APPENDIX 4 – ASSESSMENT MATRIX OF PREVIOUS TREE-RING STREAFLOW RECONSTUCTIONS

Stockton-Jacoby Colorado River Reconstruction Assessment

a) **Gage**

1. *Lee's Ferry*, AZ / Colorado River at the compact point.
   There is a potential source of confusion here: the Colorado Compact Point is located one mile downstream of the mouth of the Paria. There is no gage at this location. Flow is estimated by summing flows of the Colorado and Paria Rivers; both gages are located at Lee's Ferry and "Lee Ferry" is the name given to the compact point in all official documents.
2. *Colorado River near Cisco, UT*
3. *Green River at Green River, UT*
4. *San Juan River near Bluff, UT*

b) **Reference**: Stockton, CW and GC Jacoby. 1976. Long-term surface-water and streamflow trends in the Upper Colorado River basin based on tree-ring analysis.

c) **Length of reconstruction**

reconstruction	reconstruction period	calibration period
Colorado R at Compact, AZ	1512 – 1961	1914 – 1963
Colorado R nr Cisco, UT	1640 – 1962	(49 years)*
Green R at Green R, UT	1660 – 1962	(57 years)
San Juan nr Bluff, UT	1661 – 1969	(57 years)

* no dates provided

d) **Model: PCR**

1. Colorado R at Compact Point
   Predictors
      model A, B: PCs derived from a subset tree-ring sites plus lags
   Predictand:
      model A: estimated annual virgin flow (water year) from a USGS data set
      model B: estimated annual virgin flow (water year) from the Upper Colorado Region State-Federal Interagency group.
   Final model: average of reconstructed values from models A & B.

Streamflow gage	Predictors	Predictand
2. Colorado R nr Cisco, UT	PCs derived from a subset of tree-ring sites and their lags	estimated annual virgin flow (water year)
3. Green R at Green R, UT	same as above	same as above
4. San Juan nr Bluff, UT	same as above	same as above

e) **Stop rule: F**

1. Each predictor (PC) was evaluated individually. If the F value did not exceed 3.0, the variable was not used in the calibration equation.

f) **Accuracy:**

1. No verification.

reconstruction	R^2*
Colorado R at Compact Point	
model A	0.86
model B	0.91
Colorado R nr Cisco, UT	0.92
Green R at Green R, UT	0.80
San Juan nr Bluff, UT	0.85

* no adjustment for loss of degrees of freedom

g) **Quality issues**: **TVS**

1. **T** (detrending): negative exponential used to detrend all series; the negative exponential provides a very rigid fit to the data that may be too extreme in some cases and may introduce spurious trends.
2. **V** (variance): possible variance inflation in the earlier portions of the reconstruction due to fewer sites. For the Colorado R. at Compact Point reconstruction, earlier portions of the reconstruction may be dominated by the lower half of the region.
3. **S** (sample size): The subsample-signal-strength statistic was not used in individual chronology development.

h) **Sites**

reconstruction	number
Colorado R at Compact Point	
model A	17
model B	13
Colorado R nr Cisco, UT	4
Green R at Green R, UT	7
San Juan nr Bluff, UT	3

i) **Data reduction: P**

1. Data summarized using Principle Components Analysis. Only eigenvectors with roots greater than 1.00 and accounting for a greater percentage of variance than by chance were used.
2. Unclear how subset of data was selected to use in PCA. For example, in Lee Ferry reconstruction, 17 sites were selected from a possible total of 30.

j) **Autocorrelation adjustment**

1. Tree-ring chronologies not prewhitened; lags of t-1, t+1, and t+2 used in regression model.
2. Streamflow series not prewhitened; lags up to 3 years evaluated in regression models. Best models did not include lags.

k) **Residuals analysis**

1. lack of autocorrelation in the residuals was evaluated in model selection but no supporting statistics presented.

l) **Have data?**

Reconstruction	Missing data
Colorado R at Compact, AZ	none.
Colorado R nr Cisco, UT	escalante forks, co.
Green R at Green R, UT	wind river mountains b, wy; wind river mountains c, wy.
San Juan nr Bluff, UT	ditch canyon, nm; pueblito canyon, nm; spider rock, az.

