APPENDIX 3F – DETAILS OF RECONSTRUCTION MODELING
GAGE F – SAN JUAN RIVER NEAR BLUFF, UTAH

This reconstruction uses two sub-period models (M1 and M2), with data starting in A.D. 1280 and A.D. 1514. The predictand for modeling is water-year average daily flow in units of cms.

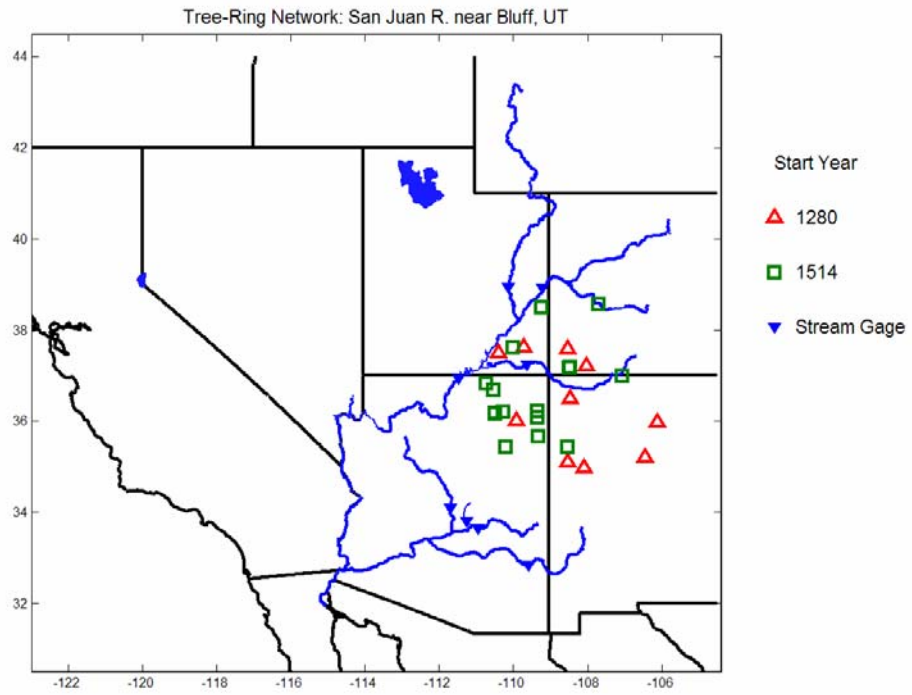

Figure 3F_1. Map showing tree-ring site locations for sub-period models used in reconstructing San Juan River near Bluff, Utah.

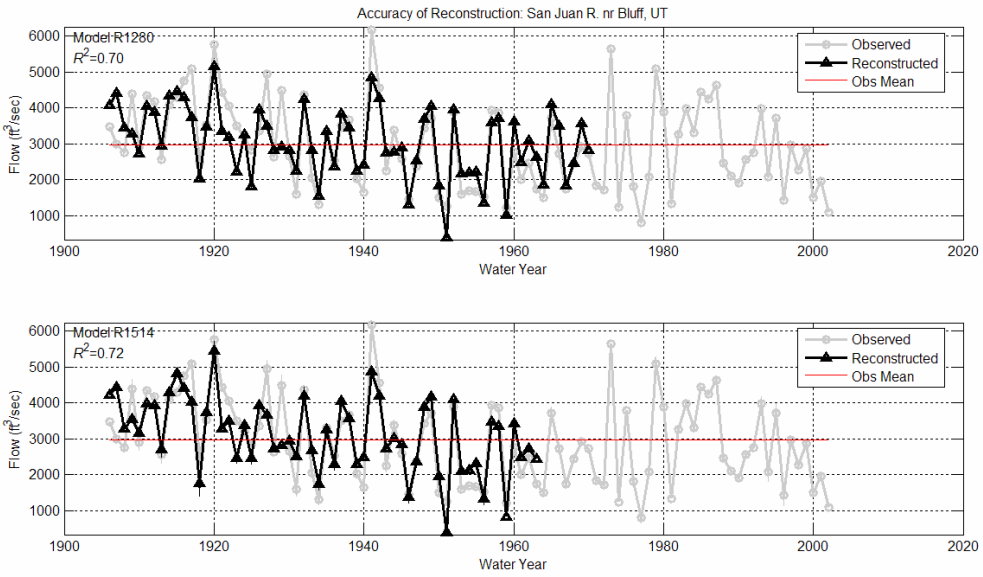

Figure 3E-2. Time series plots of observed and reconstructed flows for calibration period, San Juan River near Bluff, Utah. Top: earliest model, allowing reconstruction to A.D. 1280. Bottom: most recent model, allowing reconstruction to A.D. 1514.

