

SCIENTIFIC REPORTS



OPEN

Role of Pacific SSTs in improving reconstructed streamflow over the coterminous US

Sudarshana Mukhopadhyay, Jason M. Patskoski & A. Sankarasubramanian 

Reconstructed annual streamflows using tree-ring chronologies provide useful information on moisture availability during the pre-historic period, but they have limitations in estimating high flows due to the upper bound on soil water holding capacity and trees' metabolic growth limits. We propose a hybrid approach that uses tree-ring chronologies and climatic indices for improving high flows in 301 basins whose annual streamflows are modulated by ENSO and/or PDO. The hybrid decomposition approach relies on separating the moisture supply into the basin as outside-the-region moisture and within-the-region moisture with the former being estimated by SST indices and the latter being estimated by tree-ring chronologies. Analyses over the 301 stations from coterminous US show that the proposed approach improves the high flows and improves the overall error in the reconstructed streamflows. Potential utility of the improved reconstructed annual streamflows with improved high flows is also discussed.

Annual streamflow records provide crucial information on the available water for various uses, including water supply and irrigation and for designing reservoir size. Given little observed streamflow data is available before the 1930s in the United States, studies focused on gaining insights on annual moisture supply prior to the observational period using paleo-information¹⁻⁴. The high dependency between moisture availability and tree-ring chronologies have provided information on wet and dry periods by reconstructing streamflows over pre-instrumental periods⁵⁻⁸. While tree-ring chronologies have good skill in low flow estimation⁹⁻¹¹, they have limitations in estimating high flows. Due to the upper bound on water holding capacity of soil, trees' metabolic growth limits and increased cloud cover during wet years, high flows are not properly estimated using tree-ring chronologies^{3,12-14}. Studies have shown that incorporating sea surface temperature (SSTs) along with tree-ring chronologies resulted in improved reconstruction of streamflow and wildfire synchrony¹⁵⁻¹⁷. Recent progress in understanding ocean-atmosphere interactions and the ensued teleconnection shows that there are well organized modes of inter-annual and interdecadal variability in climate that modulate these dominant moisture delivery pathways and has significant projections on continental scale rainfall and streamflow patterns^{18,19}. Interannual modes such as the El Niño-Southern Oscillation (ENSO) resulting from sea surface conditions in the tropical Pacific Ocean primarily determine the interannual variability in precipitation over North America^{20,21}. There are also other dominant decadal and interdecadal climatic modes such as Pacific Decadal Oscillation (PDO) and North Atlantic Oscillation (NAO) that putatively govern the interannual variability in climate over the North America²². To reduce the limitation of tree rings in reconstructing high flows²², proposed a hybrid reconstruction methodology using SST and tree-ring chronologies building upon the teleconnection signal with exogenous climatic conditions²². suggested separating the moisture contributed from exogenous climatic conditions (e.g., SSTs) as “outside-the-region moisture supply” and moisture contributed from the same hydroclimatic region as “within-the-region moisture supply”. Since the study²² focused on the Southeast US (SEUS), outside-the-region moisture transport was estimated by El-Niño Southern Oscillation (ENSO) Index, Niño3.4, the within-the-region moisture contribution was estimated by tree-ring chronologies. Given the influence of outside-the-region moisture supply on streamflow is separately estimated by exogenous SST conditions, it partially improves high flow estimation that is contributed by SST sources. The hybrid reconstruction approach, relying on utilizing multi-proxies (i.e., SST and tree-ring chronologies) for streamflow reconstruction, by²² improved overall skill and resulted in better above-normal flow estimates in comparison to the traditional tree-ring reconstructions in eight stations across

Department of Civil, Construction and Environmental Engineering North Carolina State University, Raleigh, NC, 27695-7908, USA. Correspondence and requests for materials should be addressed to A.S. (email: sankar_arumugam@ncsu.edu)

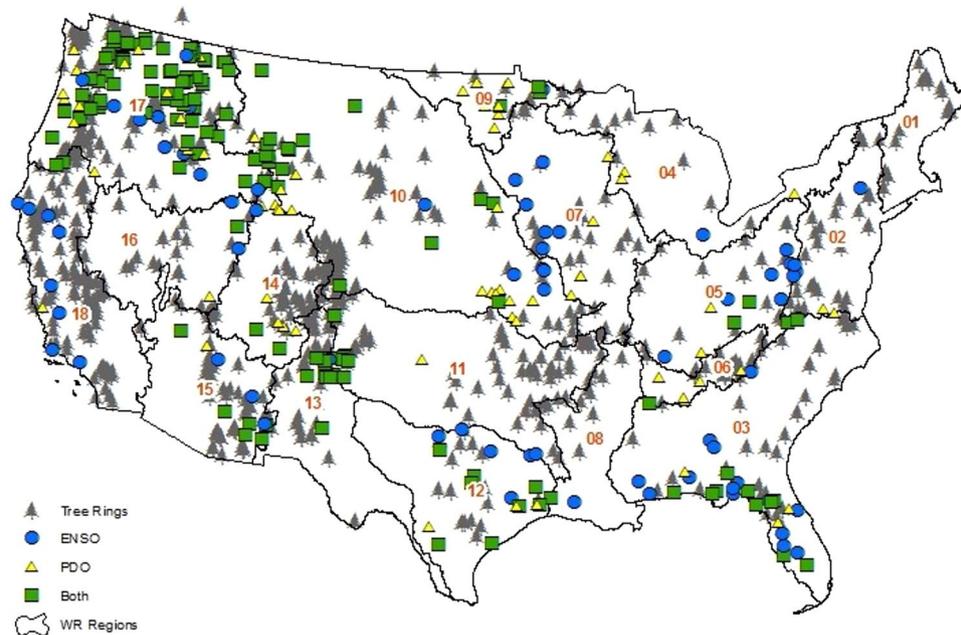


Figure 1. Locations of HCDN stations and tree-ring chronology sites across 18 water resources regions (orange) used in this study with blue circles (yellow triangles) [green squares] indicating the basins significantly influenced by ENSO (PDO) [ENSO and PDO]. This map is created using software suite ArcGIS 10.2.2 for Desktop, version number 10.2.2.3552 (<http://www.esri.com/en/arcgis/products/arcgis-pro/overview>).