Hidalgo Colorado River Reconstruction Assessment

Gage:

1. Lee's Ferry
   a) Annual unimpaired streamflow was obtained from Bureau of Reclamation (www.usbr.gov/main/index.html) for 1914 to 1963.
   b) 1913 to 1922 were compiled from 3 major tributaries of UCRB

b) **Reference**: Hidalgo, HG, TC Piechota, and JA Dracup. 2000. Alternative principle components regression procedures for dendrohydrologic reconstructions. *Water Resources Research* **36(11)**:3241-3249

c) **Length of reconstruction: AD. ~1500 - 1961**
   1. Model calibrated on 1914-1963.
   2. ~1500 to 1961. Innermost date of reconstruction not specified; approximate date obtained from graph.

d) **Model: PCR**
   1. predictors: PCs (computed for calibration period) of lagged (-1, 0, +1, +2) chronologies.
   2. predictand: "annual natural streamflow." assumed to be water year. no mention of log transformation.

e) **Stop rule: V**
   1. For each of the models tested, minimum cross-validation standard error (CVSE) was used to select models. Variables (individual chronologies) were added to the model to minimize CVSE.

f) **Accuracy: [0.824, 0.987]**
   1. Regression R^2: no adjustment needed because only one predictor used in model.
   2. Reduction of error statistic from leave-one-out cross-validation.

g) **Quality issues**: **TSC**
   1. **T** (detrending): detrending method not specified; all chronologies were standard chronologies downloaded from ITRDB.
   2. **S** (sample size): chronology development procedure unspecified but subsample-signal-strength statistic (probably) not used.
   3. **C** (coverage): unequal coverage of runoff-producing areas; final predictor site located within a band across the central portion of the UCRB.

h) **Sites**
   1. Six sites were used in the final reconstruction.
   2. Sites were distributed unequally throughout the UCRB, primarily within the middle portion of the basin.

i) **Data reduction: PCA**
   1. PCA was employed to summarize the data and eliminate multicollinearity amongst the predictor variables.

j) **Autocorrelation adjustment**
1. Standard chronologies used; not prewhitened. One of the chronologies (Uinta Mountains) used in the reconstruction had a AC of 0.47.
2. Chronologies lagged before used in model: -1, 0, +1, +2.
3. AC in streamflow series reported to be 0.30

k) **Residuals analysis: N**
1. none

l) **Have data?: Y**
1. Data for all chronologies used in the study have been downloaded from the ITRDB.

Meko-Graybill Gila River Reconstution Assessment

Gage:
1. Gila River at the head of Safford Valley near Solomon, AZ (*USGS ID# 09448500*)

b) **Reference**: Meko, D and DA Graybill. 1995. Tree-ring reconstruction of Upper Gila River discharge. *Water Resources Bulletin* **31(4)**:605-616.

c) **Length of reconstruction: AD. 1663-1985**
1. Model calibrated on full period: 1916-1985 and verified using a split sample procedure.

d) **Model: MLR**
1. Multiple linear regression with current and lagged years: -1 and +1 of six residual (prewhitened) chronologies.
2. Predictand was log10 annual (water year).

e) **Stop rule: F**
1. Stepwise regression: predictors allowed to enter the model in an order determined by F-level until adjusted R^2 reached a maximum.

f) **Accuracy:** $R^2 = 0.66$, $\sigma = 0.186$
1. Full period statistics, R^2 adjusted for degree of freedom loss and standard error of the estimate.
2. Split-sample verification statistics - product moment correlation coefficient and reduction of error statistic. See table below:

	r^2	RE
First half	0.58	0.52
Second half	0.64	0.61

g) **Quality issues**: **TSC**
1. **T** (detrending): ring-width series detrended using a straight line or modified negative exponential. The negative exponential and straight line provide a very rigid fit to the data that may be too extreme in some cases and may introduce spurious trends.
2. **S** (sample size): Subsample-signal-strength was used in development of chronologies.
3. **V** (variance): Prior to 1700, only 3 of 6 sites have adequate subsample signal strength.