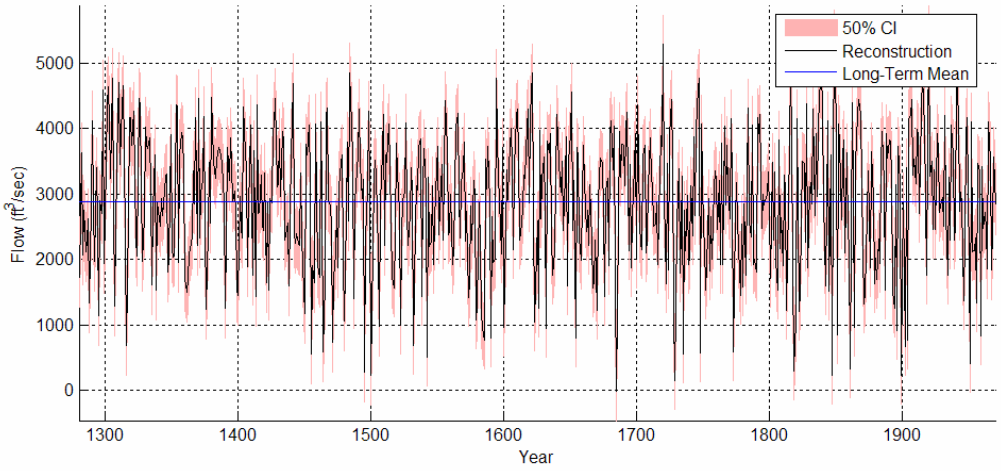

Figure 3F_3. Time series plot of reconstructed annual flows, San Juan River near Bluff, Utah. Confidence interval based on root-mean-square error of cross-validation. Reconstruction for given interval of time based on the most accurate sub-period reconstruction available for that period. Accuracy measured by root-mean-square error of cross-validation.

Table 3F_1. Summary of multi-site regression modeling for San Juan River near Bluff, Utah.

		Calibration[3]			Validation[4]		
N[1]	Start[2]	Years	n-p-q	R^2adj	m	RE	RMSE
1	1280	1906-1970	10-3-1	0.70	5	0.66	19.1147
2	1514	1906-1963	27-2-1	0.72	9	0.70	18.4580

[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 flow (not transformed)
Predictors = Principal components (covariance matrix) from PCA on full reconstruction + calibration period
Units of predictand in regression = cms
Maximum p-value of overall F for any model < 1.0E-99

Table 3F_M1_1. Chronology listing and statistics on prewhitening, model M1280.

				LOCATION[5]			TIME COVERAGE[6]		AR[7]	
N[1]	CHRONOLOGY[2]	FILE[3]	SPECIES[4]	LAT	LON	EL(M)	START	END	p	var
1	Cebolleta Me	ad1000s	many	35.1	-108.6	2114	1000(1000)	1988	2	14.3
2	El Malpais	ad1000s	PSME	35.0	-108.1	2423	1000(877)	1988	3	11.3
3	Gobernador P	ad1000s	many	36.5	-108.5	2195	1000(749)	1988	3	7.5
4	Canyon de Ch	ad1000s	many	36.0	-109.9	1830	1000(591)	1988	3	10.8
5	Durango PLUS	ad1000s	many	37.2	-108.0	2073	1000(804)	1988	3	17.2
6	Natural Brid	ad1000s	many	37.5	-110.4	1859	1000(510)	1988	3	7.7
7	Chama Valley	ad1000s	many	36.0	-106.1	2137	1000(1000)	1988	3	6.2
8	Sandia Crest	ad1000s	PIFL	35.2	-106.5	3048	1000(824)	1988	3	9.5
9	Dolores	CO067	PIED	37.6	-108.6	2195	1270(1270)	1978	3	8.3
10	Milk Ranch P	UT024	PIED	37.6	-109.7	2286	1276(N/A)	1970	3	18.1

[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 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 3F_M1_2. Summary of single-site regression/reconstruction, model M1280.

N[1]	CHRONOLOGY[2]	REGRESSION MODEL[3] LAGS	R^2	F	RE[4] A	B
1	Cebolleta Me	0,-1	0.49	38.7***	0.36	0.62
2	El Malpais	0,-1	0.34	20.8***	0.19	0.60
3	Gobernador P	0,-1,1	0.60	40.4***	0.46	0.80
4	Canyon de Ch	0	0.54	92.4***	0.46	0.66
5	Durango PLUS	0,-1	0.57	54.0***	0.59	0.71
6	Natural Brid	0,-1	0.56	51.1***	0.57	0.63
7	Chama Valley	0,-1	0.42	29.0***	0.23	0.65
8	Sandia Crest	0,-1	0.17	8.9***	0.09	0.32
9	Dolores	0,-1	0.51	36.8***	0.55	0.59
10	Milk Ranch P	0,-1	0.55	38.6***	0.43	0.61

[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 3F_M1_3. Summary of stepwise estimation of multi-site reconstruction, model M1280.

