California Fish and Wildlife 106(4):221-259; 2020

FULL RESEARCH ARTICLE

Time series modeling and forecasting of a highly regulated riverine system: implications for fisheries management

ROBERT M. SULLIVAN[1]* AND JOHN P. HILEMAN[2]

[1] *California Department of Fish and Wildlife, Region 1, Wildlife/Lands Program, P.O. Box 1185 Weaverville, California 96093*

[2] *California Department of Fish and Wildlife, Region 1, Fisheries Program, Trinity River Project, P.O. Box 1185, Weaverville, California 96093*

**Corresponding Author: robert.sullivan@wildlife.ca.gov*

Here we apply seasonal time series modeling to flow and fisheries management in a highly regulated river system. Time series modeling is commonly employed to forecast future values of streamflow and extrinsic climate-related seasonal data based on historical information. This method has not been employed in evaluating fish-flow management in highly regulated rivers that experience regular and long-term hydrological fluctuations. Forecasting annual outflow volume and predicting and evaluating its effect on variability in the thermal regime of major river systems, is vital for addressing potential impacts to anadromous and non-anadromous fisheries. Autoregressive Integrated Moving Average (SARIMA) time series analysis was used to predict and describe seasonal variation in flow volume and water temperature for the upper Trinity River in northern California. The objectives were: 1) use a multivariate approach to SARIMA modeling to describe and evaluate seasonal patterns in environmental variables associated with the historical annual time series record in response to implementation of anthropogenic flow-type hydrographs; and 2) relate results to fisheries resources in the upper Trinity River through management recommendations. Raw data representing the historical time series for volume of flow and water temperature variables were partitioned into three time series subgroups. Each subgroup represented a specific flow-type following previous research into the effects of highly managed hydrographs in the upper Trinity River. Subgroups were evaluated based on their efficiency to model and provide accurate operational forecasts of monthly environmental data. Results showed that subcomponents of the historical post-dam time series specific to managed flow-type hydrographs keyed to geomorphologic restoration actions significantly misrepresented the: 1) time series characteristics, 2) seasonal-trend decomposition patterns, and 3) forecast accuracy compared to the baseline pre-restoration time series model. These results effectively nullify use of managed hydrographs in predicting future forecasts for flow-fish planning and management purposes. Use of time series modeling without reference to continuous intrinsic periods within the historical time series data where flow is anthropogenically manipulated will be misleading when attempting to evaluate the overall characteristics and subsequent future forecasting that derive from such models. By placing environmentally-driven historical time series models into perspective relative to anthropogenically manipulated flow management the needs of both current and future water and fisheries resources will be better optimized notwithstanding the inevitable long-term effects of climate change.

Key words: fisheries resources, flow-mediated water temperature, regulated river, river restoration, SARIMA, time series modeling

---

Historically, time series analysis has been employed in building models to detect seasonal trends and forecast future values of water temperature, precipitation, air temperature, evapotranspiration, streamflow, and other climate-based data using regionally-specific historical information (Salas 1993; Adeloye and Montaseri 2002; Papalaskaris and Kampas 2017). Application of time series modeling is expanding with growing concerns about climate change and global warming. This approach requires accurate explanation of the underlying dynamics of river flow. Yet obtaining this kind of information may not always be possible by methods of statistical forecasting using conventional linear regression focused primarily on average values of data (first-order moment) but not variance in data (second-order moment; Casella and Berger 2002; Martínez-Acosta 2020; Attar et al. 2020). Use of time series modeling has not been employed in evaluating issues related to fish-flow management in riverine systems. This observation is particularly true for rivers experiencing regular and long-term hydrological fluctuations keyed to dam operational releases for in-river geomorphological restoration actions, which is the primary focus of this paper.

Variability, uncertainty, and unpredictability are hallmarks of river systems and, by extension, of any in-river restoration initiative (Wissmar and Bisson 2003). Restoration projects on large river systems in the Pacific Northwest carry substantial ecological and economic risk, which highlights the need for improved restoration science (Woelfle-Erskin et al. 2012). Adding to the uncertainty and unpredictability of riverine restoration efforts is the tendency to implement prescriptions at a local-level without consideration of the linkages at nested temporal and spatial scales (Holčík 1996). This practice minimizes insight into modeling longer-term effects that contribute to variability and unpredictability in flow-related restoration efforts and their impacts to fisheries resources (Woelfle-Erskin et al. 2012). For example, restoration flows have traditionally focused on discharge impacts over short time intervals (< 1 year). This strategy is driven primarily by annual needs for in-river geomorphic work without reference to "baseline" or "unimpaired" flow models (Shibatani 2020) synchronized with historical and current tributary flow events that automatically factor in annual variability in climate change. Shortfalls in the outcomes of river restoration have prompted calls to identify and restore processes that support and sustain biological communities, rather than focusing only on river geomorphology (Palmer and Ruhi (2019).

Stochastic Moving Average (ARMA) time series models are based on probability theory that represent the temporal uncertainty of data without a seasonal element (Martínez-Acosta et al. 2020). Modification of ARMA analyses for evaluating stochastic seasonal time series phenomenon involves Seasonal Autoregressive Integrated Moving Average (SARIMA) models. This technique in combination with the Box-Jenkins (Box and Jenkins 1976) approach, which evaluates autocorrelations among variables as well as lag-lead relationships between variables, has direct application to modeling seasonality and correlation structure in hydrological data associated with forecasting reservoir inflow, outflow discharge, and river water characteristics affecting flow-fish thermal regimes (Stergiou 1991; Pajuelo and Lorenzo 1995; Bari et al. 2015; Papalaskaris et al. 2016). Anthropogenically-induced variable hydrology and its impact on thermal regimes of major riverine systems is a critical factor affecting the physiology, ecology, and life history strategies of resident salmonids and other aquatic organisms (Olden and Naiman 2010; Hallock et al. 1970; McCullough 1999; Carter et al. 2006). Long-term patterns of flow variability have historically selected for organismal life histories related to growth, reproduction, dispersal, and the ability to persist under physical and chemical stress (Palmer and Ruhi 2019). Use of time series modeling is particularly relevant to management of fisheries resources given that most fish are ectothermic, and their life history strategies are directly and functionally affected by variable thermal regimes within their regional and local migratory landscapes (Hildebrand and Goslow 2001). When managing fisheries resources, forecasting outflow volume, and predicting its effect on variability of water temperature is particularly relevant in planning flow releases and understanding the hydrographic characteristics underlying potential impacts to anadromous fisheries resources.

Flow dynamism also is central to a functioning river system and its ability to provide ecosystem services, yet flow alteration is rarely inconsequential (Palmer and Ruhi 2019). Variability in riverine thermal regimes as a function of anthropogenic flow scheduling has the potential to adversely affect run-timing, local migratory behavior, spawning, and early development of juvenile anadromous species, including Coho Salmon (*Oncorhynchus kisutch*), spring- and fall-run Chinook Salmon, and steelhead (*Oncorhynchus mykiss*), together with resident non-anadromous Brown Trout (*Salmo trutta*) and Klamath Smallscale Suckers (*Catostomus rimiculus*; Sullivan and Hileman 2018, 2019, 2020). In northern California, populations of spring-run and fall-run Chinook Salmon (*Oncorhynchus tshawytscha*) in the Klamath Basin have declined significantly in the last 100 years as a function of several linked factors, including a series of dams on the mainstem Klamath and Trinity rivers (Romberger and Gwozdz 2018). All stocks of salmonids in the upper Trinity River are at their lowest levels since 2007. The upper Trinity River has experienced repeated fluctuations in the seasonal volume of flow and thermal regime as a function of water management policy, in-river anthropogenic actions centered around geomorphological restoration activity, and attempts to control disease using artificially augmented pulsed flows since 2003 (Sullivan and Hileman 2018). These activities directly affect migration and run-timing in both non-anadromous and anadromous salmonids linked to the expression of regionally adapted life history strategies (Sullivan and Hileman 2020).

The objectives of this research were threefold. First, a multivariate approach to SARIMA time series modeling was used to describe seasonal patterns in environmental variables potentially influenced by implementation of specific anthropogenic flow-type hydrographs. Second, we tested the hypothesis that subcomponents of the historical time series data representing specific in-river restoration actions significantly misrepresent the time series characteristics and forecast accuracy of seasonal flow volume and water temperature compared to a "baseline" pre-restoration time series model. Third, we relate results of our research to management of fisheries resources in the upper Trinity River through recommendations focused exclusively on time series modeling and analysis. Forecasting plays a critical role in fisheries management because it precedes planning which, in turn, precedes decision making (Makridakis et al. 1983; Stergiou et al. 1997). By placing environmentally-driven historical time series models into perspective relative to anthropogenically manipulated flow management, the needs of both current and future water and fisheries resources will be better optimized and managed notwithstanding the inevitable long-term effects of climate change.

METHODS

Study area

The Trinity River is in northwestern California and is the largest tributary of the Klamath River system (Figure 1A). Construction of Trinity and Lewiston dams occurred in the early 1960s.Trinity Dam creates Trinity Reservoir, storing up to 3,022 m^3 of water. Lewiston Reservoir, formed by Lewiston Dam, is located 11.8 km downstream of Trinity Dam, which serves as a re-regulating reservoir for flow to the Trinity River and diversion to the Sacramento River Basin, comprising the Trinity River Division of the Central Valley Project (Sullivan and Hileman 2019). The Trinity River system is not connected geologically to the Sacramento River system of the Central Valley. From Lewiston Dam, the Trinity River flows ~ 180 kilometers before joining the Klamath River at the township of Weitchpec, California. The Klamath River flows for an additional 70 km before entering the Pacific Ocean. Trinity River Hatchery is located immediately below Lewiston Dam. From a management perspective, the upper 63.1 rkm of the Trinity River or “mainstem,” ends at the confluence of the North Fork Trinity River and the Trinity River proper. This section of the river is the primary focus of in-river restoration activities by the Trinity River Restoration Program (TRRP 2020). The focus of our study was the upper-most section of the mainstem just below Lewiston Dam and the Trinity River Hatchery, which is the upper limit to anadromy on mainstem.

Managed flows and release schedule

The TRRP created by the Record of Decision, henceforth called “ROD” (USBR 2000), mandated a plan for restoration of 63.1 km of the upper Trinity River and its fish and wildlife populations. The Trinity River Mainstem Fishery Restoration Environmental Impact Statement was the basis for the ROD (TRFES 1999). This restoration strategy included: 1) flow management through manipulation of the annual hydrograph, 2) mechanical channel rehabilitation, 3) sediment management, 4) watershed restoration, 5) infrastructure improvements, 6) adaptive environmental assessment and monitoring, and 7) environmental compliance and mitigation. Schedules for annual flow releases for the Trinity River are established based on water year type[1] and restoration needs (Appendix I; TRRP 2020). As of water year 2020 the proportion of water scheduled to be removed from the Lewiston Reservoir and diverted to the Central Valley is 61% of the allotment. The remainder (39%)

1 The term “water year” defined by the United States Geological Survey is “the 12-month period 1 October (of any given year) through 30 September of the following year.” The water year is designated by the calendar year in which it ends, and which includes 9 of the 12 months. Thus, the year ending September 30, 1999 is called the “1999” water year. https://water.usgs.gov/nwc/explain_data.html

is scheduled to be released into the into the Trinity River. Detailed descriptions and summaries of specific flow schedules and restoration actions are provided elsewhere (TRRP 2020; Sullivan and Hileman 2019).

Data collection and sampling

To test our hypothesis, raw data representing the historical time series data for volume of flow and water temperature were partitioned (fixed portioning; Prasad 2019) into three time series subgroups. Each subgroup corresponded to specific hydrological flow-types following previous research into the effects of highly managed hydrographs on the upper Trinity River and its fisheries resources (Figure 1B[2]; Sullivan and Hileman 2020). These subgroup flow-types included: 1) "baseline" PreROD flows (1995 – 2002), 2) ROD flows (2005 – 2011, 2017), and 3) Pulse flows (2012 – 2016, 2018). ROD and Pulse flow-types represented only those continuously unbroken sequences of years in which specific managed

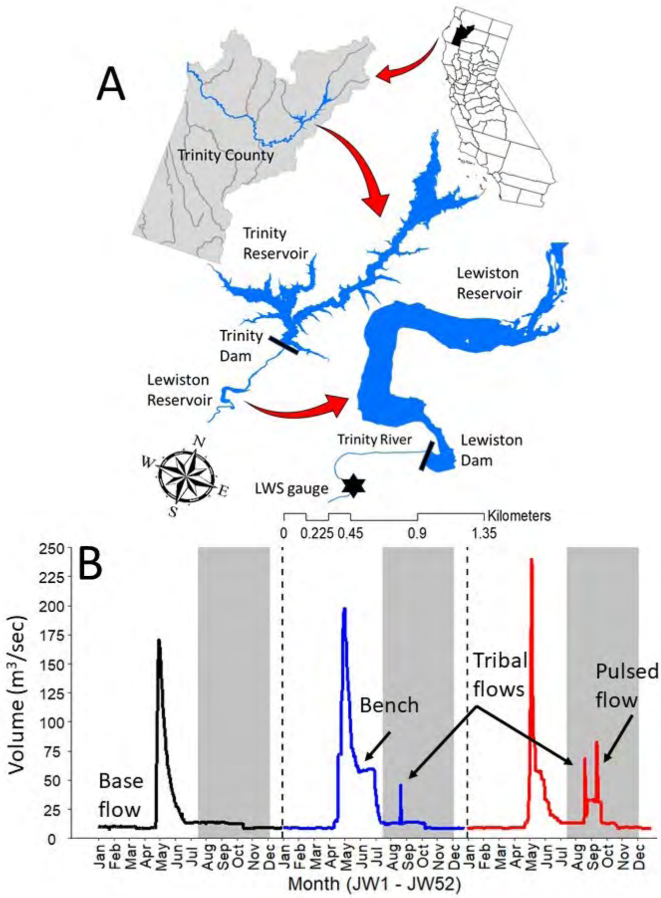


Figure 1. A) Map of Trinity County and location of Trinity Dam, Lewiston, Dam, Trinity River Hatchery, and Lewiston Water Quality Gauge (LWS) within the upper reach of traditional spawning grounds 1.7 rkm downstream of hatchery. B) Examples of hydrographs representative of different flow-types used in time series analyses. Managed hydrographs are associated with ROD and Pulse flow-types typically implemented between Julian week 13 and 40. Gray shaded areas correspond to run-time presence of various species of salmonids in the upper river.

2 Hoopa Valley Tribal Boat Dance flows are scheduled in odd years for ceremonial purposes. During this time, the United States Bureau of Reclamation increases the volume of flow from Lewiston Dam into the Trinity River in support of this ceremony in the town of Hoopa, California.

hydrographs were implemented. Daily recordings of telemetered digital data were used to assess the seasonal extent of variability in the annual thermal regime that characterizes the upper Trinity River. These data were obtained from the United States Bureau of Reclamation, Lewiston Water Quality Gauge (LWS; DWR 2020) located 1.7 rkm downriver from Lewiston Dam (river-km 178.2; UTM 516,634 m E and 4,507,678 m N; elevation 558 m). Digital data obtained included averaged daily: 1) flow volume (ADFV m^3/s), 2) minimum air temperature (MIAIR degrees Celsius [°C]), 3) maximum air temperature (MXAIR °C), 4) minimum water temperature (MIWAT °C), maximum water temperature (MXWAT °C); and 5) a measure of extreme variability in water temperature (ADWTV [average – minimum] + [maximum – average] °C). The LWS gauge was chosen because there are no other sources of inflow from the watershed or major tributaries between the LWS gauge and Lewiston Dam. As a result, this section of the river is not inadvertently influenced by other local watershed conditions. Further, this location was the "standard" used in all National Environmental Protection Act (NEPA) assessments and flow augmentation analyses of fluctuations in river water temperature, specific to the upper Trinity River out of Lewiston Dam since 1997 (Magneson and Chamberlain 2014). Importantly, this gauge provides the best location for measuring water temperature conditions nearest to the hatchery as hatchery-origin Chinook Salmon as well as other natural-origin salmonids traditionally spawn in this segment of the upper-most reach.