the SEUS, which is a region with limited tree-ring chronologies. We perform here a national-scale hybrid streamflow reconstruction for 301 stream gauges (Fig. 1) that are significantly influenced by low-frequency oscillatory signals over the coterminous United States (CONUS). In this analysis the SST streamflow components are identified using regression while²² used Singular Spectrum Analysis (SSA) for the same. Given that SSA approach requires low-frequency mode identification for separating SST signals, such an approach will be difficult in the current study which considers PDO that has a periodicity of 16–20 years. The current approach is such simpler by explaining the role of SST predictors in influencing the Box-Cox transformed streamflow.

Since the hybrid streamflow reconstruction is at the national scale, both ENSO and Pacific Decadal Oscillation (PDO) conditions are used as SST conditions that have been shown to significantly affect moisture transport over the CONUS^{21,23,24}. Details on the variability explained by ENSO and PDO on the annual flows of the selected Hydro-Climatic Data Network (HCDN) basins is given in Supplemental Information (SI) (Figs SI-1 and SI-2). For each HCDN basin, tree ring and SST predictors are chosen based on correlation, proximity and length of data (See Data and Methods (DM) and SI). Streamflow at each site is reconstructed using two models: a) Null model (NM) – the traditional principal components (PCs) based tree ring reconstruction model and (b) the Alternate Model (AM) – hybrid SST and tree ring reconstruction model. The NM reconstructs streamflow using a linear regression between PCs of the tree rings and the Box-Cox transformed flow (see DM) (Fig. 2 Orange). In contrast, the AM separates streamflow into components (Fig. 2 Purple) by regressing transformed streamflow against SSTs and regressing the residuals from the SST regression with the PCs of tree rings (Fig. 2 Blue). This is similar to the methodology in 36 but uses a much simpler approach in separating the moisture transport from outside the region (i.e., using SSTs) and local/regional moisture transport and basin storage (i.e., using SST and tree-ring chronologies) (See DM and SI).

Since the AM has more predictors than the NM, the skill of both models is compared using adjusted R^2 , \bar{R}^2 , (See SI-3 for details) to understand the improved performance is due to the information in the additional predictors (i.e., SSTs) rather than due to addition of parameters²⁵. Figure 3 shows AMs have significant \bar{R}^2 values for majority of HCDN basins over the CONUS, particularly for basins in the Pacific Northwest and Sunbelt regions, where the ENSO and PDO signals have strong influence^{21,23,24}. Both models showed insignificant skill in predicting observed flows (cyan colored sites in Figs 3 and 4). Compared to NM \bar{R}^2 (Fig. 4), the AM shows a significant improvement over NM for majority of basins over the CONUS. Even in basins where the NM performed better, the difference in \bar{R}^2 was lesser than 0.05 (Table SI-1). But, the AM shows a significant improvement across the CONUS with the largest improvement being in the Pacific Northwest and Sunbelt regions, where the ENSO and PDO signal is greatest. In the Sunbelt east (Hydrologic Unit Code (HUC 2) regions 03, 08 and 12), the AM explains an additional 5–10% variability (i.e., difference in \bar{R}^2 between 0.05 and 0.1) over majority of the basins. Further, in Sunbelt west (HUC2 regions 11, 13 and 15) and Pacific Northwest (HUC 2 regions 10 and 17), the AM explains an additional 5–10% variability in almost all sites and has several sites where the \bar{R}^2 improvement exceeds 0.15. The AM also has lower Normalized Root Mean Square Error (N-RMSE; see SI-3 for details) than the NM in 268 (89%) basins (Fig. SI-3). While the NM performs better in 11% of basins, the average improvement in

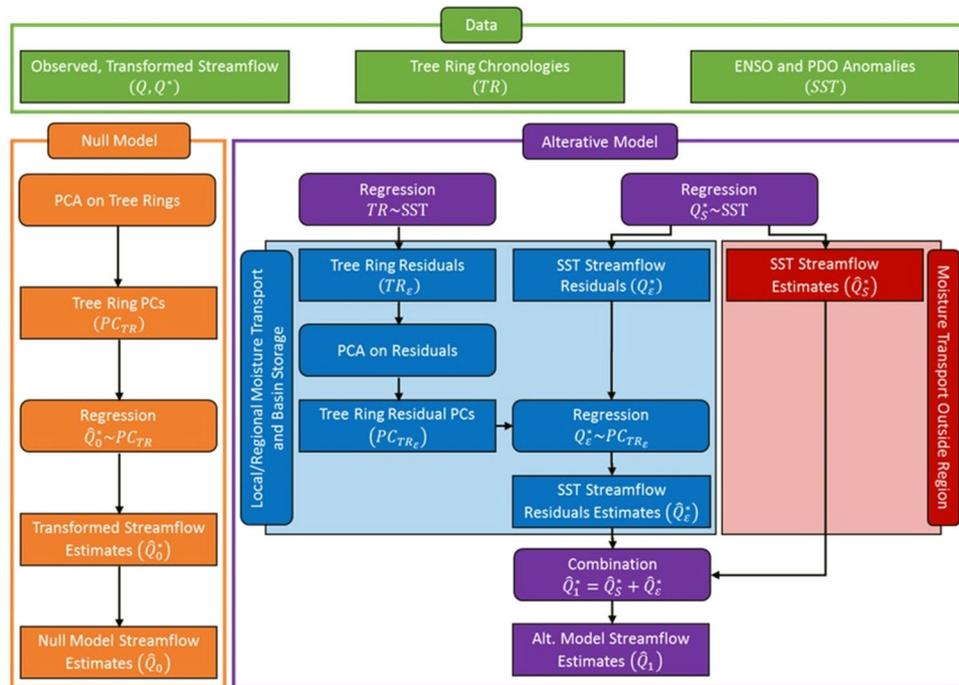


Figure 2. Data and framework for the two annual streamflow reconstruction, NM (orange) and AM (purple), considered in the study. Rectangles represent data and the operator \sim represents linear regression with the intercept model between the predictand (RHS of the operator) and the predictors (LHS).

