APPENDIX 3G – DETAILS OF RECONSTRUCTION MODELING, GAGE G – SALT+TONTO RIVERS, ARIZONA

This reconstruction uses just the earliest (M1) sub-period model, which has tree-ring data starting in A.D. 1199. The predictand for modeling is water-year average daily flow in units of $\log_{10}$ cms.

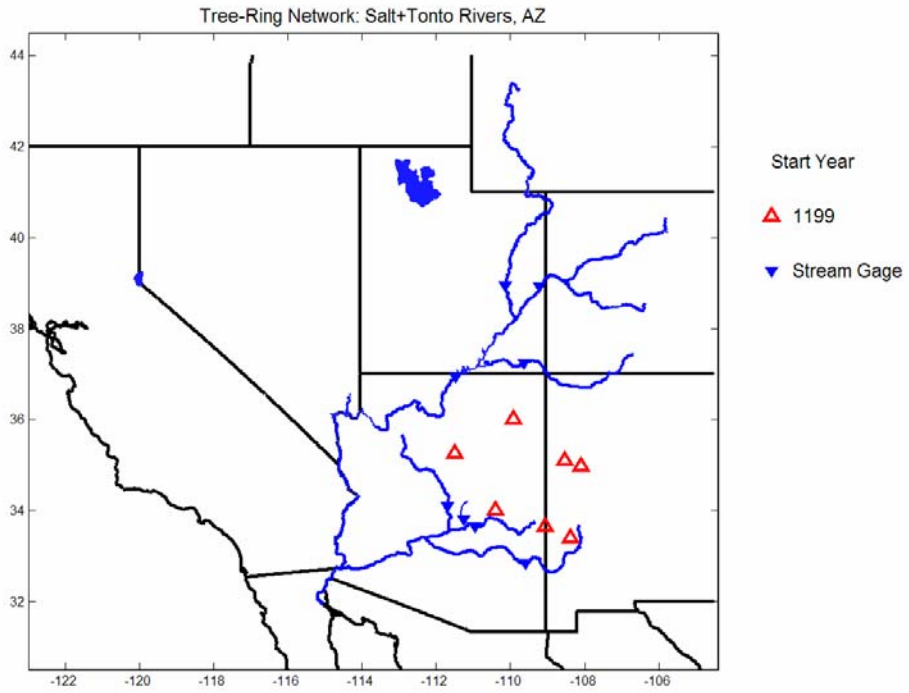

Figure 3G_1. Map showing tree-ring site locations for sub-period model used in reconstructing Salt+Tonto Rivers, Arizona.

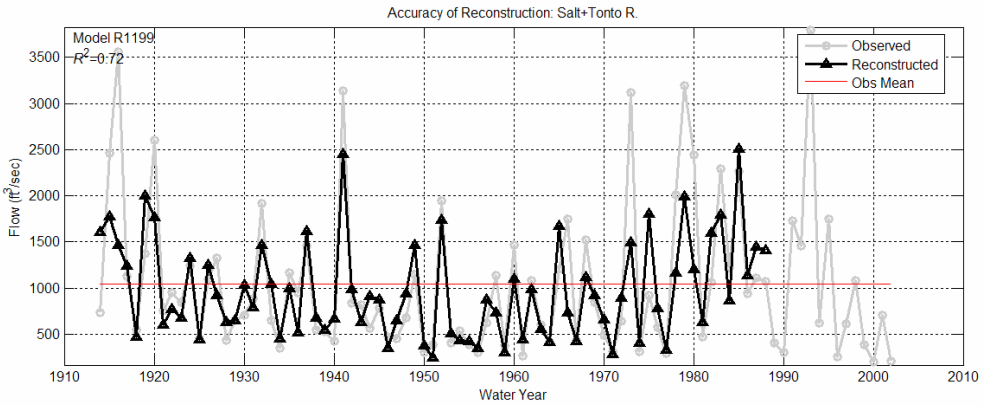

Figure 3G-2. Time series plots of observed and reconstructed flows for calibration period, Salt+Tonto Rivers, Arizona. Model R1199 indicates model allows reconstruction to A.D. 1199.