h) **Sites**
1. Tree-ring data was comprised of nine sites but four were combined into one, as they were different species from the same location. Total number of sites used in the reconstruction was 6.

i) **Data reduction = none**

j) **Autocorrelation adjustment: R/L**

1. The reconstruction model included lags on prewhitened chronologies. Prewhitening, the removal of autocorrelation, was achieved by filtering chronologies with autoregressive-moving-average models. Use of lags recognizes that too much or too little may have been removed.
2. Streamflow series was not prewhitened. Autocorrelation in this series was found to be not significant.

k) **Residuals analysis**

1. residuals were found to be approximately normally distributed, not autocorrelated and not correlated with the predicted values.

l) **Have data?**

1. Data for 5 of 9 chronologies used in the study have been downloaded from the ITRDB. Missing sites include: Agua Fria, Beaver Creek, Black Mountain white pine, and Black Mountain white fir.

\

Smith-Stockton Salt-Verde Reconstruction Assessment

a) **Gage:**

1. *Verde River*: Verde R. below Bartlett Dam (1895 – 1945) + Verde R. below Tangle Creek (1946 – 1979).
2. *Salt River*: Salt R. near Roosevelt (1914 – 1979).

b) **Reference**: Smith, LP and CW Stockton. 1981. Reconstructed stream flow for the Salt and Verde Rivers from tree-ring data. *Water Resources Bulletin* 17(6): 939-947.

c) **Length of reconstruction**.

reconstruction	reconstruction period	calibration period
Salt River	1702 – 1979	1914 - 1979
	1580 – 1979*	1914 - 1979
Verde River	1702 – 1979	1895 – 1979
	1580 – 1979*	1895 – 1979

* several reconstructions were computed for different subsets of data based on common data period for equal sample sizes. These are the shortest and longest reconstructions based on the largest and smallest number of trees, respectively.

d) **Model: MLR**

1. Predictors: Tree-ring series from respective basins; no lags.
2. Predictand: log10 discharge Salt and Verde Rivers
3. Three different periods were reconstructed: Annual (water year?), October to April, and December to March.

e) **Stop rule: coeff. significance & R^2_{adj}**

1. only predictors significant at the 99% confidence level and resulting in an increase of an adjusted R^2 were entered into the model.

f) **Accuracy.**

1. No verification.

	R^2_{adj}
Salt River	
Annual	
1702 – 1979	0.728
1580 – 1979	0.535
Oct – Apr	
1702 – 1979	0.717
1580 – 1979	0.527
Dec – Mar	
1702 – 1979	0.665
1580 – 1979	0.490
Verde River	
Oct – Apr*	
1702 – 1979	0.744
1580 – 1979	0.465
Dec – Mar	
1702 – 1979	0.701
1580 – 1979	0.469

* no annual because of surface water diversions for agriculture

g) **Quality issues**: **TVS**

1. **T** (detrending): negative exponential used to detrend all series. The negative exponential provides a very rigid fit to the data that may be too extreme in some cases and may introduce spurious trends.
2. **V** (variance): chronology lengths were taken into account in developing the reconstruction model so variance in the early part of the reconstruction should be stable.
3. **S** (sample size): subjective cutoff; chronology length for each of the sites was restricted to the time period covered by at least 4 trees; except for Verde reconstruction, where the period 1580 to 1620 was based on only *one* tree. Generally, among tree-ring sites in the Southwest, four trees will capture around 85% of the chronology signal based on the full sample.

h) **Sites**

1. Salt River: 7 sites; Verde River: 6 sites.
2. Good spatial coverage of both drainages except during earliest portion (1580 – 1620) of Verde River reconstruction when coverage is very poor (one tree).

i) **Data reduction**

1. Data not reduced.

j) **Autocorrelation adjustment**

3. Chronologies prewhitened; but no lags included in the regression model.
4. No assessment of autocorrelation in streamflow series

k) **Residuals analysis**

2. none

l) **Have data?**

Yes.

Graybill Salt+Verde+Tonto Reconstruction Assessment

a) Gage: which gage does the reconstruction apply to?