Statistical analyses

Standard statistics.—All statistical tests performed used the R-suite of statistical programs (R Core Team 2020) and statistical significance for all analyses was set at $p < 0.05$. Prior to implementing statistical analyses, visual assessment of each environmental variable was conducted using two goodness-of-fit plots: 1) theoretical density plots of histograms against fitted density functions; and 2) Q-Q plots of the theoretical quantiles against empirical ones with emphasis on lack-of-fit at the tails of distributions, which were evaluated by use of the Akaike information criterion (AIC; Package "MuMIn;" Appendix II). A follow-on statistical evaluation using the Anderson-Darling (A) test statistic (Stephens 1986) similarly showed that all environmental metrics were not normally distributed (minimum air temperature: $A = 14.7$, $p < 0.001$, $n = 5{,}752$]; maximum air temperature: $A = 80.7$, $p < 0.001$, $n = 5{,}752$]; minimum water temperature: $A = 10.9$, $p < 0.001$, $n = 8{,}055$]; maximum water temperature: $A = 14.5$, $p < p < 0.001$, $n = 8{,}055$]; extreme variability in water temperature: $A = 276.1$, $p < 0.001$, $n = 8{,}042$]; flow volume: $A = 148.5$, $p < 0.001$, $n = 8{,}401$]). Thus, all subsequent statistical analyses used non-parametric or semi-parametric[3] statistical methods (McDonald 2014; Tsiatis et al. 2006). Spearman's rank correlation rho (r_s) 2-tailed test was used to calculate the strength and direction of the relationship between two variables, expressed as a monotonic relationship, whether linear or not (Corder and Foreman 2014). Kruskal-Wallis Chi-square (χ^2) rank sum tests evaluated each designated variable accompanied by follow-on planned pairwise comparisons between each designated group using the Dunn test statistic (Z). All p-values were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg 2000; Machiwal and Madan 2006).

Principal components analysis (PCA; Program "FactoMineR") was used to describe variation, identify variable selection, discard redundant variables, and assess seasonal varia-

3 In statistics, a semiparametric model is a statistical model that has parametric and nonparametric components. https://en.wikipedia.org/wiki/Semiparametric_model

tion in each environmental variable. Component axes that accounted for > 1% of the total variation in "attribute space" were retained for further analysis among flow-types for each time series. This method is generally preferred for numerical accuracy as resulting principal components are orthogonal, thus minimizing multicollinearity between model predictors, with the goal of identifying a smaller subset of variable components that capture the majority of variance in predictors (Everitt and Hothorn 2011).

Generalized additive models.—Semi-parametric generalized additive modeling (GAM; Package "mgcv;" Wood 2017) was used in regression of each environmental variable (Hastie and Tibshirani 1990; Madsen and Thyregod 2011; Wood 2017). Response curves generated from each GAM showed the relationship between the fitted function and the response variable. Smooths were "centered" to ensure model identity and summed to zero over covariate values. Statistics reported by each GAM included: 1) F-statistic (approximate significance of smooth terms), 2) p-values and 95% confidence bands for spline lines (Nychka 1988), 2) adjusted regression coefficients for each model (R^2.Adj.), 3) estimated residual degrees of freedom (Ref.df.), and 4) proportion of null deviance explained (Dev.Exp.). The Spearman's rank correlation coefficient was used as a follow-on procedure to assess strength and significance of trends in each variable delineated by smooth terms. Ranked correlation was used because GAMs lack a statistical inference procedure and formal parameter of goodness of the fit, which makes interpretation of output potentially complicated (Package "fitdistrplus;" Diankha and Thiaw 2016). The gamma error-structure (family = "Gamma" [link = "log"]) was used to assess the error distributions in fluctuations of each environmental variable; and the AIC information criterion was used to select the most parsimonious error distribution for each environmental attribute (Package "MuMIn;" Akaike 1973; Burnham and Anderson 1998).

Seasonal-trend decomposition and adjustment.—Historical annual data were aggregated by month. Months were used instead of Julian weeks (52 weeks/year) because the historical time series data were not detailed enough for each environmental variable to justify using weekly subdivisions. Use of Julian weeks would have been preferable because they would have imparted greater detail to the analyses. Seasonal-Trend-Loess (STL) decomposition component analysis was used to remove the seasonal effect from each time series variable to facilitate understanding of trends in the dataset. The STL method using local polynomial regression was fitted by a least squares algorithm to partition the time series of each dataset into three components: 1) trend (Tt), 2) seasonality-cycle (St), and 3) remainder (Rt), written as: yt = St + Tt + Rt, for t =1 to N measured data points (Hydman and Athanasopoulos 2014). Once each component was fit by the STL model, they were subtracted from the raw time series data for each environmental variable to give the remainder component of each model. Thus, the remainder component equates to the residuals derived from the seasonal plus trend fit, or "random" time series (Cleveland et al. 1990; Cleveland et al. 1992). Ability to determine if a single measurement was unusually low or high by looking at the remainder terms is a typical use of seasonal decomposition. Locally weighted regression and scatterplot smoothing (Loess) was used to estimate nonlinear relationships, in which the entire procedure is iterated using a back-fitting algorithm. Inter-quartile range (IQR) calculations for x-values were generated by STL modeling of the trend, seasonality, and remainder components used to gain a relative measure of how spread-out points were in the original time series dataset (Zar 2010). For a normal distribution with standard deviation σ, IQR = 1.35σ.

Time series modeling and seasonality.—Goals of the time series analyses were to: 1) describe the pattern indicated in the time series data, 2) identify the nature of the phenomenon represented by the sequence of observations and integrate with seasonally variable environmental effects, and 3) forecast future values in each time series model for each environmental attribute. To accomplish these goals, time series analysis assumes that successive values in the data file represent consecutive measurements taken at equally spaced time intervals (Hill and Lewicki 2007). Thus, years 2003, 2004, and 2018 were not included in the Pulse flow time series model, nor was 2017 included in the ROD flow time series model as illustrated in Figure 1A. Importantly, the years removed were consistent with the consecutive annual monthly patterns exhibited by both managed flow-types. Retention of these years and months would have provided additional detail by contributing to the robustness of statistical tests as a function increasing sample size for each anthropogenic subcomponent of each restoration time series model.

Seasonality is a significant concern when modelling time series weather data unique to a particular region. Because all environmental data evaluated herein exhibited seasonality, SARIMA modeling (Program "astsa;" Shumway and Stoffer 2017) was used to evaluate each univariate time series and to inspect model fit diagnostics (Nau 2017; Hyndman and Athanasopoulos 2018). A SARIMA model was fitted to each environmental variable for each flow-type subcomponent of the historical time series, with the intent to discover the most appropriate match of each time series data point to previous values of the same time series, and to perform future forecasts. Seasonal components of each model were written using uppercase letters, whereas non-seasonal components were referenced in lowercase letters and written in the form: SARIMA (p, d, q) (P, D, Q)m, where p = non-seasonal order (autoregressive model AR[p]), d = non-seasonal differencing, q = non-seasonal order moving average (MA), P = seasonal (AR) order, D = seasonal differencing, Q = seasonal (MA) order, and m = number of periods per season indicating the time interval of repeating seasonal sequences. The seasonal portion of each model was comprised of components resembling the stochastic non-seasonal terms of an ARIMA model but included backshift operators of the seasonal period and multiplied with the non-seasonal elements of the model (Brownlee 2018). SARIMA models allow the user to generate synthetic time series considering cyclical variations in the observed series records (Cox and Miller 1977; Chatfield and Xing 2019). Differencing (d, D) is the difference between a value and a value with lag that is a multiple of m (Hyndman and Athanasopoulos 2018).

*SARIMA, autocorrelation, and forecastin*g.—Each SARIMA model was fitted to monthly values of maximum air temperature, extreme variability in water temperature, and flow volume for each flow-type, which takes in arguments in the following order: data, ARIMA inputs (p, d, q), SARIMA inputs (P, D, Q), and seasonal lag S. A primary advantage of SARIMA is that it automatically differences (d = 1, D = 1) each seasonal time series model consistent with the number of differences recommended by use of the ndiff() function, along with estimating the lowest AIC values and measures of variance (σ^2) for each data set as different choices of p and q were considered. This process generally results in the best-fit model for use in follow-on forecasting. The auto.arima() function without drift was used to determine the order of each SARIMA model (Package "forecast") in which the algorithm conducts numerous iterations and checks in a search for all possible models within the order constraints provided (Hyndman and Khandakar 2008). This stepwise algorithm returns the best "fit" model with the lowest AIC value and automatically differences the time series to make it stationary (Hyndman and Khandakar 2008; Wang et al. 2006). Residuals were used to investigate the relationship of each time point to each previous time point in the distribution of consecutive annual fluctuations in each environmental variable for each time series model. Follow-on assessment was conducted to ascertain if model residuals were random using autocorrelation (ACF) and partial autocorrelation (PACF) plots.

Augmented Dickey-Fuller tests gauged the extent of stationarity in each time series model and Box-Pierce tests (χ^2) determined whether any group of autocorrelations of a time series were different from zero (Box and Jenkins 1970; Fuller 1976; Ljung and Box 1978). Parsimony was used to evaluate performance and to validate each model for each time series using the minimum AIC statistic prior to forecasting. Once the best-fit SARIMA model was identified, the function sarima.for() was used to provide a forecast for the next "future" 36-month time intervals for each time series (Package forecast v8.10). Accuracy of forecasting was evaluated by use of the Mean Absolute Percentage Error (MAPE) measure, which assumes that MAPE $\leq$ 10% or 11 – 20% is considered an excellent to good forecasting estimate, respectively (Lewis 1982; Moreno et al. 2013; Hyndman and Athanasopoulos 2018). After estimating the parameters of each ARIMA model for each environmental variable by flow-type, their adequacy was evaluated graphically by visually inspecting their standardized residuals, ACF graphs, Q-Q plots, and p-values for the Ljung-Box test. Due to space limitations and the number of panels that resulted for each variable by flow-type, a graphic illustration was provided only for the PreROD flow-type.

RESULTS

Historical annual fluctuations in flow volume and water temperature

Principal components analysis of months using environmental variables merged with plot loadings showed a gradation in the seasonal distribution from cold-wet (winter-fall) to warm-dry (spring-summer) climatic conditions along the axis of PC I for the surrounding riverine corridor downriver from Lewiston Dam (Figure 2; Table 1). A total of 91.2% of the variation among months was explained on the first three PCs. As shown by the loading, relationship, and direction of each arrow, all variables vectored heavy and positive along PC I (59.3%) except flow volume, which vectored positive along PC II (23.2%) particularly during the months of May followed by June. Because of the strong correlation between individual measures of air temperature (r_s = 0.890, n = 204) and water temperature (r_s = 0.920, n = 276) only maximum water temperature, extreme variability in water temperature, and flow volume were kept for further analyses to reduce collinearity (Everitt and Hothorn 2011).

From 1994 to 2018, GAM regression identified a significant trend in the smooth term of the annual response curve in maximum air temperature for the upper-most section of the mainstem Trinity River above the Lewiston Water Quality gauge (Table 2). Yet deviance explained was not robust relative to all other environmental variables (Dev.Exp = 0.04%) and the strength of the relationship was extremely weak (r_s = 0.038, p = 0.004, n = 5,752). Significant annual trends in flow volume, maximum water temperature, and extreme variability in water temperature also were not robust. Whereas both measures of water temperature showed positive annual trends, this relationship was negative for flow volume (Figure 3). Additionally, the relationship between annual variation in maximum air temperature was significantly correlated with both water temperature variables but not with flow volume (Table 2; Figure 4). Thus, although the relationship between annual fluctuations in water temperature was significant and positive, both variables were inversely correlated with

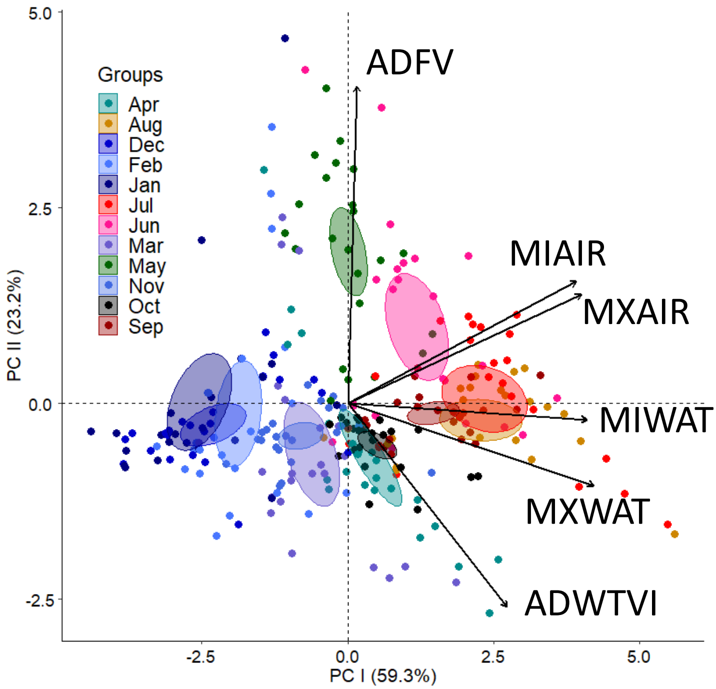


Figure 2. Principal components analysis (PCA) of monthly variation in environmental variables plotted along the first two components: MIAIR = minimum air temperature, MXAIR = maximum air temperature, MIWAT = minimum water temperature, MXWAT = maximum water temperature, ADWTV = extreme variability in water temperature, and ADFV = flow volume. All temperatures in degrees Celsius (°C) and volume of flow in m3/second.

Table 1. Principal components analysis of similarities among months merged with plot loadings. Variables were average daily environmental attributes: MIAIR = minimum air temperature, MXAIR = maximum air temperature, MIWAT = minimum water temperature, MXWAT = maximum water temperature, ADWTV = extreme variability in water temperature, and ADFV = flow volume. All temperatures in degrees Celsius (°C) and volume of flow in m³/second.

	Dimension		
Measure	PC I	PC II	PC II
Variance	3.6	1.4	0.5
Percent	59.3	23.2	8.7
Cumulative percent	59.3	82.5	91.2
	Percent	variable	contribution
Variable	PC I	PC II	PC II
ADFV	0.03	57.26	29.60
MIAIR	20.95	8.46	2.14
MXAIR	21.84	6.78	0.32
MIWAT	22.92	0.15	10.28
MXWAT	24.14	3.81	0.01
ADWTV	10.12	23.54	57.66

Table 2. Summary of generalized additive modeling using GAM regression and the Spearman rank correlation coefficient (r_s) statistics for average daily: ADFV = flow volume, 2) MXWAT = maximum water temperature, and 3) ADWTV = extreme variability in water temperature. All temperatures in degrees Celsius (°C) and volume of flow in m³/second.

