APPENDIX 3H – DETAILS OF RECONSTRUCTION MODELING, GAGE H – VERDE RIVER, ARIZONA

This reconstruction uses two sub-period models (M1 and M2), with data starting in A.D. 1202 and A.D. 1666. The predictand for modeling is water-year average daily flow in units of $\log_{10}$ cms.

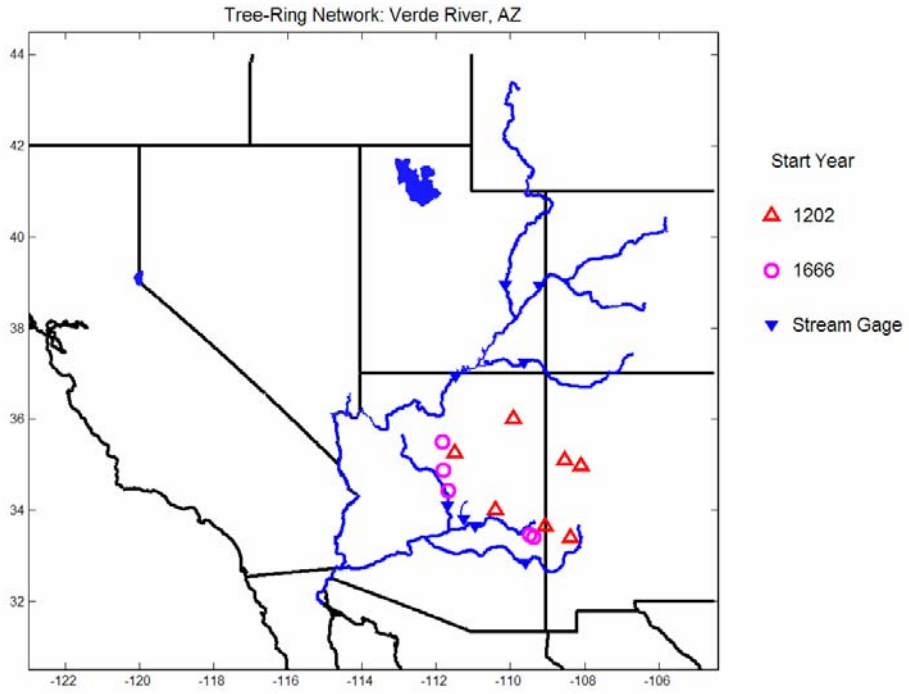

Figure 3H_1. Map showing tree-ring site locations for sub-period models used in reconstructing Verde River, Arizona.

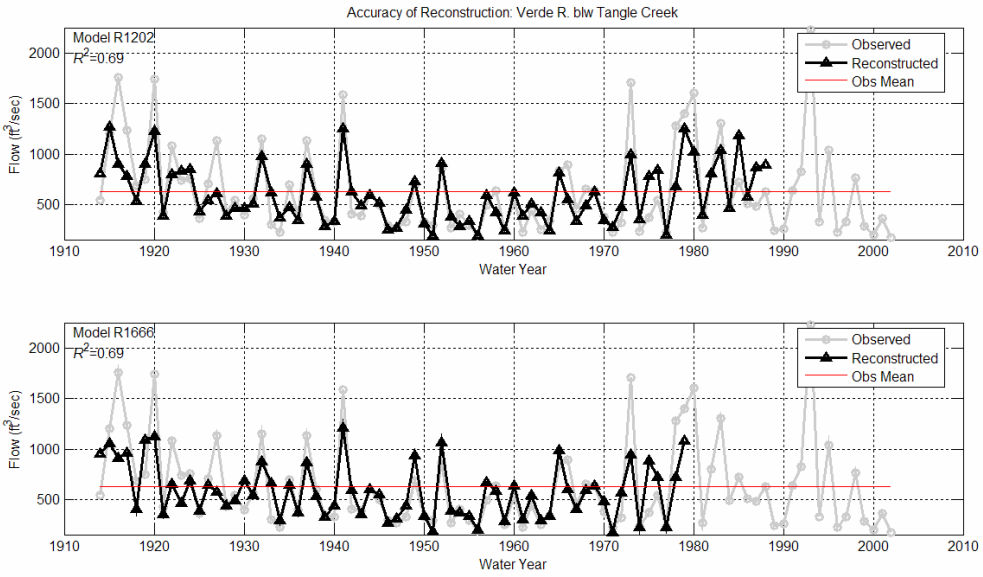

Figure 3H-2. Time series plots of observed and reconstructed flows for calibration period, Verde River, Arizona. Top: earliest model, allowing reconstruction to A.D. 1202. Bottom: most recent model, allowing reconstruction to A.D. 1666.