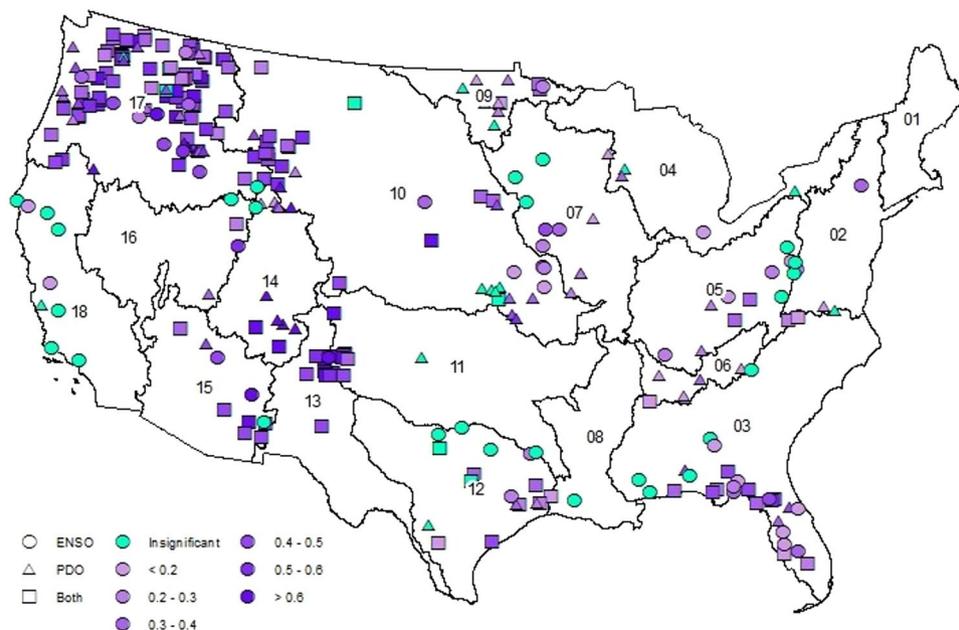


Figure 3. Adjusted R^2 values for the AM in the selected 301 basins over the CONUS. Cyan color indicates the correlation between observed and estimated streamflow was insignificant by both AM and NM. This map is created using software suite ArcGIS 10.2.2 for Desktop, version number 10.2.2.3552 (<http://www.esri.com/en/arcgis/products/arcgis-pro/overview>).

\bar{R}^2 is only 2% compared to 8% average improvement across basins where the AM performs better. In predicting above normal flows (Figs 5, SI-4), the AM performs better in 257 basins (85%) with an average improvement of 9% while the NM only performs better in 44 basins (15%) with an average improvement of 3%. In general, the improvements are better over the Sunbelt and Northwest. The reconstructed flows from the AM show good skill in predicting above-normal flows (i.e., above 67th percentile of annual flows) (Fig. 6) and also improves the overall

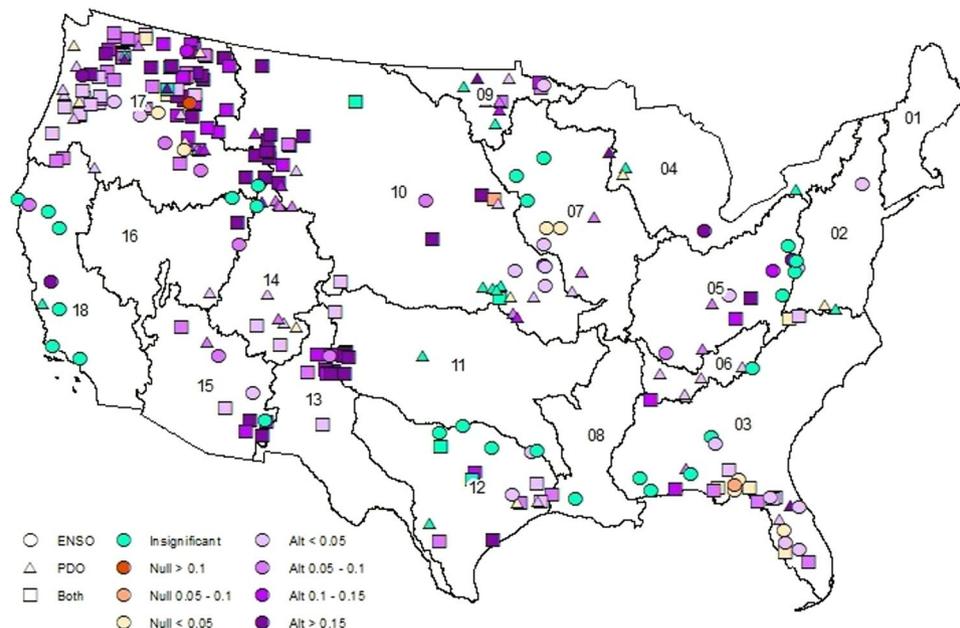


Figure 4. Comparison of adjusted R^2 from NM and AM. Sites where AM (NM) has higher adjusted R^2 indicated in darker purple (orange). Cyan indicates an insignificant correlation between estimated and observed streamflow by both models. This map is created using software suite ArcGIS 10.2.2 for Desktop, version number 10.2.2.3552 (<http://www.esri.com/en/arcgis/products/arcgis-pro/overview>).

