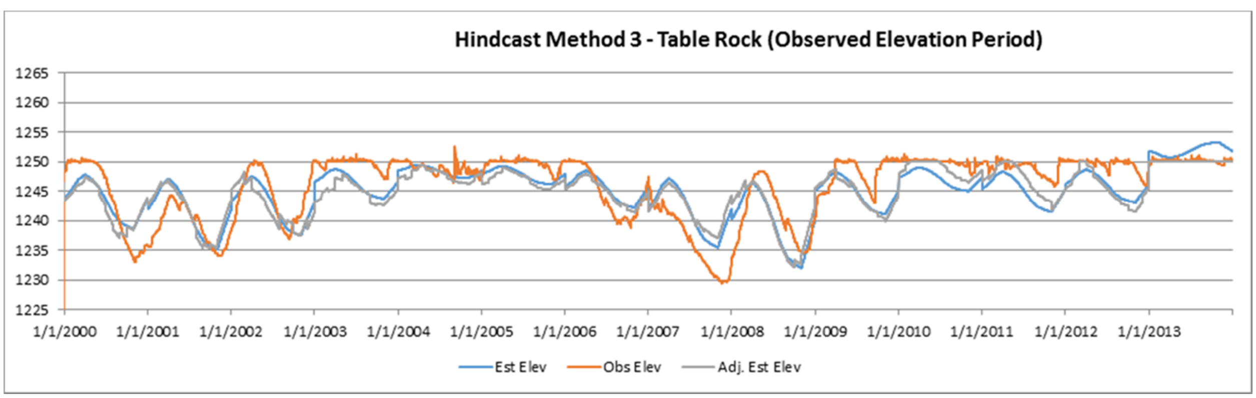


In the second example, the two methods are applied to Table Rock Reservoir. The data suggest that the reservoir draws down irregularly, and is not usually completely full. **METHOD 3 is preferred in this example** because it appears to better match the magnitude of severe drawdown, the reservoir is not usually full, and because the method provides a more credible reproduction of the overall historical pattern.

Second Example: Table Rock Reservoir







Adjustment for Variable Historic Withdrawal Rates

If data for reservoir withdrawals extend back beyond the available data of reservoir water level, adjustments can be made to the hindcast timeseries of reservoir elevation. This is because the elevation hindcasting assumes an average withdrawal pattern equal to the average withdrawals over the period of elevation records, and is aimed principally at distinguishing drawdown due to severe drought from drawdown due to normal reservoir use and operations. It does not explicitly account for drawdown due to variations in reservoir withdrawals.

In such situations, the following approach may be applied (as a supplement to Method 1, 2, or 3 above):

1. Proceed with the full reservoir hindcast procedures as specified above (Method 1, 2, or 3).

2. Compute the average monthly withdrawal over the period of ELEVATION record for each month (the average of all Januaries, the average of all Februaries, etc.)

3. Convert hindcast elevation into hindcast volume for each month using the storage-elevation relationship for the reservoir.

4. Add or subtract volume for each hindcast month based on the difference between recorded withdrawal for that specific month and average withdrawal for the corresponding months over the period of ELEVATION record (computed in Step 2).

5. Convert the adjusted volume back to elevation (but keep both timeseries, as volume is used in the UIF equation, but elevation is used for validation).

Note that this method should NOT be applied with hindcast withdrawal data. Only apply this adjustment step when there are actual operational records of withdrawals that extend back further than the records of reservoir elevation.

Also note that if the period of elevation record suggests that the reservoir does not exceed spillway elevation for extended periods of time, hindcast elevations should be capped at the spillway elevation as a maximum, with the assumption that spills happen quickly. If the period of elevation record demonstrates extended periods of time above the spillway elevation, then the hindcasting can reflect this as well, but it should not exceed the documented maximum elevation.