Environmental Science & Technology

Article

Cite This: Environ. Sci. Technol. XXXX, XXX, XXX−XXX

pubs.acs.org/est

Modeling Spatial and Temporal Variation in Natural Background Specific Conductivity

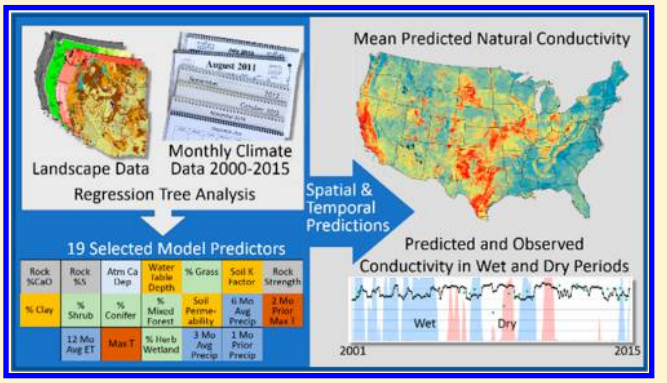
John R. Olson*,† and Susan M. Cormier‡

†California State University Monterey Bay, School of Natural Sciences, 100 Campus Center, Seaside, California 93955, United States

‡U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, 26 Martin Luther King Drive West, Cincinnati, Ohio 45268, United States

Supporting Information

ABSTRACT: Understanding how background levels of dissolved minerals vary in streams temporally and spatially is needed to assess salinization of fresh water, establish reasonable thresholds and restoration goals, and determine vulnerability to extreme climate events like drought. We developed a random forest model that predicts natural background specific conductivity (SC), a measure of total dissolved ions, for all stream segments in the contiguous United States at monthly time steps between the years 2001 to 2015. Models were trained using 11 796 observations made at 1785 minimally impaired stream segments and validated with observations from an additional 92 segments. Static predictors of SC included geology, soils, and vegetation parameters. Temporal predictors were related to climate and enabled the model to make predictions for different dates. The model explained 95% of the variation in SC among validation observations (mean absolute error = 29 μS/cm, Nash-Sutcliffe efficiency = 0.85). The model performed well across the period of interest but exhibited bias in Coastal Plain and Xeric regions (26 and 30%, respectively). National model predictions showed large spatial variation with the greatest SC predicted to occur in the desert southwest and plains. Model predictions also reflected changes at individual streams during drought.



INTRODUCTION

Total dissolved solid (TDS) concentration (measured as specific conductivity [SC] normalized to 25 °C) is an important water quality parameter that affects aquatic ecosystems. Although upper limits vary, water with SC above 1000 μS/cm is unsuitable for many industrial or human uses (e.g., boiler water needs to be <5000 μS/cm)[1] and SC values over 3000 μS/cm are unsuitable for irrigation.[2] Smaller increases in SC increases can also negatively affect aquatic life,[3] including algae,[4−6] invertebrates,[7−9] and vertebrates.[10−12]

Many human activities can increase SC,[13] including agriculture, industry, and resource extraction resulting in the loss of water resources and decreased biological integrity.[3,14] However, it is often difficult to identify where human activities increase SC compared to natural background, because of spatial and temporal variation in SC due to natural factors.[15−17] Spatially, SC naturally varies over 2 orders of magnitude among freshwater systems with variations in geology, soils, climate, and vegetation.[18] SC also responds to temporal changes in precipitation, temperature, and evapotranspiration (ET),[17,19,20] especially to extreme weather as may occur in a prolonged drought.[21−24] Modeling variability in background SC is challenging due to complex interactions among these climatically variable and static, nonclimate factors.[25,26]

A model predicting natural background would be valuable for assessing stream condition and setting restoration goals. The ability to predict natural background SC for individual streams would allow comparisons with current SC and assessment of the degree of change in SC due to human activities. Knowing the magnitude to which SC has been altered will help determine if and to what degree ecological degradation might be caused by increased salinity or changes in dissolved minerals. For example, Vander Laan et al.[27] applied natural background SC models to estimate how much SC had been altered by human activities, and quantitatively linked this alteration to changes in biological conditions. Jones and van Vliet[24] identified water availability and increased salinity as key contributors to water scarcity to drought in the southern United States.

Understanding how SC naturally varies spatially and temporally could help inform where field-based benchmark threshold values (e.g., U.S. EPA[14]) are either underprotective or overly stringent. Background SC models that reflect the effects of different climates on SC may be able to project where

Received: December 3, 2018
Revised: February 12, 2019
Accepted: March 12, 2019
Published: March 12, 2019

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

increased levels of dissolved ions could threaten freshwater biota during periods of drought. Drought can decrease flow and concentrate minerals potentially exacerbating effects from altered temperatures and flows.

Previous models of SC were not designed to predict both spatial and temporal variation in natural SC as is desired for robust and accurate estimates of SC over space and time. For example, Olson and Hawkins[18] modeled natural background SC for the western United States but did not account for temporal variation. Anning and Flynn[28] modeled TDS loads and concentrations for the contiguous United States but did not account for temporal variation. Furthermore, the model developed by Anning and Flynn[28] included both human and natural factors, and its ability to predict natural background has not been assessed. Recently developed spatially and temporally extensive data on climate are now available (i.e., Parameter-elevation Relationships on Independent Slopes Model [PRISM][29] and Moderate Resolution Imaging Spectroradiometer [MODIS] estimates of ET[30]) which allow dynamic spatial and temporal factors influencing SC to be incorporated into empirical models.

Our objective was to develop a statistical model of the natural spatial and temporal variation in SC for the contiguous United States. This model is intended to provide predicted natural background SC for each stream segment defined by the National Hydrography Data set Plus Version 2 (NHD+)[31] at monthly time steps for 2001−2015. Using these predictions, we then examine how SC varies from normal conditions during prolonged droughts.

MATERIALS AND METHODS

General Approach. To develop models that make stream-specific predictions across the contiguous United States, we used the newly developed StreamCat[32] data set and process (https://github.com/USEPA/StreamCat). The StreamCat data set is based on a network of stream segments from NHD+.[31] These stream segments drain an average area of 3.1 km^2 and thus define our spatial grain size. These small drainages are an appropriate scale for modeling because SC varies little at finer spatial scales.

The empirical background conductivity model was developed in several steps.

1. Create training and validation data sets of SC observations from minimally altered stream segments.
2. Characterize temporally and spatially specific watershed environments for each observation, including antecedent conditions.
3. Relate observed SC to environmental predictors using a machine learning technique (random forests [RF]).
4. Assess model performance and validate using multiple observations made at randomly chosen stream segments.