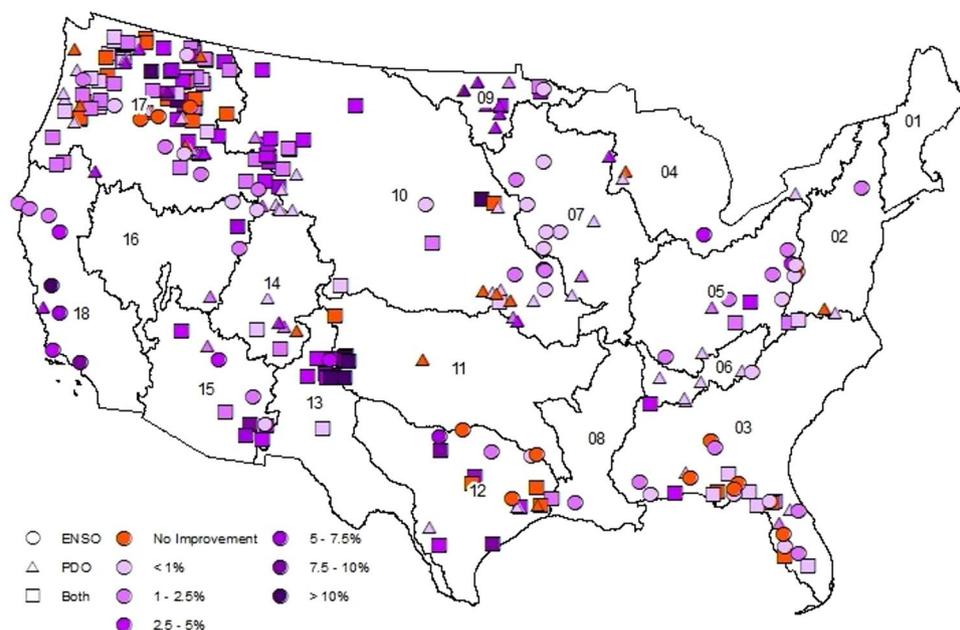


Figure 5. Comparison of above normal N-RMSE values of estimates from NM and AM. The improvement of above normal N-RMSE of AM from NM is indicated as in percentage of mean annual streamflow. Purple (orange) indicates AM (NM) had lower N-RMSE than NM (AM). This map is created using software suite ArcGIS 10.2.2 for Desktop, version number 10.2.2.3552 (<http://www.esri.com/en/arcgis/products/arcgis-pro/overview>).

skill in reconstruction. For sites with NM performing better (Fig. SI-5), the improvements in predicting above-normal flows are only marginal resulting in reduced improvements in the overall \bar{R}^2 . Further, NM mostly performs better for smaller basins with smaller mean annual discharge (Fig. SI-6). In smaller basins, it is natural to expect the role of large-scale climate teleconnection to be minimal as the basin scale hydroclimate dominates, thereby reduced role of SST in improving the reconstruction. Thus, most improvements on overall \bar{R}^2 comes from the improvements in reconstructing above-normal flows. Streamflow, being the response of the land surface, is