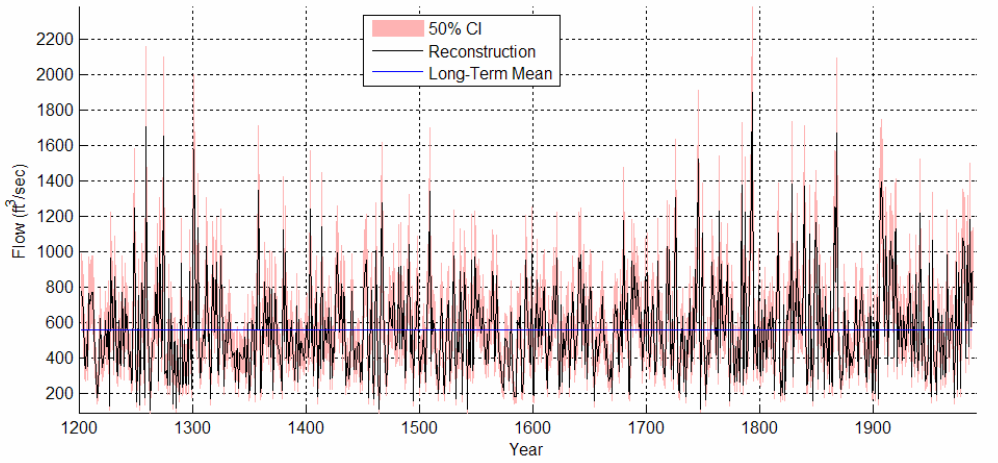

Figure 3H_3. Time series plot of reconstructed annual flows, Verde River, Arizona. Confidence interval based on root-mean-square error of cross-validation. Reconstruction for given interval of time based on the most accurate sub-period reconstruction available for that period. Accuracy measured by root-mean-square error of cross-validation.

Table 3H_1. Summary of multi-site regression modeling for Verde River, Arizona.

		Calibration[3]			Validation[4]		
N[1]	Start[2]	Years	n-p-q	R^2adj	m	RE	RMSE
1	1202	1914-1988	7-4-4	0.68	7	0.63	0.1531
2*	1574	1914-1980	14-4-1	0.65	7	0.63	0.1578
3	1666	1914-1979	19-2-1	0.69	7	0.67	0.1453

[1]Sub-period model number (1 is earliest, * indicates a model not used in final reconstruction)
[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-10 transformed flow
Predictors = Principal components (covariance matrix) from PCA on full reconstruction + calibration period
Units of predictand in regression = log10 cms
Maximum p-value of overall F for any model < 1.0E-99

Table 3H_M1_1. Chronology listing and statistics on prewhitening, model M1202.

				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 3H_M1_2. Summary of single-site regression/reconstruction, model M1202.

N[1]	CHRONOLOGY[2]	REGRESSION MODEL[3] LAGS	R^2	F	RE[4] A	B
1	Central Moun	0,-1	0.59	51.3***	0.59	0.57
2	Flagstaff	0	0.38	43.7***	0.31	0.42
3	Cebolleta Me	0	0.37	42.8***	0.27	0.46
4	El Malpais	0	0.23	21.9***	0.15	0.34
5	Canyon de Ch	0	0.57	93.9***	0.65	0.55
6	Mount Graham	0,-1	0.31	16.9***	0.19	0.47
7	Black Mounta	0,-1,-3	0.29	10.5***	0.21	0.38

[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 3H_M1_3. Summary of stepwise estimation of multi-site reconstruction, model M1202.