Create Training and Validation Data Sets of SC Observations from Minimally Altered Stream Segments. Developing an empirical model of natural background SC required SC observations from minimally disturbed sites representing the breadth of variation in environmental conditions that occur in an area of interest (see Supporting Information (SI) for details of how these data sets were developed). We first obtained over 2.4 million SC observations from across the continental United States from STORET,[33] state natural resource agencies, the U.S. Geological Survey (USGS) National Water Information System,[34] and data used in Olson and Hawkins[18] (Table S1). Although not an exhaustive collection of SC observation data, these data represent a substantial proportion of what is publicly available. We limited data to observations made between 1 January 2001 and 31 December 2015 so that MODIS satellite data (https://modis.gsfc.nasa.gov/data/) could be used as predictors in our models. Each observation was related to the nearest stream segment in the NHD+. Because our dynamic predictors used a monthly time step, we limited the data to one observation per stream segment per month. SC observations with ambiguous locations and repeat measurements along a stream segment in the same month were discarded. Using estimates of anthropogenic stress derived from the StreamCat database,[32] we selected segments with minimal amounts of human activity[35] as training data for our models. Segments with minimal human activity were selected using criteria developed for each Level II Ecoregion,[36] but in all cases, segments deemed minimally stressed had watersheds with 0−0.5% impervious surface, 0−5% urban, 0−10% agriculture, and population densities from 0.8−30 people/km^2 (Table S3). We also identified observations with large residuals in initial models and inspected these watersheds for evidence of other human activities not represented in StreamCat (e.g., mining, logging, grazing, or oil/gas extraction). Observations from disturbed watersheds were removed, as were observations that were tidally influenced or due to unusual geologic conditions like hot springs, which cause naturally high SC conditions. About 5% of SC observations in each National Rivers and Stream Assessment (NRSA) region were then randomly selected as independent validation data. The remaining observations became the large training data set for model calibration.

The final training data set used for modeling had 1785 stream segments with 11 796 observations, and the validation data set had 92 segments with 581 observations. The majority of segments had a single observation but ranged up to 165 observations per segment (Figure S1A). Reference observations were reasonably dispersed in both time (Figure S1B) and space (Figure S1C), although the Midwest had few reference segments, especially in the Corn Belt in Iowa and Illinois.

Characterize Temporally and Spatially Specific Watershed Environments for Each Observation, Including Antecedent Conditions. We derived 27 static watershed predictors from StreamCat (Table S5). These predictors focused on characterizing the naturally occurring spatial variation in geology, soils, hydrology, vegetation, topography, and atmospheric deposition among watersheds. Although acid deposition has been shown to influence chemical weathering rates and stream alkalinity in the past,[37] deposition has been decreasing and by 2009 variation in acid deposition across the U.S. was minimal.[38] Therefore, we did not include this environmental factor as a potential predictor. Temporal variation was incorporated into our models using watershed averages of four dynamic predictors available at monthly time steps for the period of interest. The four dynamic predictors were monthly average precipitation, average temperature, maximum temperature (from PRISM model[29]) and MODIS-derived evapotranspiration.[30,39] Following the same procedures used to create the StreamCat data set (https://github.com/USEPA/StreamCat), we calculated watershed averages for each NHD+ segment in the contiguous United States for each month during the period of interest (2000−2015). We then extracted the temporally and spatially specific observations of

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

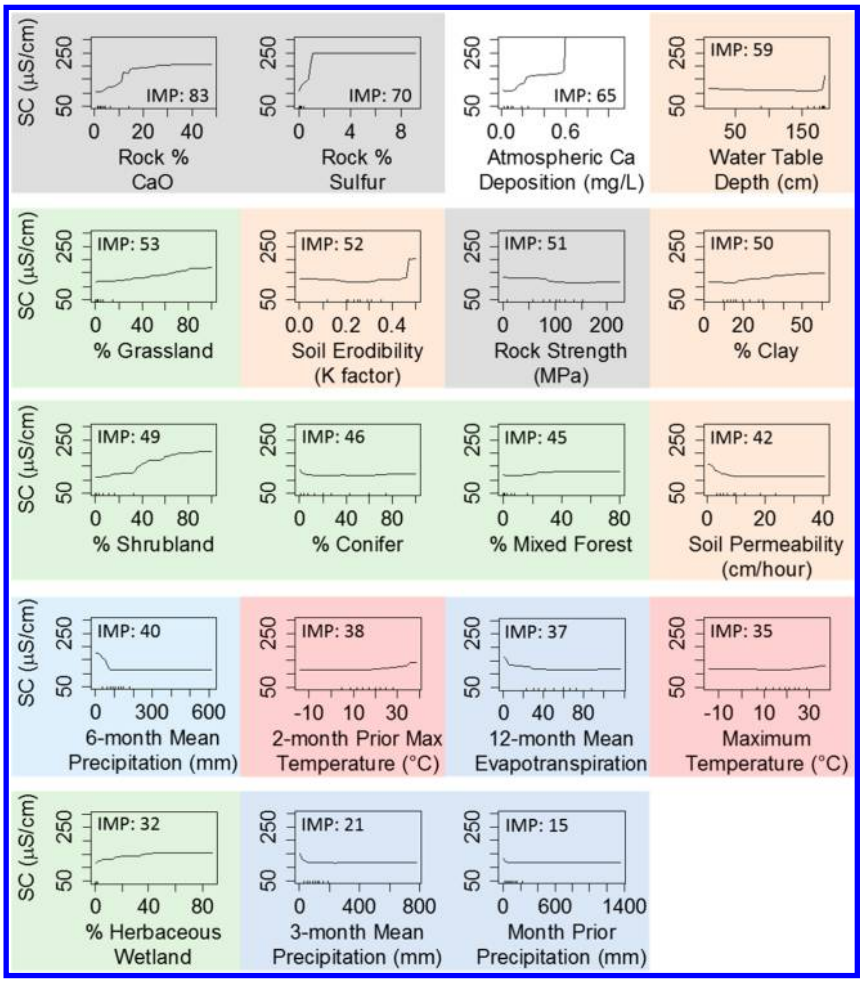


Figure 1. Partial dependence plots and importance (IMP) of selected model predictors. Partial dependence plots show how SC (μS/cm) varies in response to individual predictors while holding all other variables constant. Importance is calculated as the mean increase in error when that predictor is permuted within the model. The higher the value the greater the importance. The steeper the response curve, the more influential the variable is within that specific conductivity (SC) and variable range. Plots are color-coded by parameter type: Geological (gray), Atmospheric (white), Soil (tan), Vegetation (green), Temperature (pink), Evapotranspiration, and Precipitation (blue). Atmospheric Ca Deposition *y*-axis is truncated at 300 μS/cm to allow comparison with other predictors. SC response to Atmospheric Ca deposition plateaus at 550 μS/cm above 0.6 mg/L.

each of the four dynamic predictors (extracted precipitation, mean temperature, maximum temperature, and mean ET) that matched the time (month and year) and location (NHD+ segment) of each SC observation. In addition, we characterized conditions antecedent to each observation using estimates of each dynamic predictor from the month prior, 2 months prior, and averages of the preceding 3, 6, and 12 months. For example, an SC observation made in December 2005 was matched with watershed precipitation, temperature, and mean ET observed that month, the previous month (e.g., November), 2 months prior (e.g., October), and the averages over the previous 3 months (e.g., October–December), 6 months (e.g., July–December) and 12 months (e.g., January–December). In this way, preceding conditions were considered as well as near-term events.