Step	Variables[1]	R^2adj	RE Statistic[2] A	B	cv	RMSEcv[3]	Residuals[4] r_1	T	N
1	1	0.70	0.79	0.65	0.66	19.1147	P	-	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 1906-1937 and validation on 1938-1970, (B) calibraton on 1938-1970 and validation on 1906-1937,(cv)cross-validation with 5 observations left out at each iteration
[3]Root-mean-square error of cross-validation, in 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	80.7952	(76.1271	85.4634)
X1	0.465285	(0.387744	0.542826)

R-squared = 0.69534
F-level = 143.7864
sig <1.0 E-99

Table 3F_M1_4. Weights[1] of chronologies in principal components and final regression.

N	CHRONOLOGY	LOADINGS X1	W	W*
1	Cebolleta Me	0.348	0.1112	0.80
2	El Malpais	0.211	0.0578	0.42
3	Gobernador P	0.324	0.1143	0.82
4	Canyon de Ch	0.373	0.1225	0.88
5	Durango PLUS	0.375	0.1243	0.90
6	Natural Brid	0.323	0.1033	0.74
7	Chama Valley	0.220	0.0635	0.46
8	Sandia Crest	0.108	0.0226	0.16
9	Dolores	0.358	0.1169	0.84
10	Milk Ranch P	0.397	0.1387	1.00

[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 3F_M2_1. Chronology listing and statistics on prewhitening, model M1514.

N[1]	CHRONOLOGY[2]	FILE[3]	SPECIES[4]	LOCATION[5] LAT	LON	EL(M)	TIME COVERAGE[6] START	END	AR[7] p	var
1	Cebolleta Me	ad1000s	many	35.1	-108.6	2114	1000(1000)	1988	2	14.3
2	El Malpais	ad1000s	PSME	35.0	-108.1	2423	1000(877)	1988	3	11.3
3	Gobernador P	ad1000s	many	36.5	-108.5	2195	1000(749)	1988	3	7.5
4	Canyon de Ch	ad1000s	many	36.0	-109.9	1830	1000(591)	1988	3	10.8
5	Durango PLUS	ad1000s	many	37.2	-108.0	2073	1000(804)	1988	3	17.2
6	Natural Brid	ad1000s	many	37.5	-110.4	1859	1000(510)	1988	3	7.7
7	Chama Valley	ad1000s	many	36.0	-106.1	2137	1000(1000)	1988	3	6.2
8	Sandia Crest	ad1000s	PIFL	35.2	-106.5	3048	1000(824)	1988	3	9.5
9	Hard Rock	AZ066	PIED	36.2	-110.3	1920	1380(N/A)	1967	2	18.6
10	Tseh-Ya-Kin	AZ083	PSME	36.2	-109.3	1951	1500(N/A)	1971	3	37.9
11	Shonto Plate	AZ086	PIED	36.8	-110.7	2134	1365(1373)	1971	2	3.4
12	Tsegi Point	AZ102	PIED	36.7	-110.5	2196	1490(1448)	1972	3	11.0
13	Dead Juniper	AZ103	JUSP	36.2	-110.5	1920	1310(N/A)	1972	2	9.5
14	Snow Bowl	AZ553	PCEN	35.4	-110.2	3150	1453(1696)	1983	2	36.5
15	Schulman Old	co021	PSME	37.2	-108.5	2103	1400(1204)	1963	3	13.1
16	Black Canyon	co053	PSME	38.6	-107.7	2426	1478(1634)	1964	3	30.6
17	Dolores	CO066	PSME	37.6	-108.6	2286	1457(N/A)	1978	2	33.4
18	Dolores	CO067	PIED	37.6	-108.6	2195	1270(N/A)	1978	3	8.3
19	Spruce Canyo	CO509	PSME	37.2	-108.5	2115	1373(1389)	1978	2	16.3
20	Cross Canyon	crosscan	PIED	35.7	-109.3	N/A	1512(N/A)	1989	1	18.0
21	Gambler-Uppe	gamdin	PIED	36.2	-110.5	N/A	1400(N/A)	1983	1	14.3
22	Fort Wingate	NM031	PIED	35.4	-108.5	2268	1478(N/A)	1972	1	8.7
23	Ditch Canyon	NM503e	PSME	37.0	-107.1	2073	1487(N/A)	1978	2	16.6
24	Spider rock,	spider	PSME	36.1	-109.3	N/A	1376(N/A)	1989	2	26.6
25	La Sal Mount	ut018	PIED	38.5	-109.3	2323	1489(1597)	1972	3	20.9
26	White Canyon	UT023	PSME	37.6	-110.0	1859	1347(N/A)	1972	2	25.9
27	Milk Ranch P	UT024	PIED	37.6	-109.7	2286	1276(N/A)	1970	3	18.1