	GAM regression model					Spearman rank correlation		
Variable comparison	F-statistic	Ref.df.	p-value	R^2.Adj.	Dev.Exp.	n	r_s	p
Relationships between environmental variables versus year								
MXAIR ~ year	12.5	2.0	< 0.001	0.00	0.4%	5,752	0.038	< 0.004
MXWAT ~ year	164.1	2.0	< 0.001	0.04	3.6%	8,055	0.190	< 0.001
ADWTVI ~ year	766.3	2.0	< 0.001	0.13	13.4%	8,042	0.354	< 0.001
ADFV ~ year	19.1	1.8	< 0.001	0.00	0.8%	8,401	-0.061	< 0.001
Relationships among environmental variables								
MXWAT ~ MXAIR	2,265.0	2.0	< 0.001	0.47	47.1%	5,664	0.672	< 0.001
ADWTVI ~ MXAIR	690.4	2.0	< 0.001	0.14	20.2%	5,664	0.369	< 0.001
ADFV ~ MXAIR	231.4	2.0	< 0.001	0.07	16.5%	5,721	0.211	< 0.001
ADWTVI ~ MXWAT	2,052.0	2.0	< 0.001	0.45	34.7%	8,042	0.626	< 0.001
MXWAT ~ ADFV	111.5	2.0	< 0.001	0.03	2.6%	8,012	-0.143	< 0.001
ADWTVI ~ ADFV	770.2	2.0	< 0.001	0.08	12.3%	7,999	-0.273	< 0.001

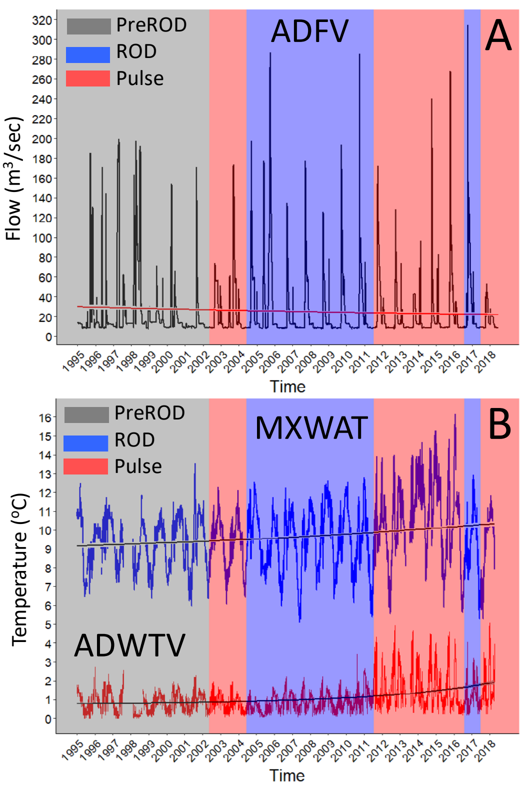


Figure 3. Generalized additive model (GAM) regression plots of annual variation in averaged daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV). Shaded smooth areas indicate 95% point-wise standard error for each curve surrounding each fitted GAM function (centered black line).

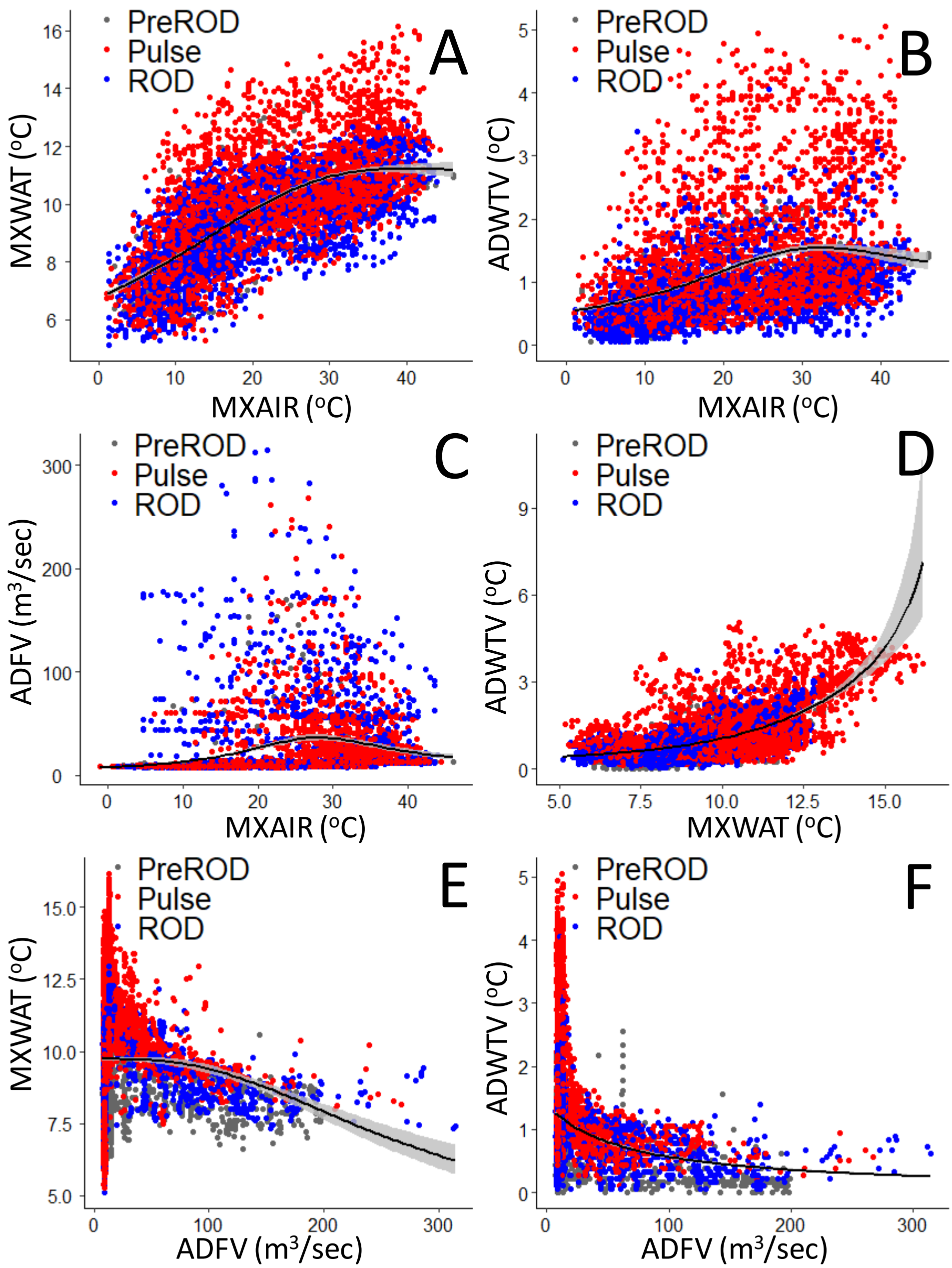


Figure 4. Generalized additive model (GAM) regression plots of annual variation in averaged daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV) versus maximum air temperature (MXAIR; A – C), and between environmental variables (D – F) for each flow-type. Shaded smooth areas indicate 95% point-wise standard error for each curve surrounding each fitted GAM function (centered black line).

fluctuations in flow volume such that when flow increases water temperature decreases as expected. Importantly, Kruskal-Wallis Chi-square (χ^2) rank sum tests revealed significant overall differences for each flow-type hydrograph for each measure of water temperature and flow volume. Follow-on post-hoc pairwise comparisons also showed significant differences between flow-types for all environmental attributes (Table 3). To recap, not only were there significant annual trends in the raw data for the complete historical time series model for each environmental variable as illustrated by GAM regression, there also were significant overall and post-hoc pairwise annual differences between flow-types for each environmental attribute.

Table 3. Comparison of historical time series data based on raw data collected using the Kruskal-Wallis rank sum test (χ^2) for environmental variables by flow-type followed by planned post-hoc Dunn test statistics (Z) of all pairwise comparisons. Variables were average daily: ADFV = flow volume, 2) MXWAT = maximum water temperature, and 3) ADWTV = extreme variability in water temperature. All temperatures in degrees Celsius (°C) and volume of flow in m³/second; p-values were < 0.05 = *, < 0.01 = **, < 0.001 = ***.

ADFV (χ^2 = 176.2, df = 2, p < 0.001***)			
Group(i)	Group(j)	Z	p.adj
PreROD (n = 2,619)	Pulse	9.7	< 0.001***
PreROD	ROD (n = 2,885)	12.8	< 0.001***
Pulse (n = 2,897)	ROD	3.2	< 0.001***
MXWAT (χ^2 = 330.0, df = 2, p < 0.001***)			
Group(i)	Group(j)	Z	p.adj
PreROD (n = 2,396)	Pulse	16.9	< 0.001***
PreROD	ROD (n = 2,908)	3.5	< 0.001***
Pulse (n = 2,751)	ROD	14.1	< 0.001***
ADWTVI (χ^2 = 594.9, df = 2, p < 0.001***)			
Group(i)	Group(j)	Z	p.adj
PreROD (n = 2,383)	Pulse	21.9	< 0.001***
PreROD	ROD (n = 2,908)	2.7	< 0.006**
Pulse (2,751)	ROD	20.1	< 0.001***

Seasonal-Trend-Loess (STL) decomposition

A comparison between the historical raw data (Figure 3) and subdivided monthly time series models using STL decomposition of each environmental variable by flow-type showed loss of detail, resolution, and therefore information content when partitioning on a monthly basis relative to Julian week or daily time series schedules (Figure 5 – Figure 7). Nonetheless, there were obvious qualitative visual differences in the subdivided monthly data among hydrographs. For example, seasonal components produced for each environmental variable revealed distinct visual differences among flow-types for each time series model (Figure 5A – 5C). Each flow-type time series showed a regular pattern of variation in each environmental variable with monthly periodicity. This periodicity indicated that each month had the same repeating pattern that changed every 30 days reflecting long-term

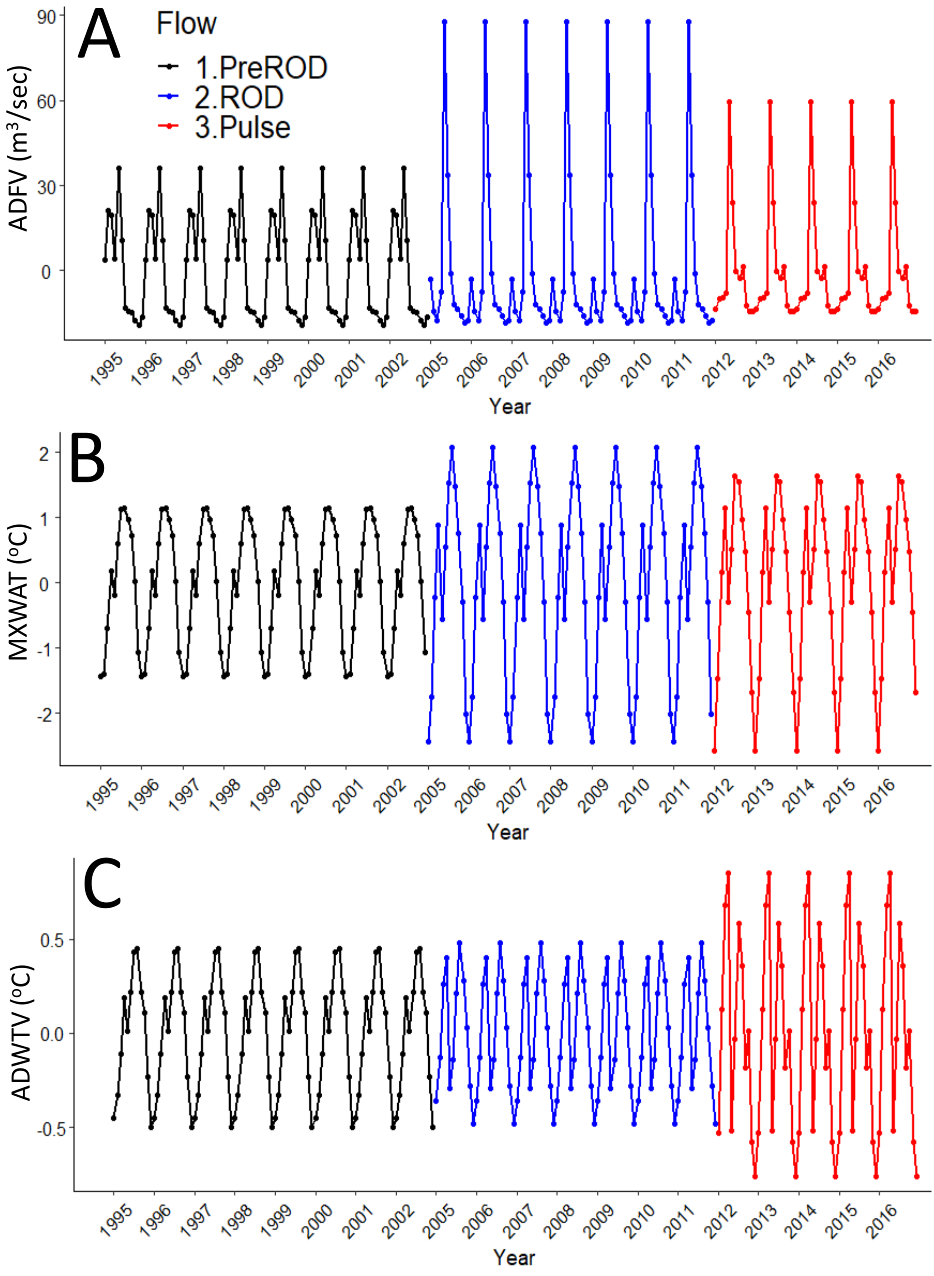


Figure 5. Results of the time series decomposition analyses showing plots of the Seasonal-Trend-Loess (STL) seasonal monthly component for each environmental variable; including average daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV) by flow-type (PreROD, ROD, Pulse).

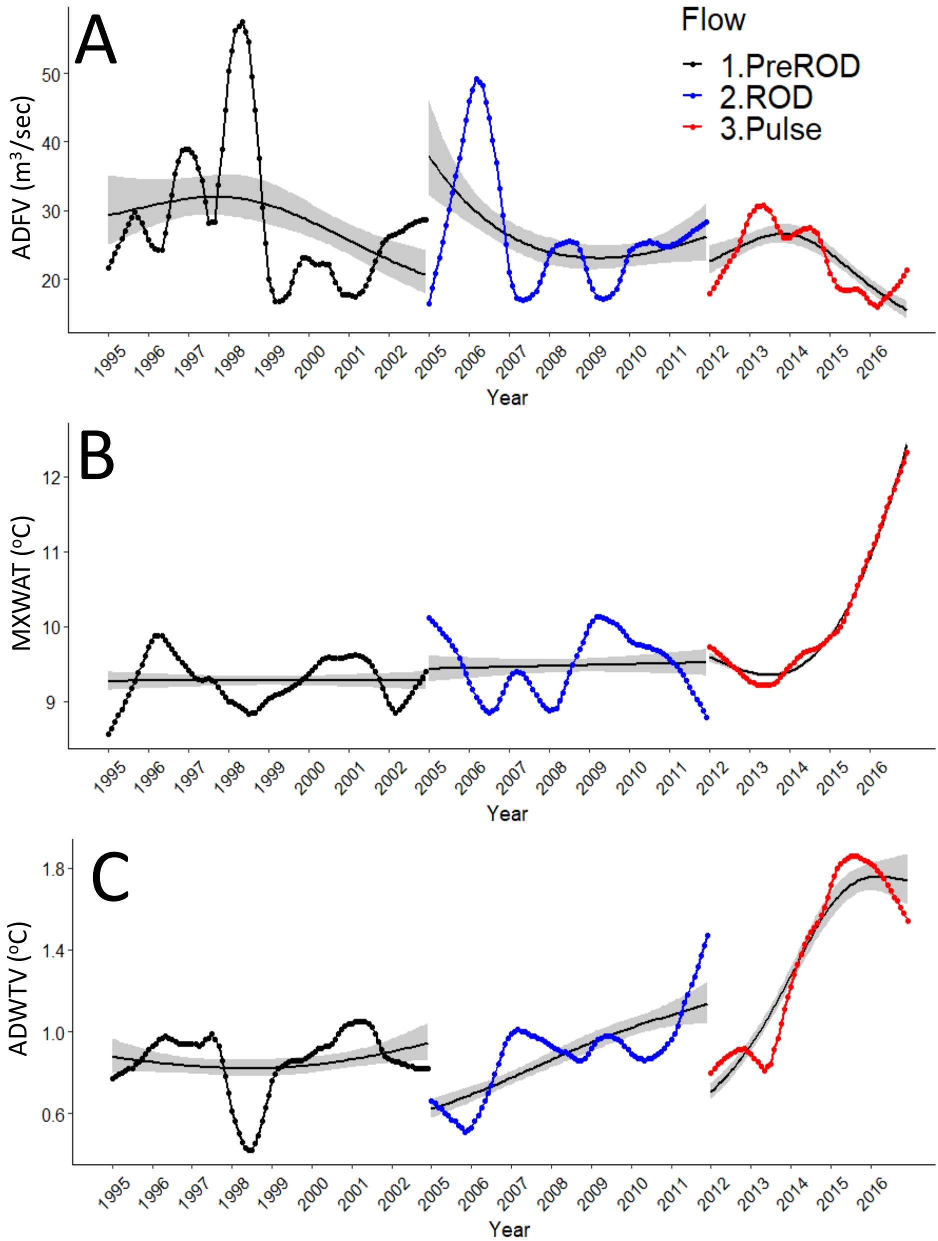


Figure 6. Results of the time series decomposition analyses showing plots of the Seasonal-Trend-Loess (STL) trend monthly component for each environmental variable; including average daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV) by flow-type (PreROD, ROD, Pulse).