Develop Random Forest Models to Relate Observed SC to Predictors. We developed RF models[40] to predict natural background stream SC. RF is a nonparametric regression and classification modeling approach that has been applied to a wide array of disciplines, including genetics, ecology, and remote sensing.[41] RF models have significant advantages over other statistical methods, including their ability to fit nonlinear relationships and high-order interactions between predictor variables without a priori specification of the shape of relationships or the presence of interactions. RF models combine predictions from numerous regression or classification trees based on bootstrapped samples of predictor and response data to produce robust models resistant to overfitting. Data not included in individual regression trees (i.e., out-of-bag training observations) are used to assess model accuracy and precision, similar to cross-validation. Models were built with the "randomForests" package in R[42] using all default settings except that we built 1500 trees and applied a bias correction feature.[43]

We selected predictors using a principal component analysis (PCA) approach that identifies uncorrelated predictors[44] with the strongest associations with SC (following a method suggested by R. A. Hill and E. W. Fox, personal communication). A PCA was constructed using centered and rescaled predictors, the number of axes needed to explain 95% of the variation was determined, and then Varimax rotation was performed on those axes. For each rotated axis, we determined which predictor had the greatest loadings and which predictor had the greatest univariate association with

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX–XXX

SC. Univariate associations were determined by fitting a classification and regression tree between SC and each predictor loading on a given axis and extracting the deviance. For axes where the greatest loading predictor was different from the predictor with the greatest association with SC, we chose the more interpretable parameter of the two. For each potential predictor, we examined the partial dependence plots showing how SC responds to that predictor while holding all other potential predictors constant. Predictors that had inconsistent or otherwise uninterpretable responses were removed. Fox et al.[45] found that variable selection only improved model performance when the majority of predictors were irrelevant. However, a parsimonious model is desirable that limits the number of required predictor variables. Therefore, the importance of different predictors was assessed as the reduction in mean square error occurring when the variable is permuted. Mathematical permutation reorders the sequence of introduction of a predictor into the regression trees while holding the other predictors constant. Individual response-predictor relations were visualized with partial dependence plots.

RF models do not distinguish between spatial and temporal variation, and do not account for temporal patterns in the data. RF models make individual predictions based only on the values of predictors associated with an observation, so spatial variation in the environment is reflected by using averages for each environmental factor for the entire upstream watershed. To account for temporal variation, we used temporally specific and antecedent observations of the four dynamic predictors (i.e., climate observations in the same month as the SC observation; Table S5) as potential predictors in our models.

Assess Model Performance and Validate Using Multiple Observations Made at Randomly Chosen Stream Segments. We assessed model performance by comparing model SC predictions for out-of-the bag observations from the training data and the external validation data to actual observations. Predictions for out-of-the bag observations were made by averaging predictions for all trees that did not use that particular observation in the creation of the tree.[46]

Measurements of model fit were calculated using the R package hydro-GOF[47] and summarized using the following four measurements of goodness-of-fit. (1) The mean absolute error (MAE) is a measure of difference between two variables that allows comparisons of predicted versus observed SC. The MAE is similar to the root-mean-square error (RMSE), except the MAE calculation does not square the errors, making interpretation of the MAE more straightforward (because it is in the same units as the model) and the statistic less sensitive to outliers.[48] Like the RMSE, the smaller the MAE value the greater the confidence in model predictions. (2) The Nash-Sutcliffe efficiency (NSE) estimates the correspondence between predicted and observed data.[49,50] An efficiency of 1 indicates equality between the predicted and observed data. (3) A coefficient of determination (R^2) describes the proportion of the variance in the observations explained by the model. R^2 ranges from 0 to 1, with higher values indicating greater explanatory power and less error. (4) Percent bias is low when over and under predictions occur randomly around the regression model.

Three sites were also hand-picked for validation with a larger than average temporal coverage from three areas affected by severe droughts during the time period covered by the model. We graphically assessed the ability of the model to predict the temporal patterns of SC at these three sites.

RESULTS AND DISCUSSION

Empirical Conductivity Model. Nineteen predictors were included in the final model, representing influences of geology, climate, soils, and vegetation on SC (Figure 1). Geology had the greatest effect on variation in SC, with SC being specifically influenced by variation in calcium and sulfur rock content (first- and second-most important predictors) as well as rock strength, which reflects resistance to physical weathering (7th most important predictor). Atmospheric deposition of calcium was also a strong predictor (3rd most important predictor), indicating its importance as a source of solutes in certain circumstances. Several vegetation types (grasses, shrubs, and mixed forests) and soils properties (water table depth, erodibility, and percent clay) were positively related to SC. Precipitation-related dynamic predictors were all negatively related to SC as expected due to dilution, but of lower importance than other factors due to spreading the signal over three separate measures of precipitation. Precipitation in the proceeding 1, 3, and 6 months was related to SC. Increasing maximum temperatures at month of measurement and 2 months prior were positively related to SC, likely reflecting the combined effects of evapo-concentration and increased weathering rates. There is general agreement of the importance of predictors used in this spatial/temporal model to those in a similar model that only used spatial predictors.[51] However, the long-term averages of temperature and precipitation were more important as predictors than the temporally specific versions used in the current model. Our variable selection process also indicated that runoff and watershed area were potentially important predictors. We chose not to include runoff because it reflected only spatial variation and was correlated with our temporal estimates of precipitation. Although, watershed area was not included in the final model because it accounted for little variation and had an inconsistent relationship to SC; however, spatial variation in the environment is reflected by using averages for each environmental factor for the entire upstream watershed.

Model Performance and Validation. The model explained most of the variation in SC and produced reasonably accurate predictions for both training data (assessed with out-of-bag predictions, MAE = 22 μS/cm, NSE = 0.92, and R^2 = 0.92) and external validation data (MAE = 29 μS/cm, NSE = 0.87, and R^2 = 0.87; Figure 2).