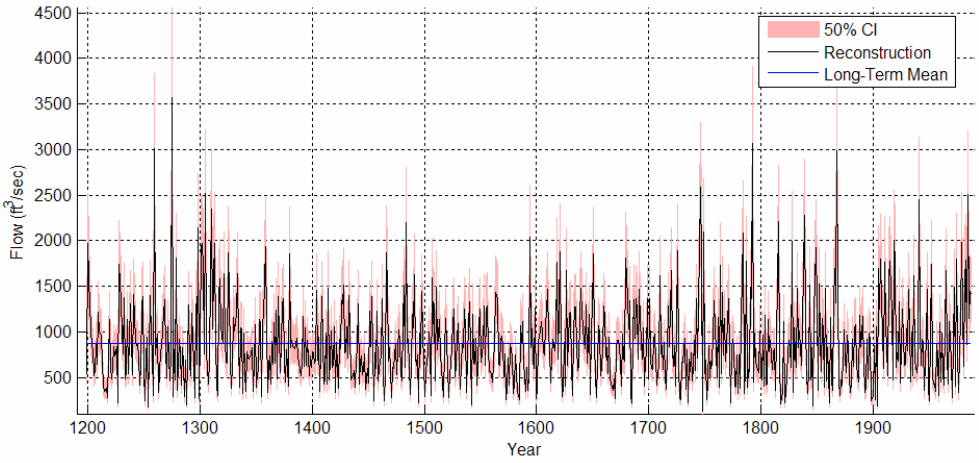

Figure 3G_3. Time series plot of reconstructed annual flows, Salt+Tonto Rivers, Arizona. Confidence interval based on root-mean-square error of cross-validation. Accuracy measured by root-mean-square error of cross-validation.

Table 3G_1. Summary of multi-site regression modeling for Salt+Tonto Rivers, Arizona.

		Calibration[3]			Validation[4]		
N[1]	Start[2]	Years	n-p-q	R^2adj	m	RE	RMSE
1	1199	1914-1988	7-4-1	0.72	3	0.70	0.1568

[1]Sub-period model number (1 is earliest)
[2]Start year of reconstruction period
[3]Calibration statistics:
  Years=calibration period
  n=number of chronologies
  p=number of potential predictors
  q=number of predictors in final model
  R^2adj = adjusted coefficient of determination
[4]Validation statistics (cross-validation)
  m = number of observations left out in "leave-m-out" cross-validation
  RE = reduction of error statistic
  RMSE = root-mean-square error of cross-validation (units of RMSE are same as units of the predictand in regression)

NOTES:
Predictand is log-transformed flow (log10 of the annual flow in cms)
Predictors = Principal components (covariance matrix) from PCA on full reconstruction + calibration period
Units of predictand in regression = log10(cms)
p-value of overall F for model < 1.0E-99

Table 3G_M1_1. Chronology listing and statistics on prewhitening, model M1199.

				LOCATION[5]			TIME COVERAGE[6]		AR[7]	
N[1]	CHRONOLOGY[2]	FILE[3]	SPECIES[4]	LAT	LON	EL(M)	START	END	p	var
1	Central Moun	swarchy	MANY	34.0	-110.4	1875	1066(1066)	1988	2	0.5
2	Flagstaff	swarchy	Many	35.3	-111.5	2130	571(571)	1988	0	0.0
3	Cebolleta Me	ad1000s	many	35.1	-108.6	2114	1000(1000)	1988	2	14.3
4	El Malpais	ad1000s	PSME	35.0	-108.1	2423	1000(877)	1988	3	11.3
5	Canyon de Ch	ad1000s	many	36.0	-109.9	1830	1000(591)	1988	3	10.8
6	Mount Graham	ad1000s	many	33.6	-109.1	2950	1000(1162)	1988	3	15.2
7	Black Mounta	bkm	many	33.4	-108.4	2710	1196(1196)	1988	3	11.2

[1]sequential site number
[2]short form of chronology name
[3]computer file (.crn) identifying chronology in ITRDB and elsewhere (e.g., ca528.crn is unique file at International Tree-Ring Data Bank). File “ad1000s” are chronologies from Ni et al. (2002).
[4]species code(see Appendix 2)
[5]latitude and longitude in decimal degrees; elevation in meters above sea level; N/A indicates information not available
[6]first year of standard chronology (first year sub-sample signal strength – see text -- exceeds 0.85); last year of chronology; N/A means not available
[7]order of autoregressive model used to prewhiten chronology, and percent chronology variance due to modeled autocorrelation

Table 3G_M1_2. Summary of single-site regression/reconstruction, model M1199.