Step	Variables[1]	R^2adj	RE Statistic[2] A	B	cv	RMSEcv[3]	Residuals[4] r_1	T	N
1	1	0.62	0.66	0.61	0.60	0.1599	P	0	P
2	1,3	0.64	0.66	0.64	0.61	0.1570	P	0	P
3	1,3,4	0.66	0.67	0.69	0.63	0.1546	P	0	P
4	1,2,3,4	0.68	0.60	0.71	0.63	0.1531	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 7 observations left out at each iteration
[3]Root-mean-square error of cross-validation, in log10 cms units
[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	1.1398	(1.1045	1.1752)
X1	0.61041	(0.50971	0.71111)
X2	-0.28001	(-0.54905	-0.01096)
X3	-0.40738	(-0.69755	-0.11722)
X4	-0.57537	(-0.96318	-0.18756)

R-squared = 0.69031
F-level = 39.0086
sig <1.0 E-99
RMSEcv = 0.15307 (before backtransforming, if applicable)

Table 3H_M1_4. Weights[1] of chronologies in principal components and final regression.

		LOADINGS					
N1	CHRONOLOGY	X1	X2	X3	X4	W	W*
1	Central Moun	0.532	0.364	-0.449	-0.161	0.3849	0.70
2	Flagstaff	0.377	-0.172	-0.552	0.510	0.1299	0.24
3	Cebolleta Me	0.382	-0.017	0.547	0.589	-0.1989	-0.36
4	El Malpais	0.218	0.120	0.330	0.139	-0.0559	-0.10
5	Canyon de Ch	0.520	-0.671	0.164	-0.497	0.5489	1.00
6	Mount Graham	0.250	0.378	0.165	-0.190	0.0505	0.09
7	Black Mounta	0.221	0.479	0.180	-0.255	0.0417	0.08

[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

Table 3H_M2_1. Chronology listing and statistics on prewhitening, model M1666.

N[1]	CHRONOLOGY[2]	FILE[3]	SPECIES[4]	LOCATION[5] LAT	LON	EL(M)	TIME COVERAGE[6] START	END	AR[7] p	var
1	Dry Creek	az517	PIED	34.9	-111.8	1377	1630(1650)	1986	2	18.8
2	Hackberry Ca	hackberr	PIED	34.4	-111.7	-9999	1664(1664)	1980	2	4.9
3	Eagle Creek	az518	PIED	33.5	-109.5	1694	1640(1738)	1987	2	17.7
4	Slate Mounta	az543	PIPO	35.5	-111.8	2194	1600(1656)	1985	2	40.2
5	Rose Peak	az541	PIPO	33.4	-109.4	2316	1660(1678)	1987	2	29.9
6	Walnut Canyo	az547	PIPO	35.2	-111.5	2057	1420(1601)	1987	3	28.9
7	Central Moun	swarchy	MANY	34.0	-110.4	1875	1066(1066)	1988	2	0.5
8	Beaver Creek	az513	PIPO	33.7	-109.2	2390	1559(1595)	1986	3	21.2
9	Flagstaff	swarchy	Many	35.3	-111.5	2130	571(571)	1988	0	0.0
10	Alpine	alpiners	PIED	33.9	-109.0	-9999	1457(1457)	1980	2	10.2
11	Cebolleta Me	ad1000s	many	35.1	-108.6	2114	1000(1000)	1988	2	14.3
12	El Malpais	ad1000s	PSME	35.0	-108.1	2423	1000(877)	1988	3	11.3
13	Canyon de Ch	ad1000s	many	36.0	-109.9	1830	1000(591)	1988	3	10.8
14	Mount Graham	ad1000s	many	33.6	-109.1	2950	1000(1162)	1988	3	15.2
15	Black Mounta	bkm	many	33.4	-108.4	2710	1196(1196)	1988	3	11.2
16	Red Butte, A	redbutte	PIED	35.8	-112.1	1940	1448(1448)	1989	2	22.1
17	Walnut Canyo	walnuted	PIED	35.4	-111.6	-9999	1309(1309)	1987	3	24.3
18	Green Mounta	AZ520	PIPO	32.4	-110.7	-9999	1460(1633)	1986	3	32.5
19	Reef of Rock	AZ557	PSME	32.5	-110.8	-9999	1321(1435)	1998	3	22.1