The model had 0 bias when applied to out-of-bag data from the training data set and 1% bias when applied to the external

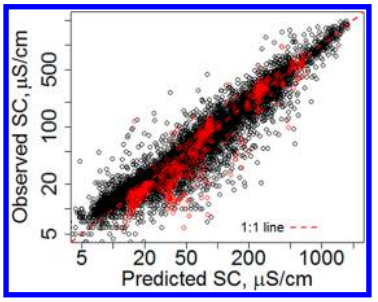


Figure 2. Plots of $\log_{10}$ observed specific conductivity (SC) vs $\log_{10}$ predicted values for out-of-bag training observations (black circles) and external validation data (red circles).

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Model Performance by Month[a]

		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
n (obs)	TNG	660	536	1005	822	1257	1161	1466	1507	1207	815	878	482
	VAL	31	13	53	48	77	49	78	82	64	32	47	7
MAE	TNG	17	21	17	24	25	25	24	25	23	26	15	20
	VAL	18	31	14	26	40	63	33	24	23	38	13	20
NSE	TNG	0.95	0.93	0.94	0.93	0.88	0.93	0.92	0.88	0.92	0.92	0.95	0.95
	VAL	0.97	0.96	0.97	0.95	0.75	0.76	0.80	0.90	0.94	0.90	0.94	0.90
R^2	TNG	0.95	0.94	0.94	0.93	0.88	0.93	0.93	0.88	0.92	0.92	0.95	0.95
	VAL	0.97	0.98	0.97	0.95	0.76	0.76	0.82	0.90	0.94	0.90	0.97	0.99
percent bias	TNG	0.1	−0.2	0.7	0.6	1	0.3	2.4	0.1	−3.7	−0.7	−2.4	0.6
	VAL	−2.8	−0.8	−0.2	2.2	6.4	−0.3	9.6	7.0	0.5	−0.5	9.8	−15.2

[a]MAE, mean absolute error; NSE, Nash-Sutcliffe efficiency; R^2, coefficient of determination; TNG, out-of-the-bag training observations; and VAL, external validation observations.

Table 2. Model Performance by NRSA Regions[a]

		CPL	NAP	NPL	SAP	SPL	TPL	UMW	WMT	XER
n (segments)	TNG	315	228	49	354	56	8	48	438	289
	VAL	18	12	3	17	3	1	3	22	13
n (obs)	TNG	2295	588	241	3459	314	32	542	3198	1127
	VAL	145	36	13	172	21	81	8	75	30
MAE	TNG	17	15	60	9	87	31	33	16	62
	VAL	18	9	67	8	97	36	41	42	117
NSE	TNG	0.88	0.64	0.63	0.84	0.81	0.40	0.78	0.85	0.92
	VAL	0.66	0.27	0.68	0.79	−0.53	−0.46	0.62	0.75	0.02
R^2	TNG	0.88	0.64	0.63	0.85	0.81	0.41	0.78	0.85	0.92
	VAL	0.78	0.52	0.80	0.81	0.11	0.11	0.70	0.76	0.32
percent bias	TNG	0.1	−1.1	−1.4	−6.6	−0.3	1.1	0.7	0.9	1.4
	VAL	32.0	19.1	−7.8	6.5	15.4	−9.7	5.7	−8.6	32.7

[a]CPL, Coastal Plain; MAE, mean absolute error; NSE, Nash-Sutcliffe efficiency; NAP, Northern Appalachia; NPL, Northern Plains; NRSA, National Rivers and Stream Assessment; R^2, coefficient of determination; SAP, Southern Appalachia; SPL, Southern Plains; TNG, out-of-the-bag training observations; TPL, Temperate Plains; UMW, = Upper Midwest; WMT, Western Mountains; VAL, external validation observations; and XER, Xeric.

validation data. Model performance remained constant across months, except for a small decrease in performance in May–July (R^2 range = 0.76–0.82) in the validation data (Table 1). However, predictions of validation data measured in December were negatively biased by 15%. This may be due to the influence of a single poor prediction among a relatively small number of validation samples collected in December (n = 7).

Model performance within individual regions was comparable to model performance across the contiguous United States, except for decreased performance in the Southern Plains (SPL) and Temperate Plains (TPL) regions (Table 2). We evaluated variation in model performance spatially using aggregated National Rivers and Stream Assessment ecoregions (following ref 52). We aggregated level II ecoregions to approximate the regions used in the National Rivers and Stream Assessment because level II ecoregions often had too few sites to reliably estimate model performance. Performance assessed with out-of-bag training observations showed some variance in performance across regions, with 4-fold increase in MAE in the SPL (MAE = 87 μS/cm) and a drop in both NSE and R^2 by over half in the TPL (NSE = 0.40, R^2 = 0.41). Bias in out-of-bag predictions was <2% for in all regions except Southern Appalachia.

Model performance among regions assessed with external validation data showed greater variability in performance than regional model performance assessed with out-of-bag training observations, perhaps a result of the smaller sample sizes from external validation data. MAE for external validation data did not differ much from that calculated for the training data, except in Xeric (XER) regions where MAE increased from 62 to 117 μS/cm. The NSE and R^2 of both the SPL and TPL indicate that the model performed very poorly when applied to the external validation data in these areas. Most of the external validation observations used in these two regions were from a single site in each region (15 of 21 observations in SPL, all 81 observations in TPL). The plots of validation versus model predictions for these regions (Figure S2) show these sites had greater temporal variability than predicted by the model, which may have been caused by greater environmental heterogeneity within these watersheds resulting in greater temporal variability than expected.[53] The Coastal Plain (CPL) and Xeric regions both showed high amounts of bias in predictions of validation observation (32 and 33%, respectively). The high bias in both cases was caused by outliers and removing two sites from each validation set improved percent bias to 13.6% CPL and 10.8% XER. Removing the two outlier sites also improved the NSE of the Xeric region validation to 0.34 and R^2 to 0.43.

Spatial and Temporal Patterns of SC. The desert southwest, southern, and northern plains, and parts of southern California exhibited the greatest mean SC, likely caused by the calcareous, evaporitic, and marine geologies interacting with high ET and low dilution from precipitation in these areas (Figure 3A). Spatial patterns of mean SC in summer and winter showed the same patterns as the annual mean (data not shown). Streams in the southern and northern plains, midwest, and most of California had the greatest amount of temporal