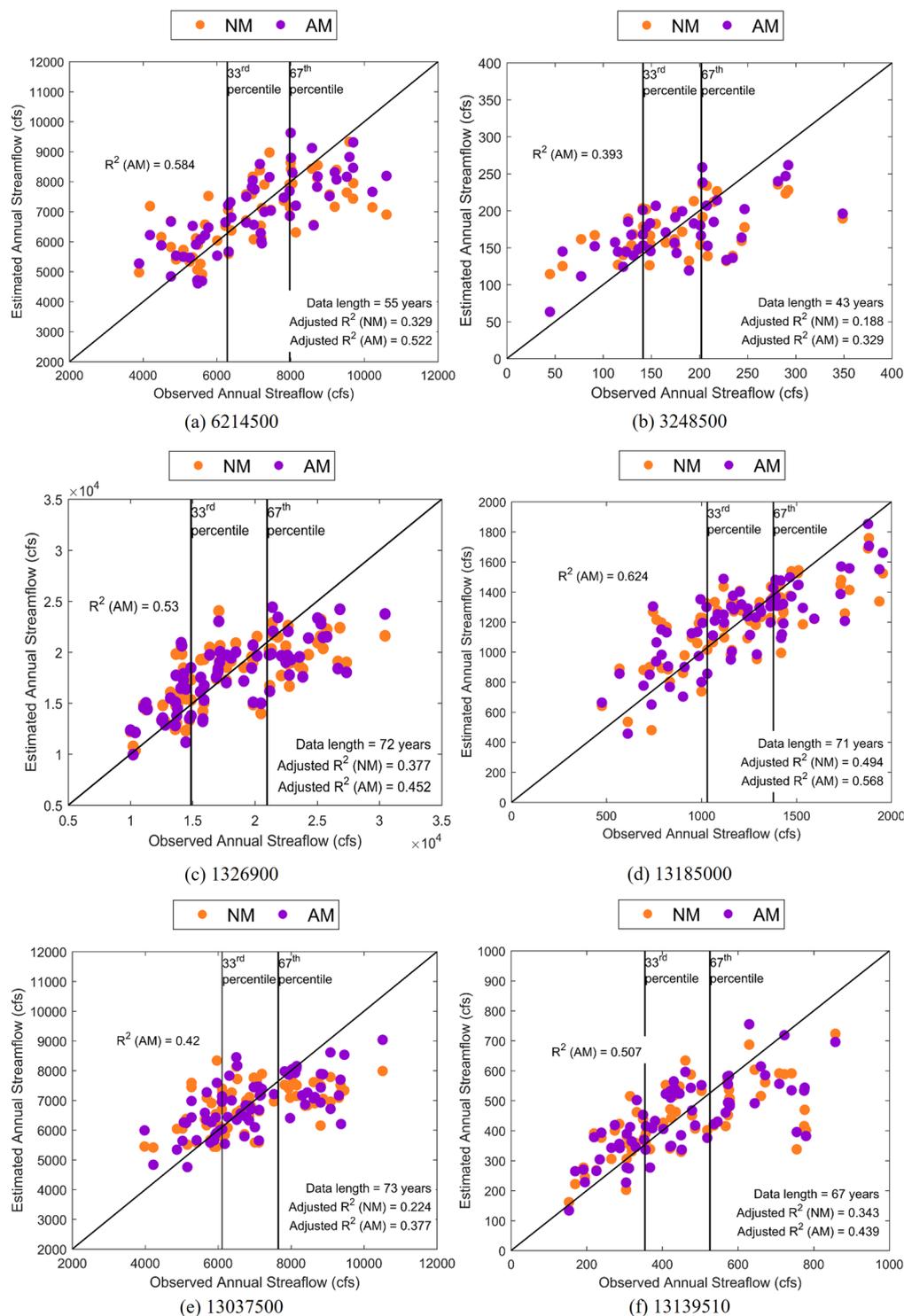


Figure 6. Comparison between observed and reconstructed annual streamflow for selected HCDN stations where AM shows good (>15%: **a,b**), moderate (5–10%: **c,d**), and marginal (<5%: **e,f**) improvements in adjusted R^2 over NM. Terciles categories (below-normal <33rd percentile; above-normal >67th percentile; Normal – 33rd to 67th percentile of annual flows) of annual flows are also marked along with data length and adjusted R^2 of the NM and AM.

controlled by other factors such as heterogeneity in forcings – precipitation and temperature – and spatial variability in soil and aquifer characteristics. Thus, the proposed hybrid approach improves streamflow reconstruction over the traditional approach, but still falls short in explaining the total variability due to spatial variability in forcings and land surface heterogeneity.

Discussion

The proposed hybrid methodology emphasizes using multi-proxies for reconstructing streamflow. Several studies have improved the reconstruction of climate fields using paleo-information from multiple proxies^{26–28}. Multi-proxies – sediment deposits and pre-observational flood information – have been used to reduce the uncertainty in flood frequency estimation^{29,30}. Findings in this study are in line with 36 in demonstrating that including SST with tree rings for streamflow reconstruction improves above normal (Fig. 5) flows and in the process improves overall reconstruction skill (Figs SI-3, SI-4). Further, the \bar{R}^2 value comparison in Fig. 4 shows that the skill improvement is due to the signal in the SSTs and not due to the addition of parameters. This is consistent with other studies that have used SSTs to improve reconstructions in snow water equivalent^{15,16} and wildfires¹⁷, in demonstrating the utility of SST in reconstruction. This hybrid approach of using SSTs could potentially improve upon other reconstruction efforts such as precipitation³¹, extreme precipitation³² and drought indices^{33–35}. In addition, the AM shows highest skill (Fig. 3) and greatest improvement (Fig. 4) over the NM in the Pacific Northwest and Sunbelt regions of the United States where low-frequency oscillations have strong influence in regional hydroclimate^{21,23,24}. Thus, regions whose hydroclimatology exhibit significant association with low frequency oscillatory signals could consider including SST anomalies as additional predictors.

Though the reconstruction was performed in virgin basins with relatively long data records from the CONUS, the hybrid reconstruction methodology is expected to provide significant benefits for basins with limited streamflow records as both SST anomalies and tree-ring chronologies are available for a longer period. Further, both these data sources extend beyond the available meteorological records (i.e., precipitation), thereby ably supporting Prediction in Ungauged Basins (PUB) efforts³⁶. In addition to ungauged basins, improved streamflow reconstructions can also benefit basins with significant anthropogenic influences (e.g. reservoir storage and groundwater pumping). For instance, existing naturalized inflow time series for basins with significant anthropogenic influences could be used with the proposed hybrid methodology for developing reconstructed flows, which could be used for capacity expansion of reservoir and for revising the current design yield considering the paleo-information. In fact, studies have shown the information from streamflow reconstruction improves reservoir sizing³⁷ and provide better insight into future streamflow conditions in comparison to near-term climate change projections³⁸. Potential to re-assess operational rule curves to incorporate environmental flows also exists using the augmented record^{39,40}.

Although the proposed hybrid methodology shows improvements in reconstruction skill, the methodology has limitations. Since the model estimates the outside-the-region moisture transport using SST anomalies, the proposed reconstruction could be applied only for basins that exhibit significant influence due to low frequency climatic signals. In addition, while the SST data is of good quality for a significant period (1940 – present), SST data prior to 1940⁴¹ is primarily estimated from ships and limited buoys. Thus, additional investigation is needed on how the uncertainty in SSTs propagates into streamflow reconstruction. One approach is to use available meteorological records from 1900s with a hydrologic model and compare with the reconstructed flows. Another approach is to replace the SST anomalies with the General Circulation Model (GCM) simulated SST time series from Coupled Model Intercomparison Project Phase 5 (CMIP5)⁴² with the hybrid methodology and compare the performance with the AM skill. However, both these additional validation efforts suffer from errors from the hydrologic model and from GCMs.

Since the hybrid methodology considers SST data for reconstruction, it also depends on the length of the SST data, which is available from 1856 onwards. Thus, it potentially limits the length of reconstructed data. Dendrochronological studies have developed reconstructed hydroclimatic data using multi-species data having different uncertainty over different time intervals^{35,43}. Similar nested reconstruction techniques could be considered to combine the reconstructed streamflows from the hybrid approach (i.e., post 1856 period) and the traditional approach (i.e., prior to 1856), thereby developing a long time series of reconstructed streamflows with different uncertainty over the entire period. Similarly, the proposed methodology also did not consider inter-site correlation as the effort here focused at the national scale. One could consider techniques⁴⁴ that explicitly consider inter-site correlation with the proposed technique to accommodate for basins exhibiting significant inter-site correlation.