1. Salt+Verde+Tonto: Salt near Roosevelt; Verde River below Tangle Creek (post 1945) and Verde River below Bartlett Dam (1945 and earlier); Tonto Creek above Gun Creek and Tonto Creek near Roosevelt
2. The two Verde gages were spliced to get longer record. The two Tonto gages, which do not overlap, were merged to get a longer record.
3. The Salt+Tonto+Verde was one of two reconstructions generated in the study; the other is the Gila at the head of Safford Valley,
4. USGS gaged data were used. October-September and April-October flows were reconstructed separately.

b) Reference: = Graybill, D. A., Gregory, D. A., Funkhouser, G. S. & Nials, F. L. (in press) Long-term streamflow reconstructions, river channel morphology, and aboriginal irrigation systems along the Salt and Gila Rivers, Prepared for inclusion in: Environmental Change and Human Adaptation in the Ancient Southwest, edited by Jeffrey S. Dean and David E. Doyel.

c) Length of reconstruction = A.D. 572-1988

1. Model calibrated on 1914-86.
2. Discrepancy in time coverage of reconstruction and of tree-ring chronologies. Tree-ring chronologies do not have data more recent than 1987. How does the model give reconstructed values for 1988?

d) Model = MLR

1. Multiple linear regression, with current and previous (t, t-1) values of regional series as predictors
2. Predictand was log10 flow summed over the water year

e) Accuracy = [0.72, 0.70]

1. Regression R^2, not adjusted for degrees of freedom loss
2. Reduction of error statistic from leave-1-out crossvalidation

f) Quality Issues = T,S,V

1. **T**: Archaeological chronologies were used for part of the record, and the detrending method for those is not discussed. If the samples were short, detrending could have resulting in loss of climate-related low-frequency variation.
2. **V**: Variance inflation toward toward the early part of the record is possible due to the variation of number of chronologies in the regional-average tree-ring. The cause would be failure to "average" out noise uniformly over the reconstruction period.
3. **S**: The subsample-signal-strength statistic was used in developing the living-tree chronologies, but it is unclear from the report what if any control over sample size effects was taken with the archaeological chronologies

g) Sites = number of tree-ring sites (chronologies) the reconstruction is based on: 1-5

1. The regional chronology is an average with a time-varying sample size. Number of sites is 1 chronology for the period 570-1039, and 5 chronologies for the period 1642-1985. A mean value function based on 5 series with noise and a common signal is likely to have lower variance than one of the component series.
2. Before A.D. 1276 the reconstructions are based on wood from archaelogical samples. No data in the report document how well these archaeological series represent the living-tree chronologies used exclusively after A.D. 1370. This could be a problem as the archaeological samples were not taken from sites selected for their likely sensitivity to drought – in fact it is not known exactly where the trees for the archaeological samples were growing
3. The ring widths were detrended conservatively so that wavelengths on the order of half the sample length would probably be retained. This may still be a problem with the archaeological samples, however, because they are generally short series. Thus the early part of the reconstruction may not be able to capture fluctuations lasting many decades.

4. Sampling of runoff regions : the basins are small enough that a sensitive tree-ring site in any of the watersheds should respond to important climate variations relevant to streamflow.

h) Data reduction =A

1. All available chronologies in any year were averaged together to get a regional tree-ring chronologies.
2. Implicit assumption that the chronologies more or less equally valuable for their hydrologic signal, and the signal is represented by well by the mean
3. Effective for removing need to handle multicollinearity, but ignores possibility that some chronologies may be much better than others as hydrologic proxy. Also creates serious problem in early part of reconstruction in that the series governing the reconstruction were not calibrated against the flow.

i) Autocorrelation adjustment = P/L

1. The reconstruction model included lags on prewhitened chronologies. Use of prewhitened chronologies removes autocorrelation. Use of lags in the model recognizes that too much or too little may have been removed.
2. GGFN note the observed flow series were not autocorrelated, and took reasonable steps to guard against building artificial autocorrelation into the reconstruction in the modeling.

j) Stop rule = F

1. Not specified in report how the second predictor was allowed to enter the model, or whether the second predictor was just assumed to be relevant. Gary Funkhouse (pers. comm.) says required a significant F-level for entry.

k) Residuals analysis = N*

1. No formal analysis of residuals is described in report. Unknown whether residuals normally distributed, autocorrelated, or whether they are correlated with the predicted values of flow (heteroscedasticity). Authors did check autocorrelation of gaged flows.

l) Have data?= Y

1. The annual (water year) reconstructed flows in acre-ft were provided by Gary Funkhouser an Excel spreadsheet. We converted those to average daily flow in cms (for processing), and have conversion routines to convert to cfs for reporting data and analyses to SRP.