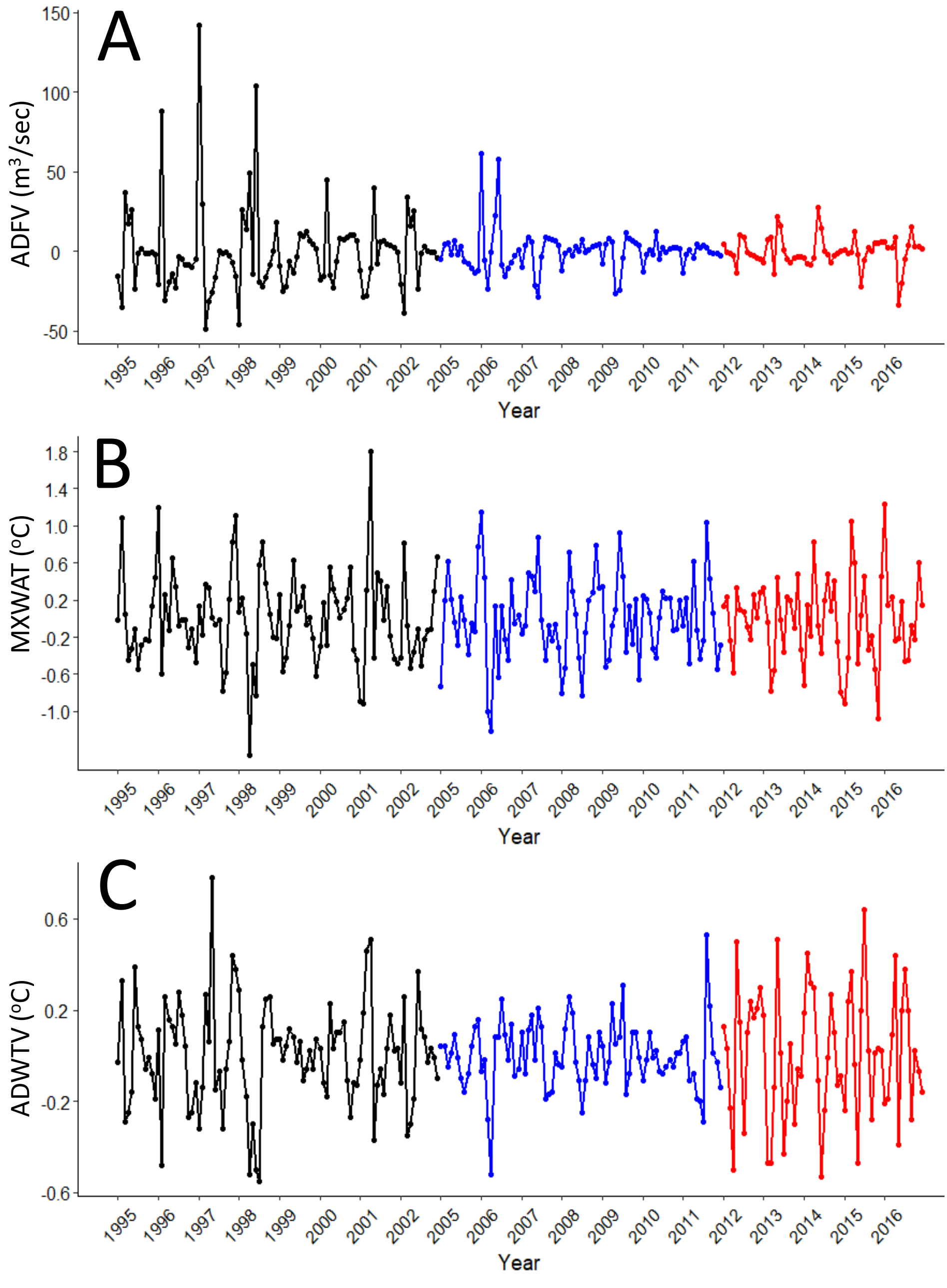


Figure 7. Results of the time series decomposition analyses showing plots of the Seasonal-Trend-Loess (STL) remainder monthly component for each environmental variable; including average daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV) by flow-type (PreROD, ROD, Pulse).

annual patterns in each environmental attribute. Annual variation in the seasonal monthly component of flow volume was much more subdued and proportionally diverse in baseline PreROD flow-types compared to managed hydrographs. Managed flow-types exhibited a pattern of dramatic single spikes for flow volume, particularly in ROD flows followed by Pulse flows; a pattern not characteristic of baseline PreROD hydrographs (Figure 5A). In contrast, annual variation in the seasonal components of both water temperature attributes showed monthly spikes with: 1) greater variation, 2) more proportional diversity in secondary spikes, and 3) spikes with greater magnitude in managed flow-types relative to baseline PreROD hydrographs (Figure 5B and 5C).

Analysis of STL decomposition for the trend component also discovered conspicuous differences among flow-types for each environmental attribute. Not only did the magnitude in volume of flow and water temperature fluctuate dramatically but there were increased annual trends in both water temperature attributes in each managed flow-type hydrograph not observed in modeling of baseline flows (Figure 6; Table 4). For example, the trend component of flow volume exhibited a much more diversified annual pattern in baseline PreROD flow-types compared to managed hydrographs (Figure 6A). In contrast, annual trend components for both water temperature variables were more subdued in baseline flows relative to managed flow-types, with dramatic upward trends in maximum water temperature for ROD flows (2005 – 2011) and both water temperature attributes for Pulse flows (2012 – 2016; Figure 6B and 6C).

Lastly, STL decomposition of the remainder term for baseline flows showed that all environmental variables consisted of erratic monthly sequences with large positive and negative spikes. Such patterns are likely a reflection of stochastic annual flow and thermal regimes typical of seasonal climatic patterns unique to the region, which were not evident

Table 4. Results of GAM regression analyses of the Trend component produced by the Seasonal-Trend-Loess (STL) decomposition method for each environmental variable by flow-type; including average daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV).

Variable	F-statistic	Ref.df.	p-value	R^2.Adj.	Dev.Exp.	n
ADFV				0.141	23.3%	240
PreROD	13.1	2.0	< 0.000***			
ROD	11.9	2.0	< 0.000***			
Pulse	24.2	1.0	< 0.000***			
MXWAT						
PreROD	1.8	1.0	0.180	0.644	62.8%	240
ROD	0.8	1.6	0.360			
Pulse	137.6	2.0	< 0.000***			
ADWTV						
PreROD	4.2	2.0	< 0.016*	0.758	71.8%	240
ROD	41.0	1.5	< 0.000***			
Pulse	132.1	1.9	< 0.000***			

in highly regulated, and systematically and abruptly implemented managed hydrographs (Figure 7A – 7C). And, although STL seasonal and trend decomposition patterns were significantly correlated in most all hydrographs for each environmental variable, there were significance differences between hydrographs for each trend component for all variables, but not in seasonal or random components for any environmental attribute (Table 5). These results clearly indicate that the STL decomposition factor that defined the most conspicuous difference among flow-types for each environmental variable was the trend component.

Inter-quartile range (IQR) calculations

Line and boxplot comparisons of each flow-type illustrate the significant and consistent quantitative differences along the 12-month continuum for each environmental variable described by STL decomposition models and the IQR analyses for each time series segment (Figure 8A – I). The seasonal term for the STL decomposition models showed that IQR values for flow volume in PreROD flows exceeded all other flow-types by a considerable margin. ROD flows showed the largest IQR value for maximum water temperature and Pulse flows had a considerably larger IQR value for extreme variability in water temperature (Table 6). For the trend component, STL decomposition models showed that Pulse flows had considerably larger IQR values for both water temperature variables than any other flow-type hydrograph. For the remainder term, STL decomposition models showed that baseline PreROD flows had the largest IQR value for flow volume consistent with a more random "natural" expression of the historical post-dam flow pattern prior to implementation of managed hydrographs in 2003. In contrast, the large IQR value for the remainder term in extreme variability in water temperature for Pulse flows was likely a biproduct of altered seasonal fluctuations observed in the trend data caused by manipulated hydrographs post-2003 (Figure 6B and 6C).

Time series

Visual assessment of time series plots for PreROD and ROD flow-types showed that an additive model was most appropriate for each environmental attribute because variation remained relatively constant over time and did not depend on the level of the time series (Figure 3). However, results indicate that for Pulse flow hydrographs an additive model was not appropriate for describing variation in each water temperature variable because of the size and increasing annual trend in values amplified with the level of the time series (Figure 3B; Hyndman and Athanasopoulos 2018). In other words, seasonality in the thermal regime of Pulse flows at the beginning was small but became larger in later years. This pattern suggested that a multiplicative decomposition model for both water temperature variables was appropriate for the Pulse flow time series. Thus, in developing a follow-on time series model using SARIMA modeling for forecasting, both water temperature variables required a natural log transformation of the original data for Pulse flows (Hyndman and Athanasopoulos 2018).

Autocorrelation and partial autocorrelation functions examined for each environmental variable by flow-type revealed that each time series model was significantly non-stationary as there were numerous autocorrelations lying outside the 95% confidence limits for all environmental attributes (Figure 9). For each environmental variable, the range in Ljung-Box

Table 5. Kruskal-Wallis rank sum test (χ^2) and Spearman ranked correlation coefficients (r^s) showing overall significance of each decomposition component (season, trend, random), followed by planned post-hoc Dunn test statistics (Z) for all pairwise comparisons between flow-types (PreROD [n = 94], ROD [n = 84] , Pulse [n = 60]). Variables were average daily: ADFV = flow volume, 2) MXWAT = maximum water temperature, and 3) ADWTV = extreme variability in water temperature. All temperatures in degrees Celsius (°C) and volume of flow in m^3/second. Spearman rank correlation coefficients (r^s) are below the diagonal and probabilities above the diagonal; p-values were < 0.05 = *, < 0.01 = **, < 0.001 = ***.

Decomposition components for season patterns									
Flow-type	ADFV			MXWAT			ADWTV		
	χ^2 = 7.3, df = 2, p = 0.030			χ^2 = 0.0, df = 2, p = 1.000			χ^2 = 0.4, df = 2, p = 0.830		
Group(i)	Group(j)	Z	p-adj	Group(j)	Z	p-adj	Group(j)	Z	p-adj
PreROD	ROD	1.5	0.103	ROD	0.2	0.631	ROD	0.1	0.448
PreROD	Pulse	1.4	0.080	Pulse	0.2	1.000	Pulse	0.5	0.448
ROD	Pulse	2.7	0.010*	Pulse	0.1	0.477	Pulse	0.6	0.855
Spearman rank correlations (r_s) for seasonal components									
Flow-type	PreROD	ROD	Pulse	PreROD	ROD	Pulse	PreROD	ROD	Pulse
PreROD	-------	0.001***	0.001***	-------	0.001***	0.001***	-------	0.001***	0.001***
ROD	0.571	-------	0.001***	0.950	-------	0.001***	0.750	-------	0.001***
Pulse	0.540	0.780	-------	0.910	0.970	-------	0.550	0.810	-------
Decomposition components for trend patterns									
Flow-type	ADFV			MXWAT			ADWTV		
	χ^2 = 10.6, df = 2, p < 0.001***			χ^2 = 46.7, df = 2, p < 0.001***			χ^2 = 48.1, df = 2, p < 0.001***		
Group(i)	Group(j)	Z	p-adj	Group(j)	Z	p-adj	Group(j)	Z	p-adj
PreROD	ROD	1.8	0.0587	ROD	3.2	< 0.001***	ROD	1.5	0.062
PreROD	Pulse	3.2	0.002**	Pulse	6.8	< 0.001***	Pulse	6.8	< 0.001***
ROD	Pulse	1.6	0.056	Pulse	3.8	< 0.001***	Pulse	5.2	< 0.001***
Spearman rank correlations (r_s) for trend components									
Flow-type	PreROD	ROD	Pulse	PreROD	ROD	Pulse	PreROD	ROD	Pulse
PreROD	-------	0.176	0.136	-------	0.002**	0.001***	-------	0.001***	0.495
ROD	0.150	-------	0.001***	-0.330	-------	0.001***	0.380	-------	0.801
Pulse	0.190	0.460	-------	-0.610	0.580	-------	-0.090	0.030	-------
Decomposition components for remainder patterns									
Flow-type	ADFV			MXWAT			ADWTV		
	χ^2 = 2.8, df = 2, p = 0.250			χ^2 = 0.1, df = 2, p =0.950			χ^2 = 0.1, df = 2, p = 0.950		
Group(i)	Group(j)	Z	p-adj	Group(j)	Z	p-adj	Group(j)	Z	p-adj
PreROD	ROD	1.5	0.199	ROD	0.3	1.000	ROD	0.1	0.609
PreROD	Pulse	1.3	0.144	Pulse	0.0	0.486	Pulse	0.2	0.465
ROD	Pulse	0.1	0.477	Pulse	0.2	0.615	Pulse	0.3	1.000
Spearman rank correlations (r_s) for remainder components									
Flow-type	PreROD	ROD	Pulse	PreROD	ROD	Pulse	PreROD	ROD	Pulse
PreROD	-------	0.613	0.349	-------	0.745	0.389	-------	0.856	0.397
ROD	0.060	-------	0.484	0.040	-------	0.273	-0.020	-------	0.419
Pulse	-0.120	0.090	-------	-0.110	0.140	-------	-0.110	0.110	-------