The inability of tree-ring chronology to capture the moisture availability during wet years due to soil and metabolic limits results in reduced interannual variability in the reconstructed annual streamflows. Given the limitation in quantifying wetter conditions, it is expected that the uncertainty in tree-ring chronologies to be larger in comparison to instrumental records in quantifying the moisture availability over the basin. So, our hybrid approach of using multi-proxies – SST and tree-rings combined – improves the streamflow reconstruction. We also considered other teleconnection indices such as North Atlantic Oscillation and Atlantic Multidecadal Oscillation in influencing the annual streamflows, but only fewer basins showed significant skill in influencing the annual streamflows. Studies²⁶ have shown that both these indices influence the hydroclimatology of the basins predominantly at the regional level. Further, we did not consider it for reconstruction due to non-availability of these indices dating back to 1856. Extending these indices back to 1856 and incorporating them in the proposed hybrid methodology of using multi-proxies could further improve the reconstructed streamflows for basins exhibiting influence with AMO and NAO.

Data and Methods. Streamflow data used for reconstruction methodology are part of the HCDN database^{45,46} and are considered as undeveloped/virgin basins with no impacts due to upstream storage or groundwater pumping. The tree-ring chronologies used for streamflow reconstruction in this study are from the National Atmospheric and Oceanic Administration (NOAA) International Tree Ring Data Bank (ITRDB) (Fig. 1). Tree ring chronologies were selected as predictors for a streamflow site if the tree-ring chronologies and HCDN gauge: (1) belong to the same water resources (HUC 2) region (Figs 1,2) are located within 400 km from each other

and (3) have 20 or more years of common data (Fig. SI-2(b)). Only HCDN streamflow basins with at least two tree-ring chronologies identified were considered for this study.

Given the analysis focused at the national scale, both ENSO and PDO indices were considered as predictors^{21,23,24}. ENSO conditions, denoted by Nino3.4, were obtained from 1856 from International Research Institute for Climate and Society (IRI) data library using the Kaplan's SST⁴¹. Time series of PDO from 1856 was also obtained from the NCEP teleconnections database⁴⁷ using Extended reconstructed SST (ERSSTv4)⁴⁸. Both Kaplan's and ERSSTv4 SST database utilize optimal interpolation methods to develop a unified global SST database with uncertainty bounds from 1856 to till date by combining SST observations from buoys and ships. For details see^{41,48,49}. Basins were classified as ENSO and PDO affected if the Spearman's rank correlation between the observed streamflow and Nino 3.4 and PDO indices respectively was significant at a 95% confidence level, and HCDN basins affected by neither were not considered for this study. Correlations between Nino 3.4 and PDO and annual streamflow are shown (Figs SI-1 and SI-2) in SI. Out of the 301 stations, 69 (85) were affected by ENSO (PDO) and 147 were impacted by both ENSO and PDO.

Figure 2 presents the reconstruction methodology for the AM and NM. Similar to 36, the hybrid reconstruction model separates streamflow into components of moisture transport from outside the region, which will be explained by SST, and local/regional moisture transport and basin storage, which will be explained by tree ring chronologies. Both reconstruction models utilize linear regression, so the Box-Cox transformed predictands follow a normal distribution. Using Lilliefors test^{50,51}, the normality of the Box-Cox -transformed flows at each site was first tested based on a 95% significance level (Fig. SI-7). [see SI for additional information and Fig. SI-8 for lambda values] was performed and the transformed time series was used as a predictand in Fig. 2. The proposed hybrid reconstruction methodology used in this study is similar to the reconstruction methodology presented in²², which separate streamflow into moisture transport from outside the region, estimated by SST anomalies, and local moisture transport and basin storage, estimated by tree-ring chronologies after SST signal using Singular Spectrum Analysis (SSA)⁵². SSA is a dimension reduction analysis which identifies periodic components of a time series by comparing lagged versions of the time series⁵², only considered the ENSO signal on streamflow which has a periodicity of three to seven years. In contrast, PDO has a periodicity of 10–20 years, making the signal difficult to identify, especially for HCDN basins with less than 30 years of data. Given the complications of PDO in this study, we pursued the approach presented in Fig. 2 for quantifying outside-the-region moisture transport.

Null Model methodology. The NM (orange) utilizes only tree-ring chronologies identified as predictors for each site to explain all components of annual streamflow. Since tree ring chronologies across the sites exhibit high correlation, principal component analysis (PCA) was performed on the tree ring chronologies to obtain predictors for the regression model^{4,7,53}. Using tree-ring principal components (PCs) explaining 90% of the variance and the transformed streamflow (Q_0^*), linear regression⁵⁴ coefficients along with the intercept are estimated for the NM ($Q_0^* \sim PC_{TR}$ – the operator \sim represents the linear regression model with the intercept) resulting in transformed streamflow estimates (\hat{Q}_0^*). The NM reconstructed streamflow (\hat{Q}_0) are obtained by inverse of the transformation (log or Box-Cox) used for the site.

Alternative Model Methodology. AM (purple) (Fig. 2) separates the outside the region moisture transport and the local/regional moisture transport and basin storage components. Components are identified and separated by linearly regressing the transformed streamflow against the identified SST anomaly(s) as described previously ($Q_s^* \sim SST$). Using the linear regression coefficients, the SST-explained streamflow (\hat{Q}_s^*) and SST streamflow residuals (Q_ε^*) are assumed as the moisture transport from outside the region (red) and the local/regional moisture transport and basin storage (blue) components of streamflow respectively. Since tree ring chronologies also have the signal from the moisture transport from outside the region, SST influence on chronologies is removed by regressing each tree ring chronology against the SST anomaly(s) identified for the basin ($TR \sim SST$) and the resulting residuals (ε_{TR}) are retained for PCA. Similar to the NM, these residuals from different tree ring chronologies will be highly correlated, so a PCA is performed and the components explaining 90% of variance is retained (PC_{TR_ε}). The SST streamflow residuals (Q_ε^*) are then regressed with the retained PCs ($Q_\varepsilon^* \sim PC_{TR_\varepsilon}$) to obtain the within basin streamflow component (\hat{Q}_ε^*). The AM transformed streamflow estimates (\hat{Q}_1^*) are calculated by adding the SST streamflow estimates (\hat{Q}_s^*) and residual estimates (\hat{Q}_ε^*). Finally, the AM streamflow estimates (\hat{Q}_1) are also obtained by transforming the transformed streamflow estimates (\hat{Q}_1^*) back to the original space using the inverse of the log or Box-Cox transformation.