Table 1. Assessment Matrix for Previous Tree-Ring Streamflow Reconstructions

				Accuracy[1e]								
Gage[a]	Reference[b]	Length of Recon.[c]	Model[d]	R^2	RE	Quality Issues[f]	Sites[g]	Data. Red.[h]	Autoc. Adjust.[i]	Stop rule[j]	Res. An.?[k]	Have Data?[l]
Colorado R. at Lees Ferry, AZ	Stockton & Jacoby (1976)	1520-1961	PCR	0.87	N	S,T	17	P	L	F	N*	Y
	Michaelsen et al. (1990)											
	Hidalgo et al. (2000)	~1500 – 1961*	PCR	0.82	0.99	S,T,C	6	P	L	V	N	Y
Colorado R. nr Cisco, UT	Stockton & Jacoby (1976)	1640-1962	PCR	0.92	N	S,T,V	4	P	L	F	Y*	N
Green R. nr Green River, UT	Stockton & Jacoby (1976)	1660-1962	PCR	0.80	N	S,T,V	7	P	L	F	Y*	N
San Juan R. nr Bluff, UT	Stockton & Jacoby (1976)	1661-1969	PCR	0.85	N	S,T,V	3	P	L	F	Y*	N
Salt R. nr Roosevelt, AZ	Smith & Stockton (1981)	1702-1979*	MLR	0.73	N	T	7	N	R	F*	N	N
Tonto Cr., abv Gun Creek, AZ	None—no reconstruction											
Verde R. blw Tangle Cr, AZ	Smith & Stockton (1981)	1702-1979*	MLR	0.74*	N	S,T	5	N	R	F*	N	Y
Salt+Verde+Tonto	Graybill et al. (in press)	572-1988	MLR	.72	.70	S,T,V	1-5	A	R/L	F	N*	Y
Gila R., at head of Safford Valley, AZ	Meko & Graybill (1995)	1663-1985	MLR	0.58*	0.52*	S,T,V	6	N	R/L	F	Y	Y
	Graybill et al. (in press)	534-1988	MLR	0.60	0.57	S,T,V	1-5	A	R/L	F	N*	Y

*=special conditions – refer to narrative section of appendix

KEY --

[a]**gage(es) to which reconstruction applies**

[b]**documentation (published or unpublished) for the reconstruction**

[c]**length of reconstructed flow series = start and end year**

[d]**model used for reconstruction.** MLR= multiple linear regression. PCR=principal components regression. CRG=canonical regression

[e]**accuracy of reconstruction and validation**. To allow comparison, regression R^2 and reduction of error statistic (RE) are used unless noted. N=no formal validation reported.

[f]**quality issues = possible problems with data quality.** S=sample-size deterioration over time, either in the number of trees making up the chronologies, or the number of chronologies used in the reconstruction. T=possible spurious low frequency variations or trends in reconstruction imparted in chronology development. C=coverage of runoff-producing regions inadequate or sparse, either for entire reconstruction or part of reconstruction

[g]**sites = number of tree-ring chronologies used in reconstruction**; number of chronologies used in the reconstruction

[h]**data reduction = method used to reduce redundancy in the tree-ring network:** A=arithmetic average over chronologies. P=principal components transformation of chronologies

[i]**autocorrelation adjustment = method used to deal with excessive autocorrelation in tree-ring chronologies**. R=use of ARMA residual chronologies as tree-ring data. L=use of lags in regression model. May have combinations (e.g., R/L)

[j]**stop rule = stopping rule for termination of entry of predictors in model.** F=minimum acceptable F-level for entry in stepwise. V=no gain in validation skill. N=none (all potential predictors included as final predictors)