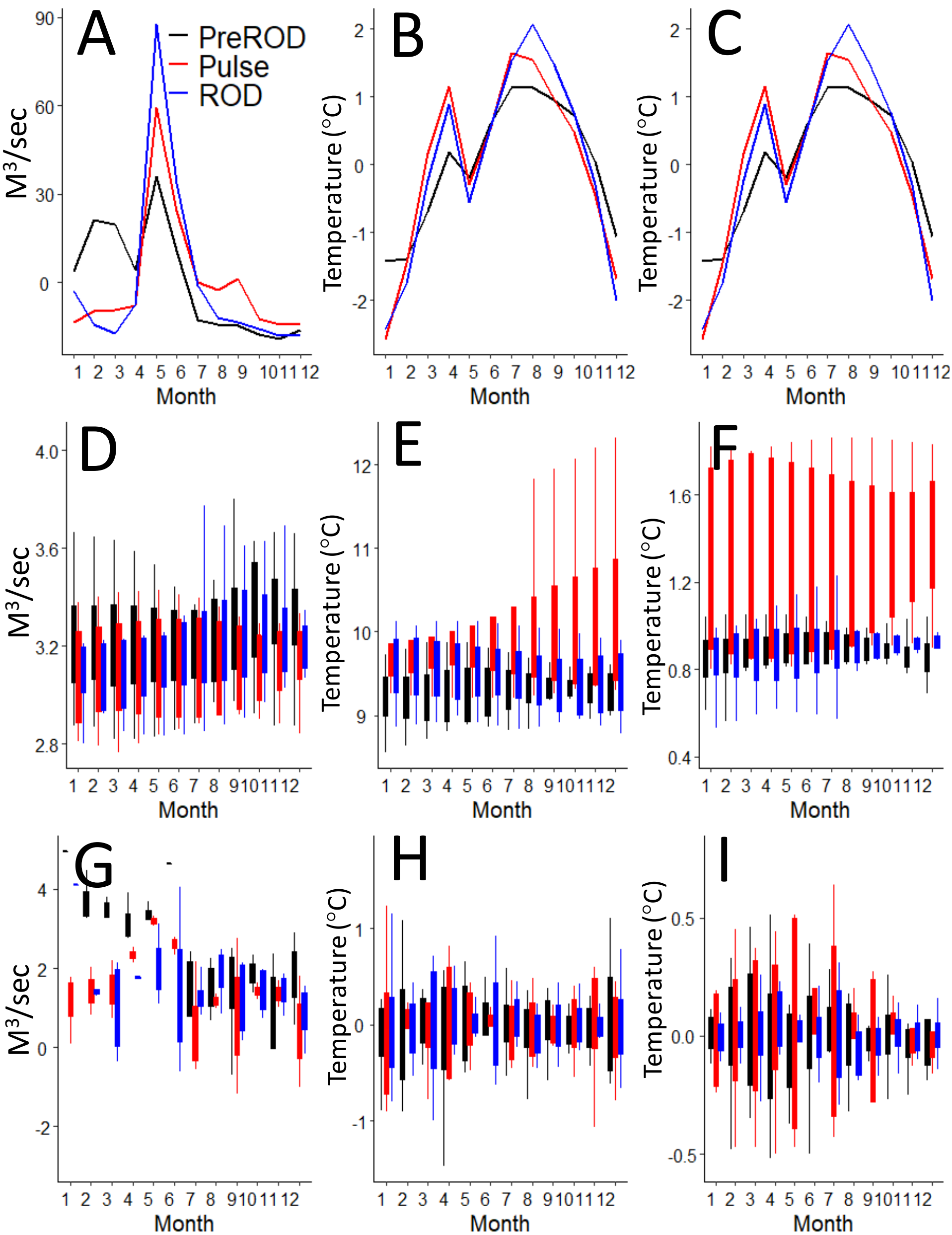


Figure 8. Line and box-plot comparisons of flow-types for each environmental variable based on the Seasonal-Trend-Loess (STL) decomposition analysis, which reflects the level of statistical significance presented in Table 4.

Table 6. Seasonal-Trend-Loess (STL) decomposition component summary of the computed interquartile range (IQR) of x-values, which measure how spread-out points were in the original time series data set for each environmental variable time series by flow-type. Variables were average daily: 1) ADFV = flow volume, 2) MXWAT = maximum water temperature, and 3) ADWTV = extreme variability in water temperature. All temperatures in degrees Celsius (°C) and volume of flow in m^3/second. The higher the IQR the more spread-out the data points; the smaller the IQR the more aggregated the data points are around the mean. Right-hand bars on each STL plot were based on IQRs and allow a relative comparison of the magnitude of variation in each component.

Flow-type	Seasonal			Trend			Remainder (random)		
	1st Quartile	3rd Quartile	IQR	1st Quartile	3rd Quartile	IQR	1st Quartile	3rd Quartile	IQR
ADFV									
PreROD	-15.00	12.80	27.80	22.10	30.60	8.50	-15.80	7.10	22.90
ROD	-16.40	-2.60	13.70	20.50	27.40	6.90	-4.80	4.40	9.10
Pulse	-12.70	0.30	13.00	18.50	26.90	8.50	-4.00	4.60	8.60
MAXWAT									
PreROD	-0.79	0.78	1.56	9.03	9.56	0.53	-0.32	0.30	0.62
ROD	-0.86	1.03	1.89	9.11	9.79	0.69	-0.30	0.25	0.55
Pulse	-0.70	1.02	1.72	9.42	10.58	1.16	-0.33	0.24	0.57
ADWTV									
PreROD	-0.26	0.22	0.48	0.82	0.95	0.13	-0.13	0.12	0.25
ROD	-0.28	0.27	0.55	0.85	0.98	0.12	-0.08	0.08	0.16
Pulse	-0.52	0.41	0.94	0.90	1.75	0.85	-0.20	0.20	0.40

statistics were: 1) flow volume: $\chi^2 = 8.0$ in PreROD flows (df = 1, p = 0.005) to $\chi^2 = 10$ in Pulse flows (df = 1, p = 0.003); 2) maximum water temperature: $\chi^2 = 50$ in PreROD flows (df = 1, p < 0.001) to $\chi^2 = 30$ in Pulse flows (df = 1, p < 0.001); and 3) extreme variability in water temperature: $\chi^2 = 40$ in PreROD flows (df = 1, p < 0.001) to $\chi^2 = 20$ in Pulse flows (df = 1, p < 0.001). All variables by flow-type showed a slow decay at multiple lags of 12 suggesting coherent variance in the relationship indicative of monthly seasonal and cyclic variation in each environmental variable for all flow-types. These results support the earlier assertion of seasonality in each time series model necessitating the need for seasonal differencing with a period of 12. Therefore, rather than manually fitting by ARIMA modeling, best fit models for each flow-type time series were generated using the auto.arima function summarized in Table 7. Following this procedure Ljung-Box statistics indicated that none of the autocorrelations in the time series were different from zero and visual assessments showed that each time series model was stationary (Figure 10). All model parameters were then verified with a graphic illustration described above for each environmental variable grouped by flow-type (Table 7). These results provided no evidence to reject the hypothesis that the distribution of residuals in any of the time series models were not normal. Instead, for each flow-type the distribution of residuals for each environmental model was Gaussian (white noise), which: 1) statistically justified each proposed model, 2) demonstrated how an analysis of time series data may be done accurately, and 3) allowed continued processing of data with the ultimate goal of forecasting estimates of each environmental variable by flow-type using SARIMA.

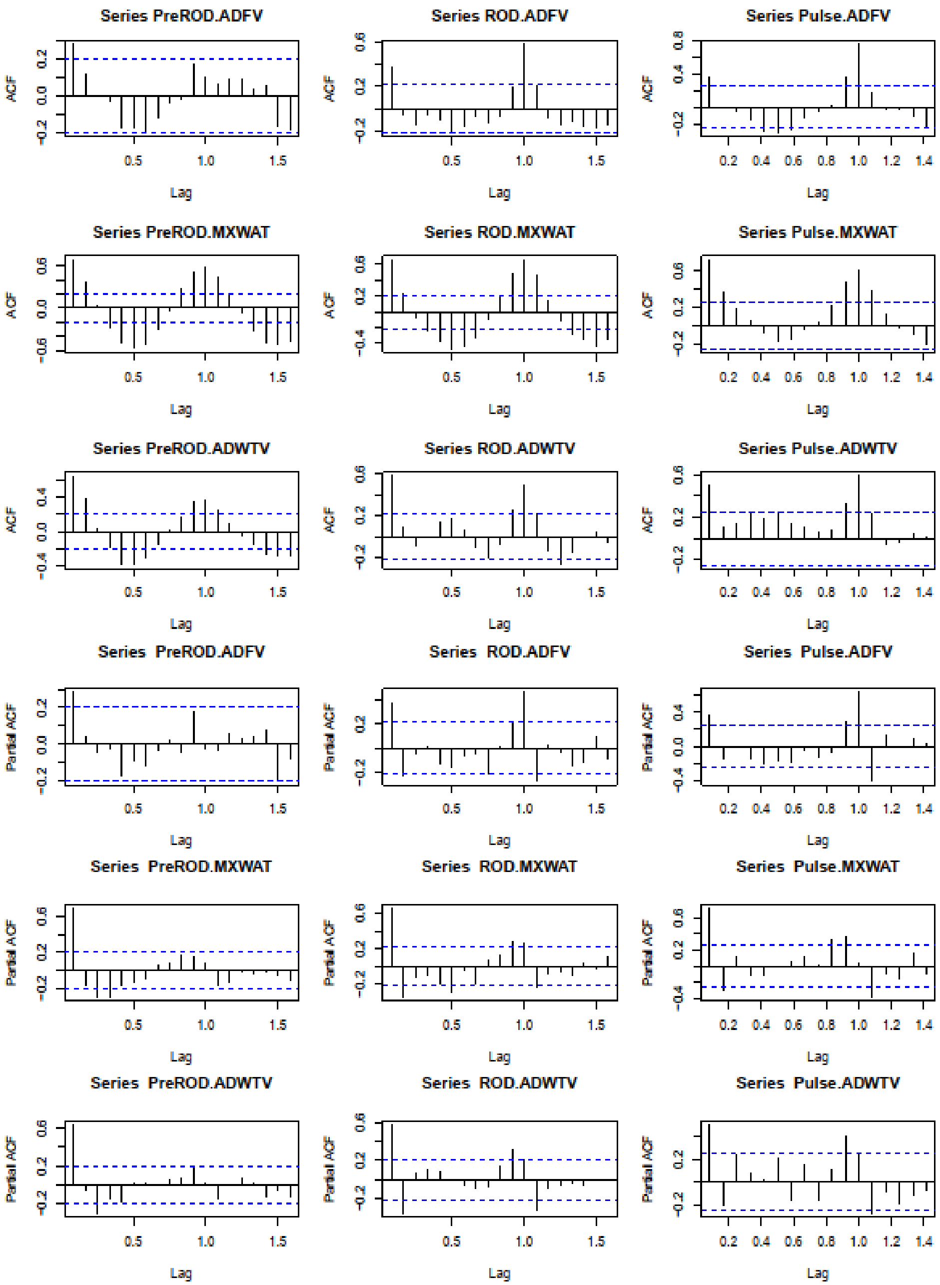


Figure 9. Auto- and partial- autocorrelation functions (ACF, PACF) plots (correlograms) of raw data for averaged daily: flow volume (ADFV), maximum water temperature (MXWAT), and extreme variability in water temperature (ADWTV) by flow-type (PreROD, ROD, Pulse). Plot shows serial correlations that may change over time in each time series dataset where an error at one point in time travels to a subsequent point in time.

Table 7. Parameters used to evaluate the adequacy of each optimal ARIMA model for environmental variables by flow-type: 1) Akaike information criterion (AIC), 2) estimates for variance σ^2 for different SARIMA models, 3) augmented Dicky-Fuller test (ADF), and 4) Box-Pierce tests (χ^2) determined if any group of autocorrelations of a time series were different from zero. Estimated variance and degrees of freedom (df) for each model are also provided.

Flow-type	Variable	Model	AIC -value	Estimate for σ^2	df	ADF test	Ljung-Box test (χ^2)
PreROD	ADFV	$(1,0,0) \times (1,1,0)_{12}$	9.1	1,560.0	82	4.2, lag order = 4, $p < 0.01$	0.05, df = 1, $p = 0.825$
ROD	ADFV	$(1,0,0) \times (1,1,0)_{12}$	7.6	284.0	69	2.7, log order = 4, $p = 0.27$	0.01, df = 1, $p = 0.916$
Pulse	ADFV	$(1,0,0) \times (0,1,0)_{12}$	6.5	159.6	46	3.6, log order = 3, $p = 0.04$	< 0.01. df = 1, $p = 0.998$
PreROD	MXWAT	$(1,0,0) \times (2,1,0)_{12}$	2.0	0.5	80	4.1, lag order = 4, $p < 0.01$	0.63, df = 1, $p = 0.427$
ROD	MXWAT	$(1,0,0) \times (2,1,0)_{12}$	1.9	0.4	68	3.0, lag order = 4, $p < 0.01$	0.04, df = 1, $p = 0.800$
Pulse	MXWAT	$(2,1,1) \times (0,1,1)_{12}$	1.9	0.4	43	6.0, lag order = 3, $p < 0.01$	2.00, df = 1, $p = 0.600$
PreROD	ADWTV	$(1,0,0) \times (2,1,0)_{12}$	0.6	0.1	80	4.2, lag order = 4, $p < 0.01$	0.40, df = 1, $p = 0.525$
ROD	ADWTV	$(1,0,0) \times (2,1,1)_{12}$	0.1	0.0	67	4.0, lag order = 4, $p < 0.01$	2.00, df = 1, $p = 0.200$
Pulse	ADWTV	$(1,0,1) \times (0,1,0)_{12}$	0.8	0.1	45	3.0, lag order = 3, $p < 0.01$	0.09, df = 1, $p = 0.764$

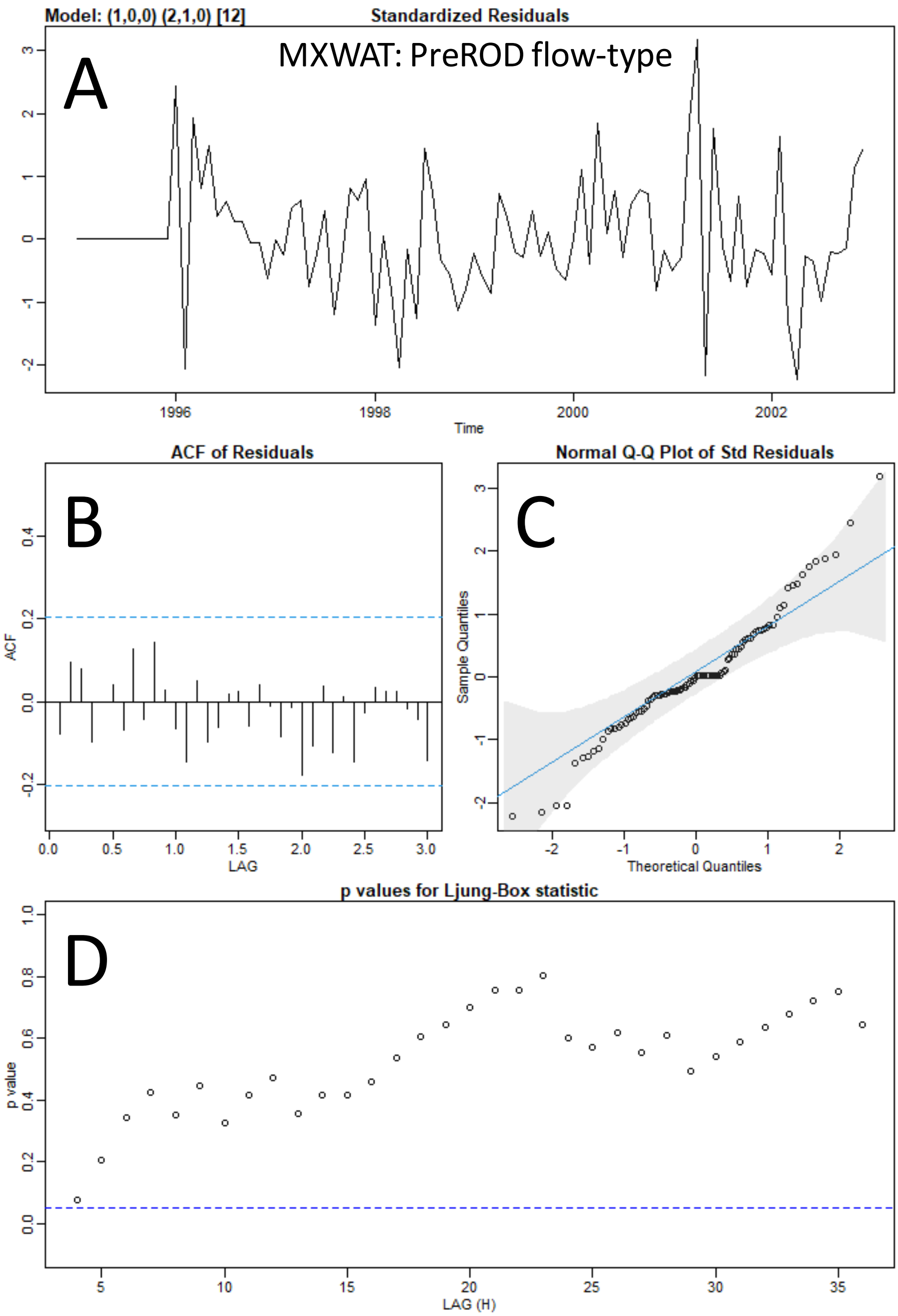


Figure 10. Graphic illustration of the adequacy of parameters used to estimate SARIMA models for each environmental variable using the PreROD flow-type as an example: A) differenced time series data; B) auto correlation function (ACF) plot of standardized residuals showing that all autocorrelations were close to zero as no lag exceeded confidence limits of $p > 0.05$; C) normal Q-Q-plot of standardized residuals along with 95% confidence limits surrounding the diagonal of data points; and D) p-values for the Ljung-Box statistic.