Data availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Cook, E. R. & Jacoby, G. C. Potomac river streamflow since 1730 as reconstructed by tree rings. *J. Climate Appl. Meteorol.* **22**, 1659–1672 (1983).
2. Devineni, N., Lall, U., Pederson, N. & Cook, E. A tree-ring-based reconstruction of delaware river basin streamflow using hierarchical bayesian regression. *J. Climate* **26**, 4357–4374 (2013).
3. Meko, D., Stockton, C. W. & Boggess, W. R. The tree-ring record of severe sustained drought. *J. Am. Water Resour. Assoc.* **31**, 789–801, <https://doi.org/10.1111/j.1752-1688.1995.tb03401> (1995).
4. Woodhouse, C. A. A tree-ring reconstruction of streamflow for the Colorado front range. *J. Am. Water Resour. Assoc.* **37**, 561–569, <https://doi.org/10.1111/j.1752-1688.2001.tb05493.x> (2001).
5. Fritts, H. C. *Tree Rings and Climate* (Academic Press, London, 1976).
6. Fritts, H. C. *Reconstructing Large-scale Climatic Patterns from Tree-Ring Data: A Diagnostic Analysis.* (University of Arizona Press, Tucson, 1991).

7. Hidalgo, H. G., Piechota, T. C. & Dracup, J. A. Alternative principal components regression procedures for dendrohydrologic reconstructions. *Water Resour. Res.* **36**(11), 3241–3249, <https://doi.org/10.1029/2000WR900097> (2000).
8. Stockton, C. & Jacoby, G. Long-term surface-water supply and streamflow levels in the Upper Colorado River basin. Lake Powell Res. Proj. Bulletin, Number 18, Inst. of Geophys. and Planet. Phys., Univ. of Calif., Los Angeles, Calif., 70 pp. [Available from University of California, Los Angeles, 3845 Slichter Hall, Mail Code 156704, Los Angeles, CA 90095-1567.] (1976).
9. Cleaveland, M. K. A 963-year reconstruction of summer (JJA) stream flow in the White River, Arkansas, USA from tree-rings. *The Holocene* **10**(1), 33–41 (2000).
10. Stahle, D. W. *et al.* Tree-ring data document 16th century megadrought over North America. *Eos Trans. Am. Geophys. Union* **81**(12), 121–125 (2000).
11. Stahle, D. W., Fye, F. K., Cook, E. R. & Griffin, R. D. Tree-ring reconstructed megadroughts over North America since A.D. 1300. *Climatic Change* **83**, 133–149, <https://doi.org/10.1007/s10584-006-9171> (2007).
12. Fritts, H. C. *Tree Rings and Climate*. (Blackburn Press, Caldwell, NJ, 2001).
13. Gangopadhyay, S., Harding, B. L., Rajagopalan, B., Lukas, J. J. & Fulp, T. J. A nonparametric approach for paleohydrologic reconstruction of annual streamflow ensembles. *Water Resour. Res.* **45**, W06417, <https://doi.org/10.1029/2008WR007201> (2009).
14. Margolis, E. Q., Meko, D. M. & Touchan, R. A tree-ring reconstruction of streamflow in the Santa Fe River, New Mexico. *J. Hydrol.* **397**, 118–127 (2011).
15. Anderson, S. R., Aziz, O., Tootle, G. A., Grissino-Mayer, H. & Barnett, A. Using Pacific Ocean climatic variability to improve hydrologic reconstructions. *J. Hydrol.* **434**, 69–77, <https://doi.org/10.1016/j.jhydrol.2012.02.037>. (2012b).
16. Anderson, S. R. *et al.* Snowpack reconstructions incorporating climate in the Upper Green River Basin (Wyoming). *Tree Ring Res.* **68**(2), 105–144 (2012a).
17. Kitzberger, T., Brown, P. M., Heyerdahl, E. K., Swetnam, T. W. & Veblen, T. T. Contingent Pacific–Atlantic Ocean influence on multicentury wildfire synchrony over western North America. *Proc. Natl. Acad. Sci. USA* **104**, 543–548 (2007).
18. Trenberth, K. E. & Guillemot, C. J. Physical Processes involved in the 1988 Drought and 1993 Floods in North America. *Journal of Climate* **9**, 1288–1298 (1996).
19. Cayan, D. R., Redmond, K. T. & Riddle, L. G. ENSO and Hydrologic Extremes in the Western United States. *J. Climate* **12**, 2881–2893 (1999).
20. Ropelewski, C. F. & Halpert, M. S. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.* **115**, 1606–1626 (1987).
21. Halpert, M. S. & Ropelewski, C. F. Surface-temperature patterns associated with the Southern Oscillation. *Journal of Climate* **5**(6), 577–593 (1992).
22. Patskoski, J., Sankarasubramanian, A. & Wang H. Reconstructed Streamflow using SST and Tree-ring Chronologies over the Southeastern United States. *Journal of Hydrology* **527**, 761–775 (2015).
23. McCabe, G. J. & Dettinger, M. D. Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *Int. J. Climatol.* **19**(13), 1399–1410 (1999).
24. Tootle, G. A. & Piechota, T. C. Relationships between Pacific and Atlantic Ocean Sea surface temperatures and U.S. streamflow variability. *Water Resour. Res.* **42**, W07411 (2006).
25. Ramanathan, R. *Introductory Econometrics: With Applications*. (Harcourt, Brace & Company, Orlando, FL, 1998).
26. Li, B., Nychka, D. W. & Ammann, C. M. The value of multi-proxy reconstruction of past climate. *J. Am. Stat. Assoc.* **105**, 883–911 (2010).
27. Tingley, M. P. & Huybers, P. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: Development and applications to paleoclimate reconstruction problems. *J. Climate* **23**, 2759–2781 (2010a).
28. Tingley, M. P. & Huybers, P. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: Comparison with the regularized expectation–maximization algorithm. *J. Climate* **23**, 2782–2800 (2010b).
29. Costa, J. E. Paleohydrologic reconstruction of flash-flood peaks from boulder deposits in the Colorado Front Range. *GSA Bulletin* **94**(8), 986–1004, <https://doi.org/10.1130/0016-7606> (1983).