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX–XXX

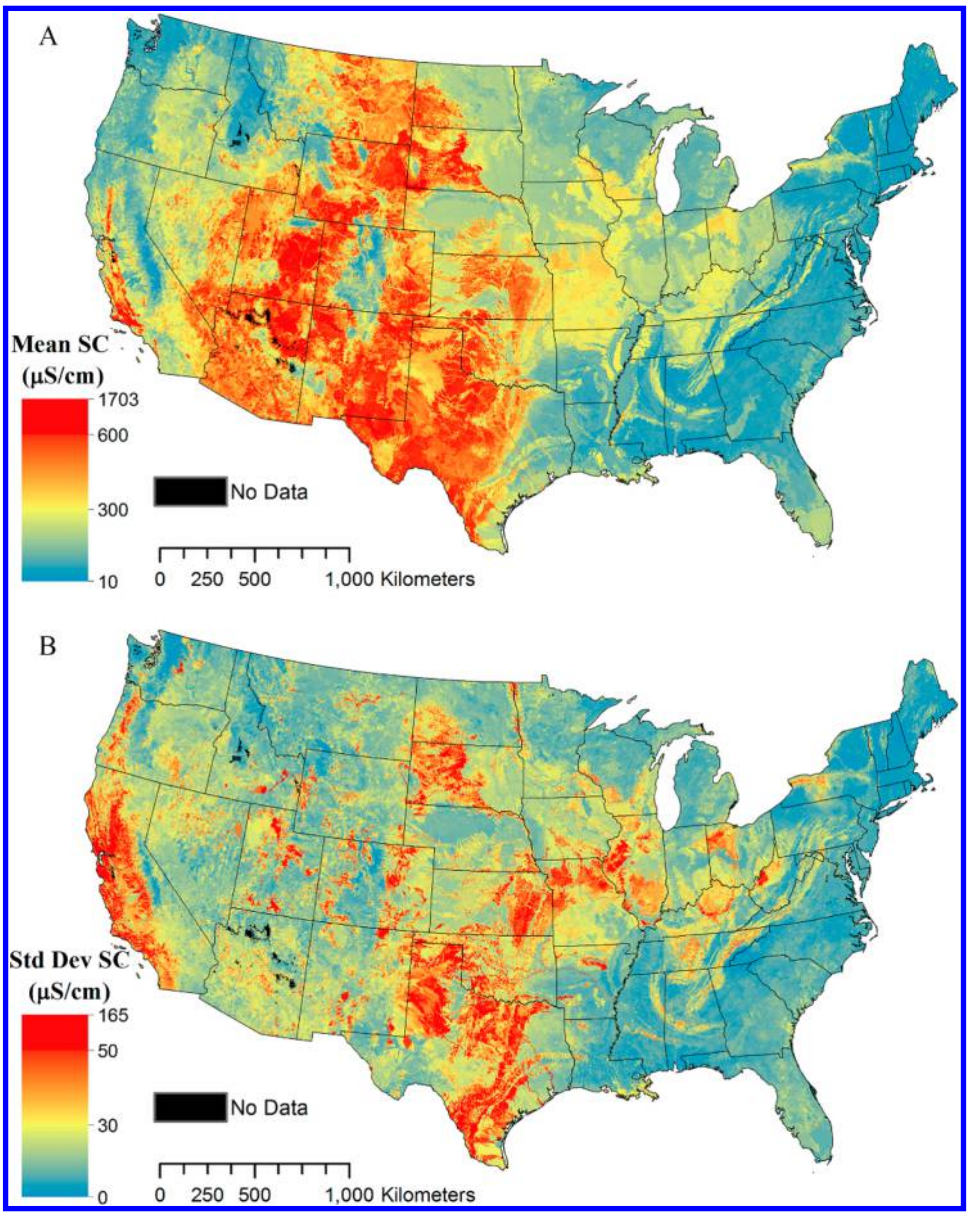


Figure 3. Maps of (A) the predicted mean monthly specific conductivity (SC) for streams in the contiguous United States between 2001−2015, and (B) the standard deviation of predicted SC across the same time period. Note different SC scales. SC, specific conductivity; and Std Dev, standard deviation.

variation measured as standard deviation across the time period (Figure 3B).

We compared temporal predictions to observed SC at three sites chosen because the SC measurements were made during drought and nondrought times (Figure 4). Although predictions at each site showed similar temporal patterns with observed SC (e.g., decreasing during wet years and seasons), there were periods in which the predictions did not agree with the observed data. For example, in the Sisquoc River, CA, almost monthly monitoring indicated a wider range of changes in SC than the SC predicted by the model (Figure 4A). Model predictions at this site were also consistently 25% lower than observed. In periods leading up to the drought (2011−2012), predicted SC is lower than but generally parallel to the observed SC. At the end of the drought (2015), predicted SC was 16% below observed SC adjusted for the underprediction. Model predictions of SC appear to underestimate climatic effects during prolonged and severe drought (43 months of flows averaging 5% of the 10-year average) and where inputs from upstream reaches might be variable due to intermittent flows. In the Sisquoc River, CA, the departure of observed SC from predicted may have been due to the river drying just above the measurement point, so most or all flow measured at the end of the drought was likely from deep groundwater. SC in the closest well (USGS 345034120131301, c. Three miles away) was 1240 μS/cm. The long contact time of deep groundwater results in increased weathering.[54] Flow completely dominated by deep groundwater is an uncommon situation not well represented by the model.

In Hondo Creek, TX, model predictions were comparable to empirical measurements except during the 2009 drought (Figure 4B), where observed SC during this drought was approximately 200 μS/cm below both the predicted and the long-term average. Empirical SC measurements may have been lower than modeled SC estimates due to water being added to the system from groundwater withdrawal. Hondo Creek draws