Forecasting using SARIMA

Analysis of the mean absolute percentage error showed that MAPE values for selected flow-type models were: 1) PreROD flow (flow volume = 81.7%, maximum water temperature = 4.9%, extreme variability in water temperature = 45.2%); 2) ROD flow (flow volume = 37.8%, maximum water temperature = 5.0%; extreme variability in water temperature = 15.0); and 3) Pulse flow (flow volume = 23.9%, maximum water temperature = 3.7%, extreme variability in water temperature = 20.0%). These empirical results signaled that the prediction derived from PreROD flow-types was poor for flow volume and extreme variability in water temperature relative to error estimates produced for ROD and Pulse flow hydrographs. This pattern was likely a reflection of the more stochastic nature of a "natural" post-dam and relatively unmanaged flow regime except under extreme flood conditions, which would be expected in a rare "Emergency of Dams" release, which is what happened in 1974. In contrast, error estimates for ROD and Pulse flows were considerably "better" as a reflection of the systematic and regular managed releases linked with anthropogenic hydrographs. The same explanation can also be applied to extreme variability in water temperature for managed flow-types, all of which appear reasonably good as models for each managed time series dataset. Conversely, prediction models based on maximum water temperature were all highly accurate (< 5.0% error). And the empirical results indicated that each model was able to accurately represent the baseline as well as each managed hydrograph time series model.

The predictive power of each SARIMA model was illustrated graphically by forecasting 36 months-ahead of the time series for each environmental variable by flow-type (Figure 11; Table 7). Overlapping blue lines for each "fitted" model (black lines) were substantial as predictions fit well to each time series dataset. Levels of prediction were calculated at 80% and 95% prediction confidence intervals as indicated by the light and dark shaded gray areas surrounding the red prediction line. Forecast values (red lines) were close to real values (black lines) and within the confidence intervals (grey shading). Thus, all monthly points plot very close to the actual prediction. In most models as time progresses beyond the first predicted point, uncertainty tends to increase, hence the prediction boundaries increase in amplitude. Importantly, follow-on forecasting of environmental variables for each flow-type over the next 3 years (36 months) showed: 1) that there were significantly different predictions among segmented flow-type time series; 2) the time span for each model was relatively long, and 3) results provided reasonably accurate predictions. Generally, forecasting further into the future will become less reliable particularly in a highly managed river system. Additionally, the fitted values for each forecast model were relatively close to the observed values. This means that the SARIMA models can be used to forecast future values because their forecasting accuracy is acceptable. Also, some lower predicted confidence limits, but not points, were negative for volume of flow, which is impossible for the flow of a river. In practice, confidences limit below zero are generally truncated when presenting them.

Consistent with graphic illustrations, SARIMA model predictions of future values over the next 36 months exhibited significant differences between flow-types in both the overall distribution of each environmental variable and in planned post-hoc paired comparisons of population mean ranks as reflected in mean values for each variable (Table 8). Noticeably for PreROD and ROD flow-types mean values of flow volume were greater than in Pulse flows, and projections of future ROD flows exceeded both other flow-types. Mean values of maximum water temperature for both PreROD and ROD flow-types were similar but the

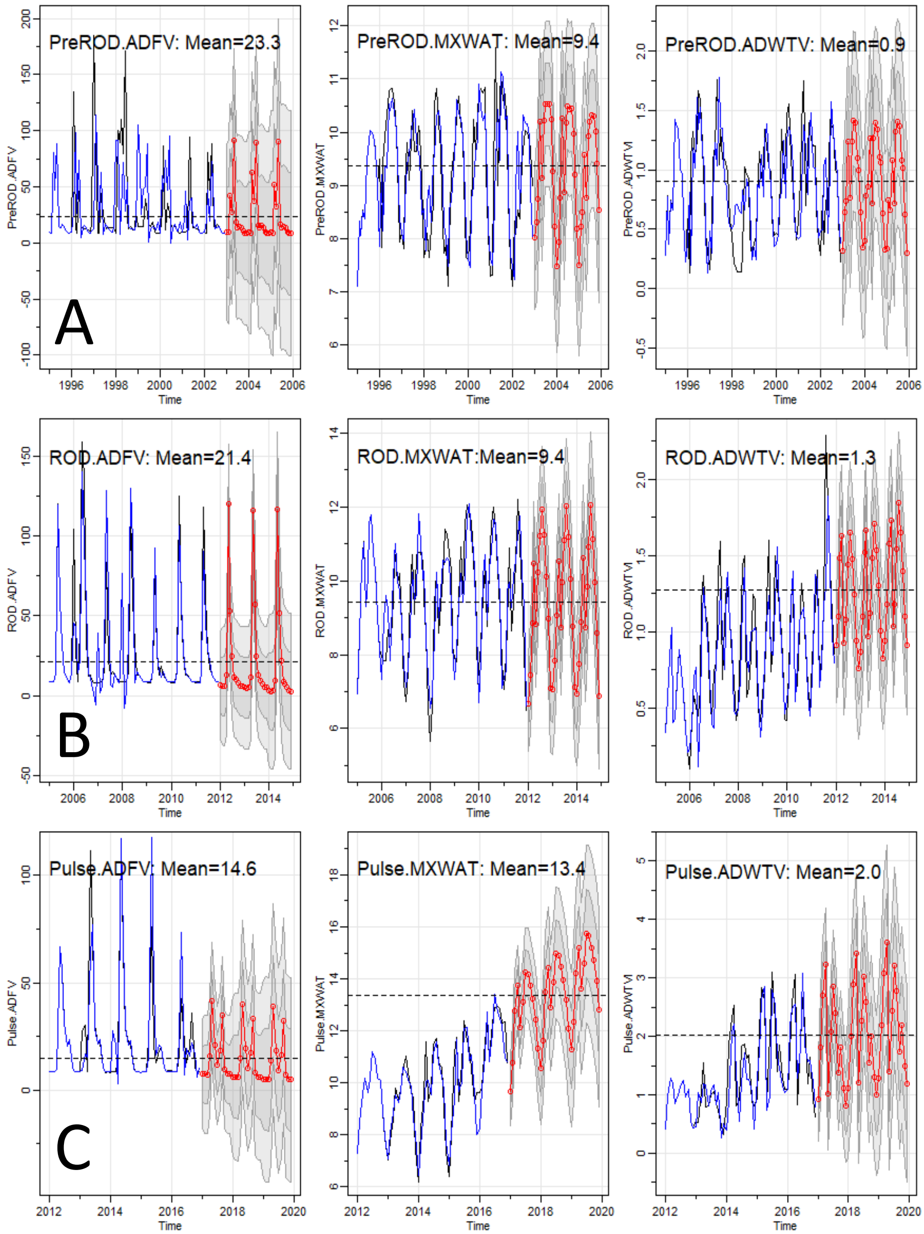


Figure 11. Plots of future values based on SARIMA model forecasting of environmental variable values predicted for the next 36 months. Black colored lines are the original observed time series values and blue lines are the "fitted" values overlaid on top for comparison against the series itself. Means are given for each variable and their position on the graph indicated by a dashed black horizontal line. Dark gray shading represents 80% confidence intervals, light gray shading 95% confidence interval, and red lines and open circles represent predictions for future months. Temperature is found along the y-axis for each variable; A = PreROD flows, B = ROD flows, and C = Pulse flows for each environmental variable (ADFV = flow volume, MXWAT = maximum water temperature, ADWTV = extreme variability in water temperature).

Table 8. Kruskal-Wallis rank sum test (χ^2) and Spearman ranked correlation coefficients (r_s) showing differences in forecasting 36-months into the future using predictions of SARIMA modeling illustrated in Figure 11, followed by post-hoc Dunn tests (Z) of all planed pairwise comparisons between flow-types (PreROD, ROD, Pulse). Variables were average daily: ADFV = flow volume, 2) MXWAT = maximum water temperature, and 3) ADWTV = extreme variability in water temperature. All temperatures in degrees Celsius (°C) and volume of flow in m³/second. Spearman rank correlation statistics are found below the diagonal and probabilities above the diagonal; n = 36 for all flow-type comparisons; p-values were < 0.05 = *, < 0.01 = **, < 0.001 = ***.

Flow-type	**ADFV**			**MXWAT**			**ADWTV**		
	χ^2 = 9.7, df = 2, p = 0.010			χ^2 = 64.7, df = 2, p < 0.001***			χ^2 = 42.4, df = 2, p < 0.001***		
Group(i)	Group(j)	Z	p-adj	Group(j)	Z	p-adj	Group(j)	Z	p-adj
PreROD	ROD	3.0	0.004**	ROD	7.1	< 0.001***	ROD	3.2	0.001***
PreROD	Pulse	2.1	0.026*	Pulse	6.8	< 0.001***	Pulse	6.5	< 0.001***
ROD	Pulse	0.9	0.178	Pulse	0.3	0.380	Pulse	3.4	0.001**
Spearman rank correlation (r_s)									
Flow-type	PreROD	ROD	Pulse	PreROD	ROD	Pulse	PreROD	ROD	Pulse
PreROD	-------	0.001***	0.001***	-------	0.001***	0.001***	-------	0.001***	0.001***
ROD	0.660	-------	0.001***	0.940	-------	0.001***	0.750	-------	0.001***
Pulse	0.600	0.900	-------	0.770	0.820	-------	0.750	0.850	-------

range of variation in projected ROD flows (2006 – 2014) was greater than in future projections of PreROD flows (1995 – 2002). Width of confidence intervals in both flow-types tended to remain narrow throughout the predicted sequence of months, suggesting that the accuracy of the forecast for each PreROD and ROD flow-types effectively held over time through the prediction range.

In Pulse flows, however, not only was the mean and range of variation in maximum water temperature considerably greater than in other hydrographs, but an increasing trend was also evident based on dependent compounding effects of previous observations and errors. In extreme variability in water temperature the same overall pattern of flow-mediated variance was projected in both ROD and Pulse flow-types, which predicted trends of higher and increased extreme variation in water temperatures over the next 36 months, a pattern not seen in PreROD flow-type projections. Finally, width of confidence intervals in both managed flow types for each water temperature variable increased in future predictions as a direct function of greater variability associated with annually managed hydrographs. Generally, but not always, as the period between the date of a flow forecast and the actual forecast period narrows there also will be a corresponding reduction in model error.

DISCUSSION

In this study we focused on time series analysis and forecasting seasonal river flow volume and water temperature using SARIMA modeling. Results show that use of subcomponents of the historical post-dam time series specific to managed flow-type hydrographs keyed to in-river geomorphologic restoration actions significantly misrepresented the: 1) time series characteristics, 2) seasonal-trend decomposition patterns, and 3) forecast accuracy compared to the baseline pre-restoration time series model. These results effectively nullify use of managed hydrographs in predicting future forecasts for flow-fish planning and management purposes. Particularly revealing were significant differences in environmental time series data between flow-types in: 1) seasonal patterns among subcomponents of the historical hydrograph as illustrated in GAM regressions; 2) nonparametric methods; 3) STL decomposition analyses of season, trend, and random effects using both IQR values and nonparametric methods; and 4) future values of the thermal regime predicted by SARIMA forecasting

In virtually every comparison, managed flow-type hydrographs showed significant increases in maximum water temperature and extreme variability in water temperature. These patterns were particularly characteristic of Pulse flows relative to baseline PreROD flows. Grabowski et al. (2014) suggested that while small shifts in flow releases may be a function of climate change, major shifts usually reflect human interventions, with dam hydropeaking[4] being a distinct indicator of artificiality in the flow regime that impact monthly and daily flows (Greimel 2018). Such flow-effects equate roughly to the implementation of ROD flow hydrographs individually and in combination with companion Pulse flows into the upper mainstem of the Trinity River for geomorphic restoration purposes since 2003, due to longer in duration and volumetric proportion relative to individual pulsed flow augmentation releases.

Based on our results, we suggest that use of time series modeling and forecasting of seasonal trends in environmental data associated with large riverine systems without reference to continuous intrinsic periods within the historical time series model, where flow is anthropogenically manipulated, will be misleading when managers attempt to evaluate the overall characteristics and subsequent future forecasting that derive from such models. Results of our analysis have the potential to provide resource agencies with additional insight into strategic flow-mediated thermal planning that mimics a "natural" or regime standard for water temperature necessary to: 1) facilitate efficient use of water resources notwithstanding changing climate; and 2) prioritize management strategies and scheduling of flows for in-river restoration activity to increase efficiency in management of fisheries, aquatic habitat conservation, and water resources (Stanford et al. 1996; Poff et al. 1997; Fausch et al. 2002; Poole et al. 2004)

Time series and forecasting considerations

The efficacy of these results suggests that time series analysis using SARIMA methodology is appropriate for modeling hydrological and water temperature data in the upper Trinity River in which the data exhibit autocorrelation with time in combination with semi-parametric regression (GAM). Application of linear models generally do not allow identification of nonlinear characteristics of hydrological data. This distinction is important when working with changes in the variance of environmental variables that fluctuate overtime, as their application may not be suited for modeling stochastic nonstationary data (Machiwal and Jha 2006; Nazir et al. 2018). River flow and other hydrologically-related variables frequently exhibit nonlinear behavior. In modeling, resource managers should consider both deterministic (algebraic) and stochastic parts of this parameter when making

4 Hydropeaking is a unique form of flow regulation, in that it introduces frequent, short duration, artificial flow events to the river. The impacts of hydropeaking on channel size and morphology are highly dependent on the size and frequency of hydropeaks in relation to size of geomorphological effective flows prior to regulation. http://wiki.reformrivers.eu/index.php/Hydropeaking

appropriate decisions for the purpose of fish-flow and water resource management. These conditions also apply to regionally specific assessments of climate change. Conversely, use of nonlinear models, such as GAM in combination with the Box-Jenkins method, allow options for nonlinear modeling and consideration of autocorrelations among variables as well as lag-lead relationships between variables. Adding variable effects of reach-specific air temperature, water temperature, precipitation, or other auxiliary environmental co-variates using a multivariate approach to forecasting also will assist in attaining a solid presumption of a cause-effect relationship in historical time series analysis for flow and restoration managers to act upon.

Understanding variability and the limitations associated with management of riverine systems is essential for addressing the ever-increasing anthropogenic needs for water and the inescapable reality of climate change, which has advanced a theoretical construct known as "functional flows" (Zimmerman et al. 2020; CEFF 2020). This paradigm seeks to mimic natural flow regimes by incorporating regionally-specific ecological, geomorphic, and biogeochemical processes into a flow prescription that protects and supports relevant foundational physical and ecological processes, which it is hoped will sustain resident biological communities in some generally unknown, but viable capacity (Poff et al. 1997; CEFF 2020; Zimmerman et al. 2020). Increasingly, the extent of flow-mediated variance in water temperature also is a key component of regulations derived from valuations of water quality (Coutant 1999). Research focused on the implications of altered thermal variability can provide improved identification of conservation priorities, management of dam-derived hydropower, and sustainable fisheries management and restoration outcomes. Use of metrics that measure the extremes in thermal variance or evaluate the range of mean values rather than only the mean is also an important consideration in understanding variability in flow dynamics. This need is particularly relevant when evaluating the potential effects of short-duration pulsed flows, which have not received the attention needed regarding flow-fish effects (Sullivan and Hileman 2020).