30. Lam, D., Thompson, C., Croke, J., Sharma, A. & Macklin, M. Reducing uncertainty with flood frequency analysis: The contribution of paleoflood and historical flood information. *Water Resour. Res.* **53**, 2312–2327, <https://doi.org/10.1002/2016WR019959> (2017).
31. Le Quesne, C. *et al.* Ancient *Austrocedrus* tree-ring chronologies used to reconstruct Central Chile precipitation variability from A.D. 1200 to 2000. *J. Climate* **19**, 5731–5744 (2006).
32. Steinschneider, S., Ho, M., Cook, E. R. & Lall, U. Can PDSI inform extreme precipitation?: An exploration with a 500 year long paleoclimate reconstruction over the U.S. *Water Resour. Res.* **52**, 3866–3880, <https://doi.org/10.1002/2016WR018712> (2016).
33. Ho, M., Lall, U. & Cook, E. R. Can a paleodrought record be used to reconstruct streamflow?: A case study for the Missouri River Basin. *Water Resour. Res.* **52**, 5195–5212, <https://doi.org/10.1002/2015WR018444> (2016).
34. Kauffman, G. J. & Vonck, K. J. Frequency and intensity of extreme drought in the Delaware basin, 1600–2002. *Water Resour. Res.* **47**, W05521, <https://doi.org/10.1029/2009WR008821> (2011).
35. Maxwell, R. S., Hessel, A. E., Cook, E. R. & Pederson, N. A multispecies tree ring reconstruction of Potomac River streamflow (950–2001). *Water Resour. Res.* **47**, W05512, <https://doi.org/10.1029/2010WR010019> (2011).
36. Blöschl, G., Sivapalan, M. & Wagener, T. Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales. p. 484. Cambridge University Press, (2013).
37. Patskoski, J. & Sankarasubramanian, A. Improved Reservoir Sizing Utilizing Observed and Reconstructed Streamflows within a Bayesian Combination Framework. *Water Resour. Res.*, <https://doi.org/10.1002/2014WR016189> (2015).
38. Patskoski, J. & Sankarasubramanian, A. Reducing Uncertainty in Stochastic Streamflow Generation and Reservoir Sizing by 1 Combining Observed, Reconstructed and Projected Streamflow (under revision SERRA).
39. Poff, N. L. R. *et al.* Sustainable water management under future uncertainty with eco-engineering decision scaling. *Nature Climate Change* **6**(1), 25–34, <https://doi.org/10.1038/nclimate2765> (2016).
40. Wang, H., Brill, E. D., Ranjithan, R. S. & Sankarasubramanian, A. A Framework for Incorporating Ecological Releases in Single Reservoir Operation. *Adv. Water Res.* **78**, 9–21 (2015).
41. Kaplan, A. *et al.* Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.* **103**, 18567–18589 (1998).
42. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.* **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1> (2012).
43. Pederson, N. *et al.* Is an epic pluvial masking the water insecurity of the Greater New York City region? *J. Climate* **26**, 1339–1354 (2013).
44. Bracken, C., Rajagopalan, B. & Woodhouse, C. A Bayesian hierarchical nonhomogeneous hidden Markov model for multisite streamflow reconstructions. *Water Resour. Res.* **52**, 7837–7850, <https://doi.org/10.1002/2016WR018887> (2016).
45. Slack, J. R., Lumb, A. & Landwehr, J. M. Hydro-Climatic Data Network (HCDN) Streamflow Data Set, 1874–1988, *U.S. Geological Survey Report* (1993).
46. Vogel, R. M. & Sankarasubramanian, A. The Validation of a Watershed Model without Calibration. *Water Resour. Res.* **39**(10), art. no.1292 (2003).
47. Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M. & Francis, R. C. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.* **78**, 1069–1079 (1997).

48. Huang, B. *et al.* Extended Reconstructed Sea Surface Temperature Version 4 (ERSST.v4). Part I: Upgrades and Intercomparisons. *J. Climate* **28**, 911–930, <https://doi.org/10.1175/JCLI-D-14-00006.1> (2015).
49. Hamlet, A. F. & Lettenmaier, D. P. Columbia River streamflow forecasting based on ENSO and PDO climate signals. *J. Water Resour. Plann. Manage.* **125**, 333–341 (1999).
50. Lilliefors, H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Am. Stat. Assoc.* **62**, 399–402 (1967).
51. Lilliefors, H. W. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *J. Am. Stat. Assoc.* **64**, 387–389 (1969).
52. Vautard, R. & Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenomena* **35**(3), 395–424 (1989).
53. Almanaseer, N., Sankarasubramanian, A. & Bales, J. Improving groundwater predictions utilizing seasonal precipitation forecasts from GCMs forced with SST forecasts. *J. Hydrol. Eng.* [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000776](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000776) (2012).
54. Weisberg, S. *Applied Linear Regression*. (John Wiley, New York, 1985).

Acknowledgements

This research is supported in part by the National Science Foundation under two grants: CBET- 0954405, CBET-1442909. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author Contributions

S.M, J.M.P. and A.S. developed the hybrid streamflow reconstruction model. S.M handled the computational work, wrote the code, analyzed the output, developed the figures. J.M.P. assembled the input data and wrote the manuscript. A.S. designed and administered the work. All authors contributed to the supplementary information.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-23294-6>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018