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

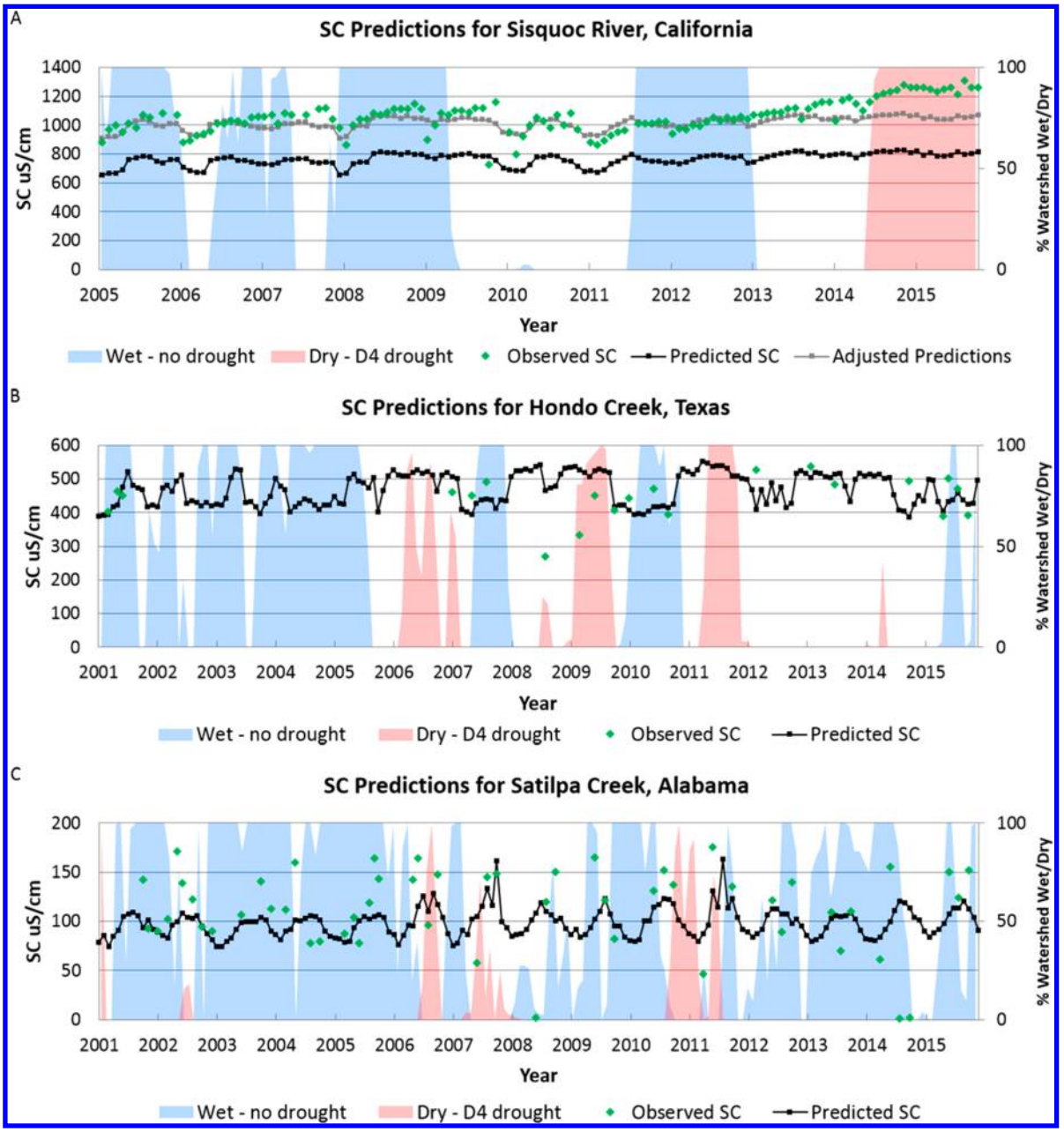


Figure 4. Observed and predicted changes in specific conductivity (SC) over time at 3 stream segments during droughts. (A) Sisquoc River, CA (Unique ID: 17625379); (B) Hondo Creek, TX (10654651); and (C) Satilpa Creek, AL (21640642). For each graph, the black squares indicate the predicted SC, the green diamonds represent observed SC, the percentage of the hydrologic unit code 8 not in drought (i.e., wet periods) are indicated by blue areas and the percentage in extreme drought (dry periods, i.e., areas classified D4 (Exceptional Drought) by the U.S. Drought Monitor) by red areas. Note that SC scales differ for each plot and illustrate how well the model estimates the SC range for each stream segment. Adjusted SC predictions are included for the Sisquoc River to account for the underprediction of SC for that site.

groundwater from the Edwards aquifer with SC < 460 μS/cm, which is similar and at times less than the average stream SC (USGS https://data.edwardsaquifer.org/documents/2006_Green-etal_KinneyUvaldeEvaluation.pdf).

Predictions from Satilpa Creek, AL were primarily between 70 and 120 μS/cm, which generally followed a seasonal pattern. During droughts in 2006, 2007, and 2010−2011, sharp peaks predicted increased SC during drought years (Figure 4C). However, observed SC was more variable than average predictions during both drought and nondrought periods, suggesting there may have been other sources of salts not detected by our screening for anthropogenic effects. Also, Satilpa Creek has a relatively low SC regime and may reflect the temporal discriminatory precision of the model in low SC streams during drought.

We also used our model to identify and examine areas predicted to have the largest potential SC increases during the extended drought in California between 2012−2017 (Figure 5). We compared the predicted natural background SC expected in July during a wet year (2005) to the SC predicted in 2015 during the drought in California. Some parts of California were predicted to have >125 μS/cm SC increases during the drought. Areas susceptible to increases are those that depend on dilution from snowmelt compared to Xeric

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

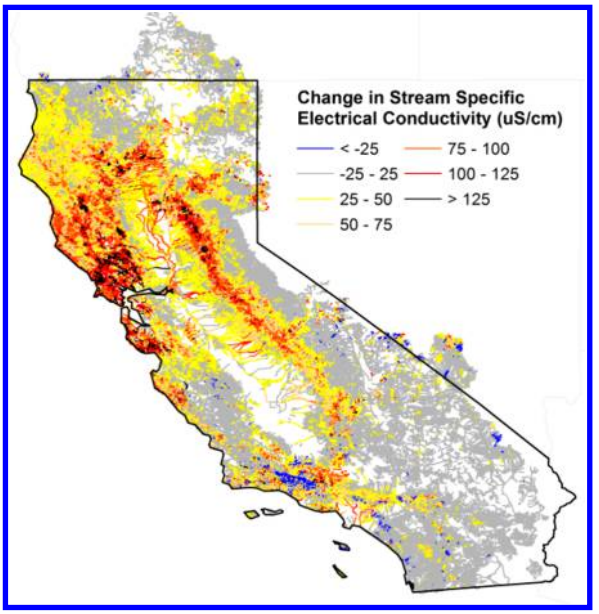


Figure 5. Difference in predicted specific conductivity between a wet year (July 2005) and a drought year (July 2015) assuming streams are unaffected by anthropogenic inputs.

areas that were predicted to have little or no change in SC during drought. The model only predicts the effects of drought on natural background SC, but streams with SC regimes altered by anthropogenic activity may experience compounding increases in SC than is predicted for minimally disturbed streams given minimal dilution and continuing discharge inputs, which commonly contain elevated ion concentrations.

The empirical model presented here provides both temporal and spatial estimates of the natural background SC of streams in the contiguous United States. At the national level, the model was strong (R^2 = 0.92), validated (R^2 = 0.95), and mirrored trends of observed stream SC over the long-term (Figure 2). This model improves on a previous model published by Olson and Hawkins (2012) in a few ways. First, this new model was developed to improve predictions in the natural background range of SC for most freshwaters (i.e., < 1000 μS/cm[15]). Second, by including precipitation variables with different time lags, the model is now capable of predicting SC over time when precipitation data are available. Lastly, although this model is empirical, it is coherent with factors expected to mechanistically influence the availability and mobility of ion delivery to stream networks.[55,56] Therefore, on this large scale, this empirical model output along with real world data may help to improve mechanistic understanding of geophysical processes influencing stream SC.