Inevitability this approach will require accurate understanding of the underlying dynamics of riverine flow that may not be possible by use of statistical forecasting methods using conventional linear regression. River restoration and conservation planning would benefit from consideration of the natural temperature regime of riverine systems in their full complexity, rather than create predictions based on total temperature units delivered in a short period of time or at lethal thresholds (Poole et al. 2004; Steel et al. 2012; Romberger and Gwozdz 2018). Future climates will invariably differ from current climates with respect to thermal variance. Yet there is remarkably little research exploring these effects on individual or population-level fitness of in-river spawning anadromous salmonids or the complexity and dynamics of aquatic community food webs subjected to anthropogenic management of riverine flows in the Trinity River and elsewhere (Hughes and Murdoc 2017).

Forecasting outflow volume and temperature variability are essential tools for planning and management of dam operations as these activities directly influence the availability of water and the variability in the resulting thermal regimes created. Efforts to return natural variability to regional riverscapes through formation of a diverse mosaic of habitats within riverine and riverscape communities is a vital precursor to being realistic in our expectations of the range of possible restoration outcomes (Peipoch et al. 2015; Yarnell et al. 2015). Such efforts are irrespective of whether the goal is to: 1) re-naturalize flows and thermal regimes as a basis for understanding variability, uncertainty, and unpredictability in concert with restoration initiatives; or 2) implementation of a functional flow design to address realistic anthropogenic needs for water resources, while simultaneously attempting to maintain the regional biotic and physical integrity of the riverscape. Thus, it is imperative to understand the dynamics of riverine flows over both historical and recent time scales as a prerequisite to restoration and effectiveness monitoring of our fisheries resources.

Resource management considerations

Accurate forecasting of streamflow is a fundamental issue of interest to water resources engineers, hydrologists, and fisheries scientists. Identification of accurate and reliable analyses to model future river flow is an important precondition for successful planning and management of water resources upon which fisheries resources depend. From a practical perspective, use of reliable models to forecast river flow could be instrumental for regional fisheries management and water resources planning keyed to the upper Trinity River watersheds. Detection of trends and stationarity is a major focus of past hydrological and climatological time series analyses with a wide application of semi-parametric regression methods such as GAM. Yet forecasting highly accurate estimates of the volume and variability in the thermal regime of riverine flow remains problematic given the nonlinearity and uncertainty hidden in the historical data, which requires an approach with high forecasting precision and efficiency for effective application.

Future needs of water resource and fisheries management in the Trinity River would benefit significantly by placing environmentally-driven time series models into perspective relative anthropogenically-driven flow management hydrographs as part of flow-fish and river restoration management. Research designs that incorporate multivariate and dynamic conditional correlation methods would provide additional insights into the relationships between more “naturally” managed flow and thermal regimes relative to those forced upon the system by a highly manipulative anthropogenically-induced flow strategy (Pool 2004; Caissie 2006). In the Trinity River managed flows have not been mirrored historical regional hydrographic patterns linked to upland watersheds of the Trinity River and Klamath Basin. Such considerations need to be fully vetted with the overarching management goals of flow-mediated hydrographs designed to mimic an “unimpeded”, “natural”, e-flow, or “functional” flow post-dam management strategy (Yarnell et al. 2015), while simultaneously accommodating riverine: 1) thermal criteria, 2) in-river restoration, 3) conservation of biotic communities, and 4) water conservation policies unique to the particular river system in northern California and elsewhere.

Palmer and Ruhi (2019) correctly state that effective river restoration requires advancing our mechanistic understanding of how flow regimes affect biota and ecosystem processes. Any attempt to derive insight into sustainable flow-fish management strategies for hatchery- and natural-area spawning salmonids based on time series modeling of post-2003 ROD or Pulse hydrology will be less reliable than those based on long-term data collected simultaneously from: 1) unobstructed headwaters of the Trinity River above the Trinity Reservoir in combination with 2) the primary free flowing tributaries of the mainstem Trinity River below Lewiston Dam. A prerequisite to management of riverine systems should be the incorporation of these kinds of “baseline” data into long-term restoration strategies coincidental with natural and historical events that shape the regions riverine hydrology to which life histories of resident salmonids have adapted. Distinguishing these elements in flow management could help managers restore ecologically important facets of the flow regime.

We recommend that in projecting future flow conditions, restoration managers should consider use of environmental information that optimizes the operational time span of continuous uninterrupted and sequential collection of data to be as long-term as possible. Reliable predictions may be obtained over relatively shorter time spans by continuous operation, monitoring, and data gathering at gauging stations within the focal reach (Papalaskaris and Kampas 2017). Sequentially, the most accurate and complete data records come from river gauging stations, followed by supplemental information obtained from remotely sensed data and documentary sources (Grabowski et al. 2014). Numerous indicators may be extracted from river flow records (e.g., average, extreme flows, and their timing) and used to estimate hydrological alteration. Gauging station records minimally spanning several decades (e.g. $\geq$ 30 years) are typically necessary for this type of analysis. This approach will allow the entire historical time series can be analyzed to investigate temporal trends, in magnitude, frequency, timing, duration, and rate of change in the flow-dynamic (Grabowski et al. 2014).

Further, flows can be subdivided into time periods related to significant changes in the historical flow regime (e.g., baseline pre-dam construction, post-dam construction, in-river restoration flows, or fish-flows). Subdivided flows can also be applied to an “unimpeded”, “natural”, e-flow, “functional” flow post-dam management strategy, or other prescriptions attributable to natural flow abstractions. For the upper Trinity River, the post-dam “natural” riverine environment evaluated by the STL modeling using annual flow and thermal variance attributes was illustrated in the seasonal, trend, and remainder patterns of monthly variation produced by baseline PreROD flow-type models. However, we caution that use of subdivided historical raw data can result in loss of detail, resolution, and therefore information content when partitioning monthly data compared to use of Julian weeks or a daily schedule. Clearly, results of out study would have been benefited had the data been detailed enough to accommodate smaller subdivisions.

Once established, data can be feed into a standardized modeling procedure as additional raw data becomes available to develop a more accurate and reliable long-term model for use in monitoring seasonal variation in flow-fish thermal regimes (Grabowski et al. 2014). These datasets will facilitate annual: 1) monitoring, modeling, and gaming of hydrological conditions; 2) monitoring of riverine thermal regimes for use in flow-fish management, assessment of in-river fitness metrics (fertility, productivity) for both hatchery- and natural-area spawning salmonids in cooperation and coordination with hatchery operations; 3) facilitate scheduling of specific river restoration actions; and 4) permit annual monitoring and assessment of climate change effects within the mainstem of Trinity River in support of the above management recommendations.

In conclusion, time series modeling and forecasting with flow and thermal data developed during the period of managed hydrographs (post-2003) will greatly detract from the ability to accurately predict future thermal regimes in the upper Trinity River. Sustainable flow-fish management for hatchery- and natural-area spawning salmonids should consider watershed and unobstructed tributary characteristics in assessing the impacts to resident salmonids, which automatically incorporate the annual effects of climate change on riverine systems within the greater Trinity Basin. Time series modeling using such criteria would facilitate development of a flow management strategy for river restoration that fit the historical characteristics of the regional watershed while simultaneously accommodating management of fisheries and other aquatic resources experiencing changing climatic conditions now and into the future.

ACKNOWLEDGEMENTS

We thank Robert Nau (Duke University) for discussions of the use of consecutive sequence versus disjunct sequences in the time series analysis of managed flows. We are also grateful for the editorial comments and suggestions from several anonymous reviewers.

LITERATURE CITED

Adeloye, A. J., and M. Montaseri. 2002: Preliminary streamflow data analyses prior to water resources planning study. Hydrological Sciences Journal 47:679–692.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 in B. N. Petrov and F. Csáki, editors. 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, Budapest, USSR.

Attar, N. F., Q. B. Pham, S. F. Nowbandegani, M. Rezaie-Balf, C. M. Fai, A. N. Ahmed, S. Pipelzadeh, T. D. Dung, P. T. T. Nhi, D. N. Khoi, and A. El-Shafie. 2020. Enhancing the prediction accuracy of data-driven models for monthly streamflow in Urmia Lake Basin based upon the autoregressive conditionally heteroskedastic time-series model. Applied Science 10:1-20. Available from: http://www.mdpi.com/journal/applsci

Bari, S. H., M. T. Rahman, M. M. Hussain, and S. Ray. 2015. Forecasting monthly precipitation in Sylhet City using ARIMA model. Civil and Environmental Research 7:69–77.

Benjamini, Y., and Y. Hochberg. 2000. On the adaptive control of the false discovery rate in multiple testing with independent statistics. Journal of Educational and Behavioral Statistics 25:60–83.

Box, G. E. P., and G. M. Jenkins. 1976. Time series analysis: forecasting and control. Holden-Day, San Francisco, CA, USA.

Brownlee, J. 2018. How to create an ARIMA model for time series forecasting in Python. Machine Learning Mastery. Available from: https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, NY, USA.

Caissie, D. 2006. The thermal regime of rivers: a review. Freshwater Biology 51:1389–1406.

California Department of Water Resources (DWR). 2020. California Department of Water Resources Data Exchange Center. Available from: http://cdec.water.ca.gov

California Environmental Flows Framework (CEFF). 2020. Functional flows approach. Available from: https://ceff.sf.ucdavis.edu/functional-flows-approach

Carter, J. L, V. H. Resh, D. M. Rosenberg, and T. B. Reynoldson. 2006. Biomonitoring in North American rivers: a comparison of methods used for benthic macroinvertebrates in Canada and the United States. Pages 203–228 in G. Ziglio, G. Flaim, M. Sillgardi, eds. Biological Monitoring of Rivers. John Wiley and Sons Ltd, New York, NY, USA.

Casella, G., and R. L. Berger. 2002. Statistical Inference. Duxbury, Pacific Grove, CA, USA.

Chatfield, C., and H. Xing. 2019. The analysis of time series an introduction with R. 7th edition. Chapman and Hall/CRC, London, UK.

Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning. 1990. STL: a seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics 6:3–73.

Cleveland, W. S., E. Grosse, and W. M. Shyu. 1992. Local regression models. Pages 309–376 in J. M. Chambers and T. Hastie, editors. Statistical Models. S. Chapman and Hall, New York, NY, USA.

Corder, G. W., and D. I. Foreman. 2014. Nonparametric statistics: A step-by-step approach, John Wiley and Sons, Inc., Hoboken, NJ, USA.

Coutant, C. C. 1999. Perspective on temperature in the Pacific Northwest's fresh water. Environmental Sciences Division, Publication No. 4849, Oak Ridge National Laboratory, TN, USA.

Cox, D. R., and H. D. Miller. 1977. The theory of stochastic processes. 1st edition. Chapman and Hall/CRC, Boca Raton, FL, USA.

Diankha, O., and M. Thiaw. 2016. Studying the ten years variability of *Octopus vulgaris* in Senegalese waters using generalized additive model (GAM). International Journal of Fisheries and Aquatic Studies 2016:61–67.

Everitt, B. S., and T. Hothorn. 2011 An introduction to applied multivariate analysis with R. Springer, New York, NY, USA.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: Bridging the gap between research and conservation of stream fishes. BioScience 52: 483–498.

Fuller, W. A. 1976. Introduction to Statistical Time-series. John Wiley and Sons, New York, NY, USA.

Grabowski, R. C., N. Surian, and A. M. Gurnell. 2014. Characterizing geomorphological change to support sustainable river restoration and management. WIREs Water 1:483-512. doi: 10.1002/wat2.1037. Available from: https://www.academia.edu/7772927/Characterizing_geomorphological_change_to_support_sustainable_river_restoration_and_management?email_work_card=view-paper

Greimel, F., L. Schülting, W. Graf, E. Bondar-Kunze, S. Auer, B. Zeiringer, and C. Hauer. 2018. Hydropeaking impacts and mitigation. Pages 91–110 in S. Schmutz and J. Sendzimir editors. Riverine Ecosystem Management: Science for Governing Towards a Sustainable Future. Aquatic Ecology Series 8. Springer Open, Cham, Switzerland.

Hallock, R. J., R. F. Elwell, and D. H. Fry. 1970. Migrations of adult king salmon, *Oncorhynchus tshawytscha*, in the San Joaquin Delta, as demonstrated by use of sonic tags. California Fish and Game Bulletin 151.

Hastie, T., and R. Tibshirani. 1990. Generalized additive models. Statistical Science 1:297–301.

Hildebrand, M., and G. E. Goslow. 2001. Analysis of vertebrate structure. Wiley, New York, NY, USA.

Hill, T., and P. Lewicki. 2007. Statistics: methods and applications. StatSoft, Tulsa, OK, USA.

Holčík, J. 1996. Ecological fish production in the inland delta of the Middle Danube, a floodplain river. Environmental Biology of Fishes 46:151–165.

Hughes, M. S., and Murdoch A. R. 2017. Spawning habitat of hatchery spring Chinook Salmon and possible mechanisms contributing to lower reproductive success. Transactions of the American Fisheries Society 146:1016–1027.

Hyndman, R. J., and Y. Khandakar. 2008. Automatic time series forecasting: The forecast package for R. Journal of Statistical Software 26.

Hyndman, R. J., and G. Athanasopoulos. 2018. Forecasting: Principles and Practice. 2nd edition. OTexts, Melbourne, Australia.

Lewis, C. D. 1982. Industrial and Business Forecasting Methods. Butterworths, London, UK.

Ljung, G. M., and G. E. P. Box. 1978. On a measure of lack of fit in time series models. Biometrika 65:297–303.

Machiwal, D., and M. K. Jha. 2006. Time series analysis of hydrological data for water resources planning and management: a review. Journal of Hydrology and Hydromechanics 54:237–257.

Madsen, H., and P. Thyregod. 2011 Introduction to General and Generalized Linear Models. Chapman and Hall/CRC, Boca Raton, FL, USA.

Magneson, M. D., and C. D. Chamberlain. D. 2014. The influence of Lewiston Dam releases on water temperatures of the Trinity River and Lower Klamath River, CA, April to October 2014. U.S. Fish and Wildlife Service, Arcata Fish and Wildlife Office, Arcata Fisheries Data Series Report Number DS 2015-41, Arcata, California. Available from: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=2ahUKEwjtvbXw_o7gAhWsCTQIHVTQAccQFjADegQIBxAC&url=http%3A%2F%2Fodp.trrp.net%2FDataPort%2Fdoc.php%3Fid%3D2211&usg=AOvVaw16FSHoaPjluYNVZA5yWnJ9

Makridakis, S., S. Wheelwright, and V. McGee. 1983. Forecasting Methods and Applications. 2nd edition. Wiley and Sons, New York, NY, USA.

Martínez-Acosta, L., J. P. Medrano-Barboza, Á. López-Ramos, J. F. R. López, and A. A. López-Lambraño. 2020. SARIMA approach to generating synthetic monthly rainfall in the Sinú River watershed in Colombia. Atmosphere 112:16.

McCullough, D. A. 1999. A review and synthesis of effects of alterations to the water temperature regime on freshwater life stages of salmonids, with special reference to Chinook Salmon. Seattle, Washington, U.S. Environmental Protection Agency, Region 10. Available from: https://www.critfc.org/blog/reports/a-review-and-synthesis-of-effects-of-alterations-to-the-water-temperature-regime-on-freshwater-life-stages-of-salmonids-with-special-reference-to-chinook-salmon/

McDonald, J. H. 2014. Handbook of biological statistics. Sparky House Publishing, Baltimore, MD, USA.