[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 3F_M2_2. Summary of single-site regression/reconstruction, model M1514

N[1]	CHRONOLOGY[2]	REGRESSION MODEL[3] LAGS	R^2	F	RE[4] A	RE[4] B
1	Cebolleta Me	0,-1	0.49	38.7***	0.36	0.62
2	El Malpais	0,-1	0.34	20.8***	0.19	0.60
3	Gobernador P	0,-1,1	0.60	40.4***	0.46	0.80
4	Canyon de Ch	0	0.54	92.4***	0.46	0.66
5	Durango PLUS	0,-1	0.57	54.0***	0.59	0.71
6	Natural Brid	0,-1	0.56	51.1***	0.57	0.63
7	Chama Valley	0,-1	0.42	29.0***	0.23	0.65
8	Sandia Crest	0,-1	0.17	8.9***	0.09	0.32
9	Hard Rock	0	0.43	44.8***	0.24	0.55
10	Tseh-Ya-Kin	0,-1	0.39	20.8***	0.26	0.55
11	Shonto Plate	0	0.61	97.6***	0.42	0.64
12	Tsegi Point	0	0.54	75.3***	0.29	0.61
13	Dead Juniper	0	0.40	42.3***	0.14	0.71
14	Snow Bowl	0	0.13	10.9**	0.09	0.30
15	Schulman Old	0,-1,-3	0.62	30.0***	0.58	0.58
16	Black Canyon	0	0.18	11.9**	0.06	0.26
17	Dolores	0	0.57	92.0***	0.44	0.64
18	Dolores	0,-1	0.51	36.8***	0.55	0.59
19	Spruce Canyo	0	0.45	56.5***	0.31	0.65
20	Cross Canyon	0,-1	0.38	25.5***	0.23	0.57
21	Gambler-Uppe	0	0.41	52.7***	0.33	0.46
22	Fort Wingate	0,-1	0.53	36.6***	0.23	0.68
23	Ditch Canyon	0,-1	0.61	54.7***	0.44	0.74
24	Spider rock,	0,-1	0.49	39.5***	0.36	0.59
25	La Sal Mount	0	0.31	29.2***	0.30	0.33
26	White Canyon	0	0.41	44.5***	0.35	0.45
27	Milk Ranch P	0,-1	0.55	38.6***	0.43	0.61

[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 3F_M1_3. Summary of stepwise estimation of multi-site reconstruction, model M1514.

			RE Statistic[2]				Residuals[4]		
Step	Variables[1]	R^2adj	A	B	cv	RMSEcv[3]	r_1	T	N
1	1	0.72	0.84	0.71	0.70	18.4580	P	0	F

[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 1906-1934 and validation on 1935-1963, (B) calibraton on 1935-1963 and validation on 1906-1934, (cv) cross-validation with 9 observations left out at each iteration
[3]Root-mean-square error of cross-validation, in 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	80.9668	(76.0635 85.8701)
X1	0.302677	(0.252539 0.352815)

R-squared = 0.72311
F-level = 146.2498
sig <1.0E-99

Table 3F_M2_4. Weights[1] of chronologies in principal components and final regression.

		LOADINGS		
N1	CHRONOLOGY	X1	W	W*
1	Cebolleta Me	0.210	0.0433	0.64
2	El Malpais	0.130	0.0229	0.34
3	Gobernador P	0.205	0.0458	0.68
4	Canyon de Ch	0.244	0.0513	0.76
5	Durango PLUS	0.222	0.0466	0.69
6	Natural Brid	0.216	0.0447	0.66
7	Chama Valley	0.137	0.0257	0.38
8	Sandia Crest	0.063	0.0083	0.12
9	Hard Rock	0.186	0.0371	0.55
10	Tseh-Ya-Kin	0.167	0.0322	0.48
11	Shonto Plate	0.259	0.0602	0.89
12	Tsegi Point	0.226	0.0478	0.71
13	Dead Juniper	0.170	0.0325	0.48
14	Snow Bowl	0.057	0.0060	0.09
15	Schulman Old	0.210	0.0500	0.74
16	Black Canyon	0.069	0.0088	0.13
17	Dolores	0.175	0.0408	0.61
18	Dolores	0.212	0.0442	0.66
19	Spruce Canyo	0.188	0.0414	0.61
20	Cross Canyon	0.180	0.0331	0.49
21	Gambler-Uppe	0.221	0.0439	0.65
22	Fort Wingate	0.219	0.0457	0.68
23	Ditch Canyon	0.265	0.0674	1.00
24	Spider rock,	0.220	0.0488	0.72
25	La Sal Mount	0.115	0.0189	0.28
26	White Canyon	0.160	0.0314	0.47
27	Milk Ranch P	0.244	0.0552	0.82

[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