Although the predictions of natural background SC made by this model have many uses, several limitations should be considered. First, because the model relies on the NHD+ and StreamCat data sets, streams without this data (i.e., some headwater systems and buried urban streams) will not have predictions. Second, predictions in places where SC is driven by factors not included in the model will be inaccurate. Examples include coastal areas influenced by tidal salinity and salt water intrusion or geothermally active areas. Predictions in areas with few minimally disturbed streams, especially the Temperate Plains represented by only eight streams segments, will be less accurate than the model as a whole. Because the model relies on current vegetation as a predictor (Figure 1), predictions for streams where humans have shifted vegetation between forests and either grasses or shrubs over significant portions of a watershed may not be truly representative of natural background conditions.

Despite these limitations, the model output is useful for depicting the patterns of natural background SC (Figures 3 and 4). The model is also useful for identifying potential sources of pristine fresh waters and informing the investigations regarding the vulnerability of pristine freshwaters to rainfall variability. For example, the comparison of wet and dry years in California (Figure 5) showed areas with naturally lower SC were more variable than areas with higher SC, suggesting dilution by precipitation was a key factor in seasonal changes in SC. Although we did not attempt to make predictions for scenarios during an ongoing drought or estimate the amount of rainfall required to achieve drought relief thereby lowering stream SC, the model parameters for precipitation could be varied to estimate the dilution needed to reduce SC to a desired level. To enable the widest possible use of the underlying data, model, and model outputs are all made available on the U.S. Environmental Protection Agency (EPA) Environmental Data set Gateway (https://edg.epa.gov/metadata/catalog/main/home.page).

CONCLUSIONS

The development of a national model for predicting stream SC was possible because calibration data were available through state and federal sampling efforts in addition to large digital data sets for geology, climate, and remotely sensed vegetative cover. The modeled results may be compared with measured data to quantify changes in stream SC for individual reaches or for larger regions, particularly in areas affected by anthropogenic disturbances. For example, many state and federal agencies and stakeholder groups have empirical SC data from sites with anthropogenic inputs, which may be compared to the modeled natural background SC in those same areas. This information may be used to estimate the proportion and magnitude of stream salinization in the contiguous United States. The differential between predicted background and observed SC may also be used to estimate extirpation of aquatic life.[57] Model results and additional analyses have the potential to enable planning and management of freshwater conditions in catchments from small streams to large river basins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.8b06777.

Detailed description of how the data used in modeling was selected, supplemental tables and figures, evaluated and selected predictors for random forest model, spatial and temporal distribution of observations used to train the model, and plots of observed specific conductivity versus predicted values for external validation observations (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: joolson@csumb.edu.

ORCID

John R. Olson: 0000-0003-1456-6681

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by and prepared at the U.S. EPA, National Center for Environmental Assessment, Cincinnati Division and Office of Water, Health, and Ecological Criteria Division, Washington, DC. The authors are indebted to the work of field and laboratory personnel that generated the primary data, Lei Zheng and Ann Roseberry-Lincoln for data set construction, Benjamin Jessup of TetraTech for project management, and the spatial analyses of Amber Stephens and Megan Rodenbeck. The manuscript has been subjected to U.S. EPA's peer and administrative review and approved for publication. However, the views expressed are those of the authors and do not necessarily represent the views or policies of the U.S. EPA. Michael Gallagher edited and formatted the document. Constructive comments from James Justice and Thomas Hollenhorst and from anonymous reviewers helped to substantially improve an earlier version of this manuscript. The authors declare no competing financial interest.

■ REFERENCES

(1) American Boiler Manufacturers Association (ABMA). Boiler Water Quality Requirements and Associated Steam Quality for Industrial/Commercial and Institutional Boilers. In *ABMA-Boiler 402*; American Boiler Manufacturers Association: Vienna, VA, 2005.

(2) Ayers, R. S., Westcot, D. W. *Water Quality for Agriculture, FAO Irrigation and Drainage Paper 29 Rev. 1*; Food and Agricultural Organization of the United Nations: Rome, 1985. ISBN 92-5-102263-1.

(3) Cañedo-Argüelles, M.; Hawkins, C. P.; Kefford, B. J.; Schäfer, R. B.; Dyack, B. J.; Brucet, S.; Buchwalter, D.; Dunlop, J.; Frör, O.; Lazorchak, J.; Coring, E.; Fernandez, H. R.; Goodfellow, W.; Achem, A. L. G.; Hatfield-Dodds, S.; Karimov, B. K.; Mensah, P.; Olson, J. R.; Piscart, C.; Prat, N.; Ponsá, S.; Schulz, C.-J.; Timpano, A. J. Saving freshwater from salts. *Science* **2016**, *351* (6276), 914−916.

(4) Leland, H. V.; Brown, L. R.; Mueller, D. K. Distribution of algae in the San Joaquin River, California, in relation to nutrient supply, salinity and other environmental factors. *Freshwater Biol.* **2001**, *46* (9), 1139−1167.

(5) Potapova, M. *Relationships of Soft-Bodied Algae to Water-Quality and Habitat Characteristics in the U.S. Rivers: Analysis of the National Water-Quality Assessment (NAWQA) Program Data Set*; The Academy of Natural Sciences: Philadelphia, PA, 2005. http://diatom.acnatsci.org/autecology/uploads/Report_October20.pdf.

(6) Potapova, M.; Charles, D. F. Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition. *Freshwater Biol.* **2003**, *48* (8), 1311−1328.

(7) Boehme, E. A.; Zipper, C. E.; Schoenholtz, S. H.; Soucek, D. J.; Timpano, A. J. Temporal dynamics of benthic macroinvertebrate communities and their response to elevated specific conductance in Appalachian coalfield headwater streams. *Ecol. Indic.* **2016**, *64*, 171−180.

(8) Clements, W. H.; Kotalik, C. Effects of major ions on natural benthic communities: an experimental assessment of the US Environmental Protection Agency aquatic life benchmark for conductivity. *Freshw. Sci.* **2016**, *35* (1), 126−138.

(9) Cormier, S. M.; Suter, G.W. II; Zheng, L.; Pond, G. J. Assessing causation of the extirpation of stream macroinvertebrates by a mixture of ions. *Environ. Toxicol. Chem.* **2013**, *32* (2), 277−287.

(10) Van Meter, R. J.; Swan, C. M.; Snodgrass, J. W. Salinization alters ecosystem structure in urban stormwater detention ponds. *Urban Ecosyst* **2011**, *14* (4), 723−736.

(11) Cheek, C. A.; Taylor, C. M. Salinity and geomorphology drive long-term changes to local and regional fish assemblage attributes in the lower Pecos River, Texas. *Ecol. Freshw. Fish.* **2016**, *25* (3), 340−351.