N[1]	CHRONOLOGY[2]	REGRESSION MODEL[3] LAGS	R^2	F	RE[4] A	B
1	Central Moun	0,-1	0.57	95.6 ***	0.51	0.63
2	Flagstaff	0	0.37	42.5 ***	0.32	0.41
3	Cebolleta Me	0,-1	0.60	53.6 ***	0.53	0.65
4	El Malpais	0	0.41	50.6 ***	0.33	0.50
5	Canyon de Ch	0	0.65	132.0***	0.67	0.64
6	Mount Graham	0	0.42	51.9 ***	0.29	0.59
7	Black Mounta	0	0.36	40.1 ***	0.24	0.48

[1]sequential site number
[2]chronology name (truncated)
[3]regression modeling specifications and statistics:
LAGS = lags included on predictors
R^2 = variance explained by regression, adjusted
F = F-level and significance (*, **, *** indicate 0.05, 0.01 and 0.001 alpha-levels)
[4]Reduction of error statistic for split-sample validation;
A = validation on second half of data (calibration on first)
B = validation on first half of data (calibration on second)

Table 3G_M1_3. Summary of stepwise estimation of multi-site reconstruction, model M1199.

Step	Variables[1]	R^2adj	RE Statistic[2] A	B	cv	RMSEcv[3]	Residuals[4] r_1	T	N
1	1	0.72	0.74	0.67	0.70	0.1568	P	0	P

[1]Variables included as predictors in the model at the indicated step. Variables are principal components (covariance matrix) from PCA on full period of reconstruction and calibration. Variable 1 is PC#1, variable 2 is PC#2, and so forth.
[2]Reduction of error statistics from (A) calibration on 1914-1950 and validation on 1951-1988, (B) calibraton on 1951-1988 and validation on 1914-1950,(cv)cross-validation with 3 observations left out at each iteration
[3]Root-mean-square error of cross-validation, in log10(cms)
[4]Results of analysis of residuals: r_1 is Durbin-Watson(DW) test for first-order autocorrelation of residuals; T is test for significant slope in regression of residuals on time (trend); N is Lilliefors test for normality of residuals; "P" for DW and N tests indicates “pass”, or test statistic not significant at 0.05 alpha-level; 0 indicates slope of trend line not significant at 0.05 level, while - or + indicates significant negative or positive trend in residuals

Model Equation: constant term, coefficients, confidence interval, selected statistics:

Var	Coef	95% CI
Con	1.3325	(1.2962 1.3688)
X1	0.54746	(0.46735 0.62757)

R-squared = 0.71761
F-level = 185.508
sig = sig <1.0 E-99

Table 3B_M1_4. Weights[1] of chronologies in principal components and final regression.

N	CHRONOLOGY	LOADINGS X1	W	W*
1	Central Moun	0.444	0.1841	0.83
2	Flagstaff	0.329	0.1105	0.50
3	Cebolleta Me	0.474	0.2025	0.91
4	El Malpais	0.283	0.1001	0.45
5	Canyon de Ch	0.503	0.2231	1.00
6	Mount Graham	0.265	0.0937	0.42
7	Black Mounta	0.258	0.0843	0.38

[1]Columns X1, X2,... are the principal component loadings on the chronologies. X1 denotes PC1, X2 denotes PC1, and so forth. Final, or multi-site, reconstruction was generated by regression of flow on the PC scores. The final reconstruction can be generated by applying the estimated regression equation to those PC scores. The final reconstruction can alternatively be generated from the individual filtered, scaled chronologies themselves. To generate the final from the chronologies, the applicable weights are in column "W". ("W*" are the same weights proportionally scaled so that the largest weight is 1.0.) The weights W and W* measure the relative importance of the individual chronologies to the final reconstruction. Steps for generating reconstruction from original chronologies:

1) filter and scale the original chronologies into single-site (ss) reconstructions as described in the text
2) convert ss reconstructions to Z scores, using calibration period means and standard deviations
3) multiply those z-score series by the regression weights in next-to-last column (W) above, and sum the weighted series
4) multiply resulting series by calibration-period standard deviation of flow and add the calibration-period mean observed flow