Moreno, J. J., M. Moreno, A. P. Pol, A. S. Abad, and B. C. Blasco. 2013. Psicothema 4:500–506. Available from: https://www.researchgate.net/publication/257812432_Using_the_R-MAPE_index_as_a_resistant_measure_of_forecast_accuracy.

Nau R. 2017. ARIMA models for time series forecasting. Available from: https://people.duke.edu/~rnau/seasarim.htm

Nazir, H. M., I. Hussain, I. Ahmad, M. Faisal, and I. M. Almanjahie. 2018. An improved framework to predict river flow time series data. PeerJ 7:1–22.

Nychka. 1988. Bayesian confidence intervals for smoothing splines. Journal of the American Statistical Association 83:1134–1143.

Olden, J. D., and R. J. Naiman 2010. Incorporating thermal regimes into environmental flows assessments: modifying dam operations to restore freshwater ecosystem integrity. Freshwater Biology 55:86–107.

Pajuelo, J. G., and J. M. Lorenzo. 1995. Analysis and forecasting of the demersal fishery of the Canary Islands using an ARIMA model. Scientia Marina 59:155–164.

Palmer, M., and A. Ruhi. 2019. Linkages between flow regime, biota, and ecosystem processes: implications for river restoration. Science 20:9–20.

Papalaskaris, T., T. Panagiotidis, and A. Pantrakis. 2016. Stochastic monthly rainfall time series analysis, modeling and forecasting in Kavala city, Greece, North-Eastern Mediterranean Basin. Procedia Engineering 162:254–263.

Papalaskaris, T., and G. Kampas. 2017. Time series analysis of water characteristics of streams in Eastern Macedonia–Thrace, Greece. European Water 57:93–100.

Peipoch, M., M. Brauns, F. R. Hauer, M. Weitere, and H. M. Valett. 2015. Ecological simplification: human influences on riverscape complexity. BioScience 65:1057–1065.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. Stromberg. 1997. The natural flow regime: a paradigm for riverine conservation and restoration. Bioscience 47:769–784.

Poole, G. C., J. B. Dunham, D. M. Keenan, S. T. Sauter, D. A. McCullough, C. Mebane, J. C. Lockwood, D, A. Essig, M. P Hicks, D. J. Sturdevant, E. J. Materna, S. A. Spalding, J. Risley, and M. Deppman. 2004. The case for regime-based water quality standards. BioScience 159:155–161.

Prasad, A. 2019. Prediction and analysis of time series data using tensorflow. Towards Data Science, 15 November. Available from: https://towardsdatascience.com/prediction-and-analysis-of-time-series-data-using-tensorflow-2136ef633018

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available from: http://www.R-project.org

Romberger, C. Z., and S. Gwozdz. 2018. Performance of water temperature management on the Klamath and Trinity Rivers, 2017. U.S. Fish and Wildlife Service. Arcata Fish and Wildlife Office, Arcata Fisheries Data Series Number DS 2018-59, Arcata, California. Available from: https://www.fws.gov/arcata/fisheries/reports/dataSeries/2017%20Klamath-Trinity%20water%20temp%20reportFINALWith-Appdex.pdf

Salas, J. D. 1993. Analysis and Modeling of Hydrologic Time Series. Pages 19.1–19.72 in D. R. Maidment, editor. Handbook of Hydrology, McGraw-Hill, Inc, New York, NY, USA.

Shibatani, R. 2020. Commentary: the final demise of unimpaired flows. Maven's Notebook, January 2020. Available from: https://mavensnotebook.com/2020/01/28/robert-shibatani-commentary-the-final-demise-of-unimpaired-flows/

Shumway, R. H., and D. S. Stoffer. 2017. Data sets and scripts to accompany time series analysis and its applications: with R examples. Springer Texts in Statistics. Springer, New York, NY, USA.

Stanford, J. A., J. V. Ward, W. J. Liss, C. A. Frissell, R. N. Williams, J. A. Lichatowich, and C. C. Coutant. 1996. A general protocol for restoration of regulated Rivers. United States Department of Energy Publications 43. Available from: https://digitalcommons.unl.edu/usdoepub/43

Steel, E. A., A. Tillotson, D. A. Larsen, A. H. Fullerton, K. P. Denton, and B. R. Beckman. 2012. Beyond the mean: the role of variability in predicting ecological effects of stream temperature on salmon. Ecosphere 3:104.

Stephens, M. A. 1979. Test of fit for the logistic distribution based on the empirical distribution function. Biometrika 66:591–5.

Stergiou, K. I. 1991. Short-term fisheries forecasting comparison of smoothing, ARIMA and regression techniques. Journal of Applied Ichthyology 7:193–204.

Stergiou, K. I., E. D. Christou, and G. Petrakis. 1997. Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods. Fisheries Research 29:55–95.

Sullivan, R. M., and J. P. Hileman. 2018. Annual and seasonal variation, relative abundance, and effects of managed flows on timing of migration in Brown Trout (*Salmo trutta*) in the upper Trinity River. California Fish and Game 104:99–128.

Sullivan, R. M., and J. P. Hileman. 2019. Effects of managed flows on Chinook Salmon in relation to run-timing, average annual percent fertility, and fluctuations in water temperature and flow volume. California Fish and Game 105:132–176.

Sullivan, R. M., and J. P. Hileman. 2020. Comparative analyses of upstream migration in a multispecies assemblage of fish in response to highly managed flow regimes. California Fish and Wildlife 106:99–138.

Trinity River Flow Evaluation Study (TRFES). 1999. Report by the U.S. Fish and Wildlife Service and Hoopa Valley Tribe to the Secretary, U.S. Department of Interior, Washington, D.C. Available from: https://www.fws.gov/arcata/fisheries/reports/technical/Trinity_River_Flow_Evaluation_-_TOC.pdf

Trinity River Restoration Program (TRRP). 2020. Trinity River Restoration Program website. Available from: http://www.trrp.net

Tsiatis, A. A. 2006. Semiparametric theory and missing data. Springer Series in Statistics. Springer, New York, NY, USA.

United States Bureau of Reclamation (USBR). 2000. U.S. Department of the Interior. Record of Decision Trinity River Mainstem Fishery Restoration Final Environmental Impact Statement/Environmental Impact Report, December 2000. Available from: http://odp.trrp.net/library/document/?id=227

Wang, X., K. S. Smith, and R. J. Hyndman. 2006. Characteristic-based clustering for time series data. Data Mining and Knowledge Discovery 13:335–364.

Wissmar, R. C., and P. A. Bisson. 2003. Strategies for restoring river ecosystems: sources of variability and uncertainty in natural and managed Systems. American Fisheries Society, Bethesda, MD, USA.

Woelfle-Erskine, C., A. C. Wilcox, and J. N. Moore. 2012. Combining historical and process perspectives to infer ranges of geomorphic variability and inform river restoration in a wandering gravel-bed river. Earth Surface Processes and Landforms 37:1302–1312.

Wood, S. N. 2017. Generalized additive models: an introduction with R. 2nd edition. Chapman and Hall/CRC, Boca Raton, FL, USA.

Yarnell, S. M., G. E. Petts, J. C. Schmidt, A. A. Whipple, E. E. Beller, C. N. Dahm, P. Goodwin, and J. H. Viers. 2015. Functional flows in modified riverscapes: hydrographs, habitats, and opportunities BioScience 65:963–972.

Zar, J. H. 2010. Biostatistical analysis. Prentice-Hall, Upper Saddle River, NJ, USA.

Zimmerman, J., J. Howard. J. Rosenfield, and R. Henery. 2020. Guest commentary (rebuttal): ecologists see little difference between unimpaired and (truly) functional approaches to flow. Maven's Notebook, February 2020. Available from: https://

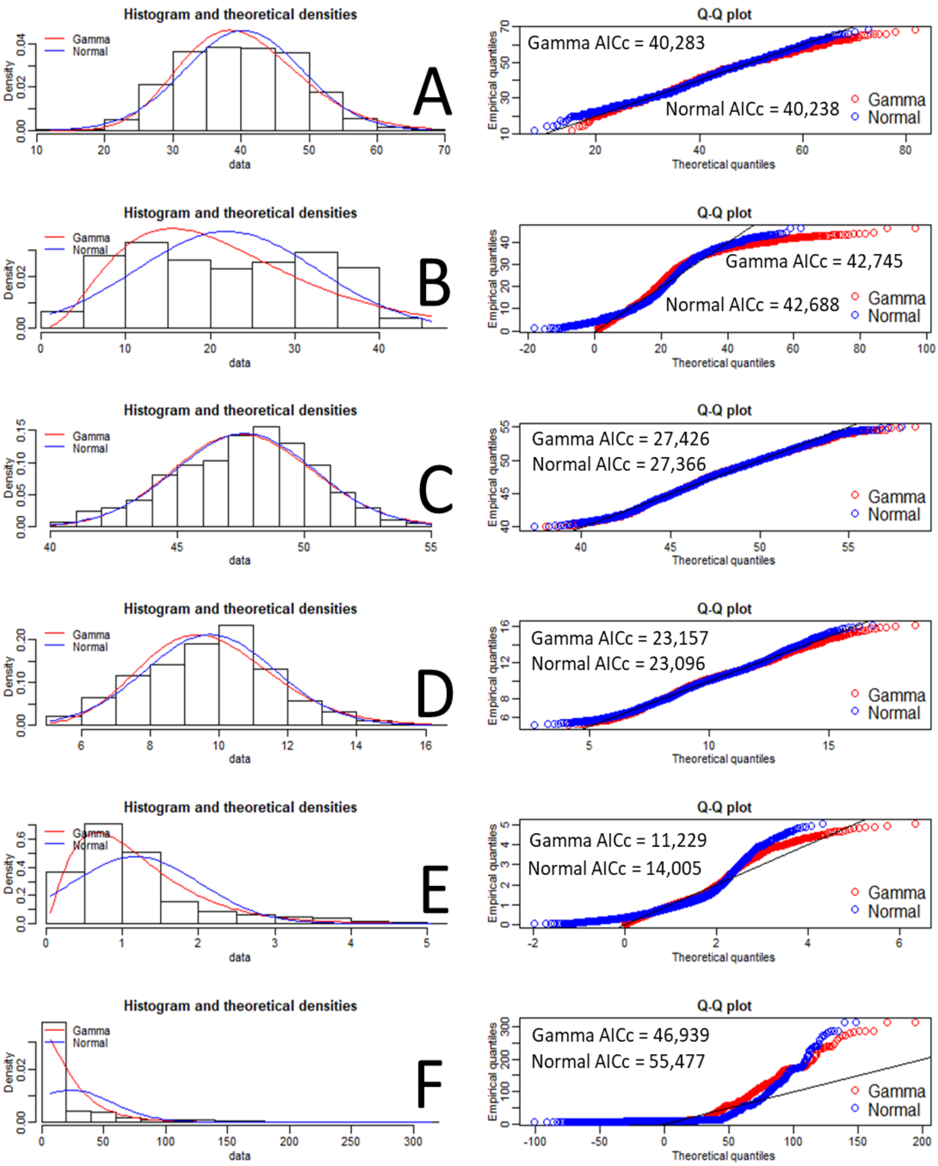
mavensnotebook.com/2020/02/04/guest-commentary-rebuttal-ecologists-see-little-difference-between-unimpaired-and-truly-functional-approaches-to-flow/ and https://ceff.sf.ucdavis.edu/functional-flows-approach

Submitted 4 May 2020
Accepted 22 September 2020
Associate Editors were L. Damon and J. Kelly

APPENDIX I

Year	Water year type	Flow-type	Natural flow at Lewiston Dam (km^3)	TRRP restoration release (km^3)	Lewiston Reservoir management releases (km^3)	Ceremonial releases (km^3)	Other releases (km^3)	Trinity River total release (km^3)	Trinity River peak rate release (m^3/sec)	Diversion to Central Valley (km^3)
2001	Dry	PreROD	1.009	0.468	0.000	0.005	0.000	0.473	60.598	0.825
2002	Normal	PreROD	1.595	0.595	0.000	0.000	0.000	0.595	186.042	0.776
Total			**2.604**	**1.063**	**0.000**	**0.005**	**0.000**	**1.068**	**246.640**	**1.601**
Average			**1.302**	**0.532**	**0.000**	**0.003**	**0.000**	**0.534**	**123.320**	**0.800**
2005	Wet	ROD	1.820	0.798	0.000	0.004	0.000	0.803	216.341	0.575
2006	Ext Wet	ROD	3.078	0.999	0.501	0.000	0.000	1.500	294.495	1.665
2007	Dry	ROD	0.928	0.559	0.000	0.005	0.000	0.564	136.204	0.758
2008	Dry	ROD	1.079	0.800	0.000	0.000	0.000	0.800	195.103	0.684
2009	Dry	ROD	1.029	0.549	0.000	0.014	0.000	0.563	131.107	0.665
2010	Wet	ROD	1.976	0.810	0.000	0.000	0.000	0.810	222.004	0.339
2011	Wet	ROD	2.322	0.890	0.000	0.013	0.000	0.903	348.297	0.583
2017	Ext Wet	ROD	2.872	1.013	0.044	0.011	0.000	1.068	339.802	0.775
2019	Wet	ROD	2.092	0.867	0.000	0.010	0.000	0.877	305.822	0.523
Total			**17.194**	**7.285**	**0.545**	**0.058**	**0.000**	**7.888**	**2189.175**	**6.568**
Average			**1.910**	**0.809**	**0.061**	**0.006**	**0.000**	**0.876**	**243.242**	**0.730**
2003	Wet	Pulse	2.304	0.553	0.084	0.007	0.042	0.686	78.721	1.057
2004	Wet	Pulse	1.864	0.803	0.100	0.000	0.045	0.947	179.812	1.218
2012	Normal	Pulse	1.326	0.798	0.000	0.000	0.048	0.846	174.998	0.875
2013	Dry	Pulse	1.052	0.557	0.000	0.012	0.023	0.592	129.974	1.051
2014	Crit Dry	Pulse	0.489	0.457	0.000	0.000	0.080	0.537	97.976	0.763
2015	Dry	Pulse	1.109	0.556	0.000	0.011	0.059	0.626	250.038	0.555
2016	Wet	Pulse	1.797	0.874	0.000	0.000	0.048	0.922	271.842	0.344
2018	Crit Dry	Pulse	0.674	0.465	0.000	0.000	0.042	0.507	57.766	0.481
Total			**10.615**	**5.062**	**0.184**	**0.031**	**0.387**	**5.664**	**1241.127**	**6.344**
Average			**1.327**	**0.633**	**0.023**	**0.004**	**0.048**	**0.708**	**155.141**	**0.793**

APPENDIX II


Histogram and theoretical densities
Gamma
Normal
Density
data
A
Q-Q plot
Gamma AICc = 40,283
Normal AICc = 40,238
Gamma
Normal
Empirical quantiles
Theoretical quantiles
Histogram and theoretical densities
Gamma
Normal
Density
data
B
Q-Q plot
Gamma AICc = 42,745
Normal AICc = 42,688
Gamma
Normal
Empirical quantiles
Theoretical quantiles
Histogram and theoretical densities
Gamma
Normal
Density
data
C
Q-Q plot
Gamma AICc = 27,426
Normal AICc = 27,366
Gamma
Normal
Empirical quantiles
Theoretical quantiles
Histogram and theoretical densities
Gamma
Normal
Density
data
D
Q-Q plot
Gamma AICc = 23,157
Normal AICc = 23,096
Gamma
Normal
Empirical quantiles
Theoretical quantiles
Histogram and theoretical densities
Gamma
Normal
Density
data
E
Q-Q plot
Gamma AICc = 11,229
Normal AICc = 14,005
Gamma
Normal
Empirical quantiles
Theoretical quantiles
Histogram and theoretical densities
Gamma
Normal
Density
data
F
Q-Q plot
Gamma AICc = 46,939
Normal AICc = 55,477
Gamma
Normal
Empirical quantiles
Theoretical quantiles