[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 key on 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 indicated information not available
[7]order of autoregressive model used to prewhiten chronology, and percent chronology variance due to modeled autocorrelation

Table 3H_M2_2. Summary of single-site regression/reconstruction, model M1666

		REGRESSION MODEL[3]			RE[4]	
N[1]	CHRONOLOGY[2]	LAGS	R^2	F	A	B
1	Dry Creek	0,-1	0.62	56.5***	0.55	0.75
2	Hackberry Ca	0	0.48	57.9***	0.39	0.58
3	Eagle Creek	0,-1	0.28	28.2***	0.18	0.39
4	Slate Mounta	0	0.33	34.7***	0.28	0.35
5	Rose Peak	0,-1	0.30	29.9***	0.22	0.39
6	Walnut Canyo	0	0.38	43.9***	0.30	0.42
7	Central Moun	0,-1	0.59	51.3***	0.59	0.57
8	Beaver Creek	0,-1,-3	0.21	9.6***	0.17	0.28
9	Flagstaff	0	0.38	43.7***	0.31	0.42
10	Alpine	0,-1	0.33	16.6***	0.32	0.37
11	Cebolleta Me	0	0.37	42.8***	0.27	0.46
12	El Malpais	0	0.23	21.9***	0.15	0.34
13	Canyon de Ch	0	0.57	93.9***	0.65	0.55
14	Mount Graham	0,-1	0.31	16.9***	0.19	0.47
15	Black Mounta	0,-1,-3	0.29	10.5***	0.21	0.38
16	Red Butte, A	0	0.37	42.1***	0.27	0.47
17	Walnut Canyo	0	0.39	45.3***	0.32	0.42
18	Green Mounta	0,-3	0.23	21.0***	0.35	0.12
19	Reef of Rock	0,-1	0.10	9.3**	0.12	0.10

[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 3H_M1_3. Summary of stepwise estimation of multi-site reconstruction, model M1666.

			RE Statistic[2]				Residuals[4]		
Step	Variables[1]	R^2adj	A	B	cv	RMSEcv[3]	r_1	T	N
1	1	0.69	0.76	0.68	0.67	0.1453	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 1914-1946 and validation on 1947-1979, (B) calibraton on 1947-1979 and validation on 1914-1946, (cv) cross-validation with 7 observations left out at each iteration
[3]Root-mean-square error of cross-validation, in units log10 cms
[4]Results of analysis of residuals: r_1 is Durbin-Watson (DW) test for first-order autocorrelation of residual; 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 test 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	1.1666	(1.131	1.2022)
X1	0.40989	(0.34061	0.47917)

R-squared = 0.68582
F-level = 139.7071
sig <1.0E-99

Table 3H_M2_4. Weights[1] of chronologies in principal components and final regression.

		LOADINGS		
N1	CHRONOLOGY	X1	W	W*
1	Dry Creek	0.343	0.1112	0.87
2	Hackberry Ca	0.242	0.0682	0.54
3	Eagle Creek	0.166	0.0374	0.29
4	Slate Mounta	0.188	0.0459	0.36
5	Rose Peak	0.169	0.0393	0.31
6	Walnut Canyo	0.286	0.0751	0.59
7	Central Moun	0.396	0.1274	1.00
8	Beaver Creek	0.103	0.0182	0.14
9	Flagstaff	0.275	0.0712	0.56
10	Alpine	0.221	0.0533	0.42
11	Cebolleta Me	0.227	0.0559	0.44
12	El Malpais	0.131	0.0254	0.20
13	Canyon de Ch	0.304	0.0900	0.71
14	Mount Graham	0.174	0.0405	0.32
15	Black Mounta	0.136	0.0314	0.25
16	Red Butte, A	0.218	0.0522	0.41
17	Walnut Canyo	0.282	0.0747	0.59
18	Green Mounta	0.121	0.0251	0.20
19	Reef of Rock	0.082	0.0112	0.09

[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