(12) Griffith, M. B.; Zheng, L.; Cormier, S. M. Using extirpation to evaluate ionic tolerance of freshwater fish. *Environ. Toxicol. Chem.* **2018**, *37* (3), 871−883.

(13) Kaushal, S. S.; Likens, G. E.; Pace, M. L.; Utz, R. M.; Haq, S.; Gorman, J.; Grese, M. Freshwater salinization syndrome on a continental scale. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E574−E583.

(14) U.S. Environmental Protection Agency (U.S.EPA). Public Review Draft Field-Based Methods for Developing Aquatic Life Criteria for Specific Conductivity; *EPA-822-R-07−010*; U.S. Environmental Protection Agency, Office of Water: Washington, DC, 2016. https://www.epa.gov/wqc/draft-field-based-methods-developing-aquatic-life-criteria-specific-conductivity-documents.

(15) Griffith, M. B. Natural variation and current reference for specific conductivity and major ions in wadeable streams of the conterminous USA. *Freshw. Sci.* **2014**, *33* (1), 1−17.

(16) Hem, J. D. Study and Interpretation of the Chemical Characteristics of Natural Water. *Water Supply Paper 2254*; Department of the Interior, U.S. Geological Survey: Alexandria, VA, 1985. https://pubs.usgs.gov/wsp/wsp2254/.

(17) Timpano, A. J.; Zipper, C. E.; Soucek, D. J.; Schoenholtz, S. H. Seasonal pattern of anthropogenic salinization in temperate forested headwater streams. *Water Res.* **2018**, *133*, 8−18.

(18) Olson, J. R.; Hawkins, C. P. Predicting natural base-flow stream water chemistry in the western United States. *Water Resour. Res.* **2012**, *48* (2), W02504.

(19) Interlandi, S. J.; Crockett, C. S. Recent water quality trends in the Schuylkill River, Pennsylvania, USA: a preliminary assessment of the relative influences of climate, river discharge and suburban development. *Water Res.* **2003**, *37* (8), 1737−1748.

(20) Smol, J. P.; Douglas, M. S. Crossing the final ecological threshold in high Arctic ponds. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (30), 12395−12397.

(21) Bohnert, H. J.; Sheveleva, E. Plant stress adaptations-making metabolism move. *Curr. Opin. Plant Biol.* **1998**, *1* (3), 267−274.

(22) Caruso, B. S. Temporal and spatial patterns of extreme low flows and effects on stream ecosystems in Otago, New Zealand. *J. Hydrol.* **2002**, *257* (1−4), 115−133.

(23) Mosley, L. M. Drought impacts on the water quality of freshwater systems; review and integration. *Earth-Sci. Rev.* **2015**, *140*, 203−214.

(24) Jones, E.; van Vliet, M. T. H. Drought impacts on river salinity in the southern US: Implications for water scarcity. *Sci. Total Environ.* **2018**, *644*, 844−853.

(25) Kundzewicz, Z. W.; Krysanova, V. Climate change and stream water quality in the multi-factor context. *Clim. Change* **2010**, *103* (3), 353−362.

(26) Hellwig, J.; Stahl, K.; Lange, J. Patterns in the linkage of water quantity and quality during low-flows. *Hydrol. Process* **2017**, *31* (23), 4195−4205.

(27) Vander Laan, J. J.; Hawkins, C. P.; Olson, J. R.; Hill, R. A. Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. *Freshw. Sci.* **2013**, *32* (3), 801−820.

(28) Anning, D. W., Flynn, M. E. *Dissolved-Solids Sources, Loads, Yields, and Concentrations in Streams of the Conterminous United States; Scientific Investigations Report 2014−5012*; U.S. Geological Survey: Reston, VA, 2014 DOI: 10.3133/sir20145012.

(29) Daly, C.; Halbleib, M.; Smith, J. I.; Gibson, W. P.; Doggett, M. K.; Taylor, G. H.; Curtis, J.; Pasteris, P. P. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.* **2008**, *28* (15), 2031−2064.

(30) Mu, Q.; Zhao, M.; Running, S. W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **2011**, *115* (8), 1781−1800.

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(31) McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., Rea, A. *NHDPlus Version 2: User Guide*; National Operational Hydrologic Remote Sensing Center: Washington, DC, 2012, ftp://ftp.horizon-systems.com/NHDplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf.
(32) Hill, R. A.; Weber, M. H.; Leibowitz, S. G.; Olsen, A. R.; Thornbrugh, D. J. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. *J. Am. Water Resour. Assoc.* **2016**, *52* (1), 120−128.
(33) U.S. Environmental Protection Agency (U.S. EPA). STORET; http://www.epa.gov/storet/. Accessed July 2016.
(34) U.S. Geological Survey (USGS). National Water Information System; http://waterdata.usgs.gov/nwis. Accessed June 2016.
(35) Stoddard, J. L.; Larsen, D. P.; Hawkins, C. P.; Johnson, R. K.; Norris, R. H. Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecol. Appl.* **2006**, *16* (4), 1267−1276.
(36) Omernik, J. M.; Griffith, G. E. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environ. Manage.* **2014**, *54* (6), 1249−1266.
(37) Kaushal, S. S.; Likens, G. E.; Utz, R. M.; Pace, M. L.; Grese, M.; Yepsen, M. Increased river alkalinization in the Eastern US. *Environ. Sci. Technol.* **2013**, *47* (18), 10302−10311.
(38) Waller, K.; Driscoll, C.; Lynch, J.; Newcomb, D.; Roy, K. Long-term recovery of lakes in the Adirondack region of New York to decreases in acidic deposition. *Atmos. Environ.* **2012**, *46*, 56−64.
(39) Reitz, M.; Senay, G. B.; Sanford, W. E. Combined remote sensing and water-balance evapotranspiration estimates (SSEBop-WB) for the conterminous United States. *U.S. Geological Survey Data Release.* **2017**, *9*, 1181.
(40) Breiman, L. Random forests. *Mach. Learn.* **2001**, *45* (1), 5−32.
(41) Cutler, D. R.; Edwards, T. C., Jr.; Beard, K. H.; Cutler, A.; Hess, K. T.; Gibson, J.; Lawler, J. J. Random forests for classification in ecology. *Ecology* **2007**, *88* (11), 2783−2792.
(42) Liaw, A.; Wiener, M.Classification and regression by randomForest.*R News*2002, *2* (3), 18−22https://www.r-project.org/doc/Rnews/Rnews_2002-3.pdf.
(43) Zhang, G.; Lu, Y. Bias-corrected random forests in regression. *J. Appl. Stat.* **2012**, *39* (1), 151−160.
(44) Jolliffe, I. T. Discarding variables in a principal component analysis. I: Artificial data. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1972**, *21* (2), 160−173.
(45) Fox, E. W.; Hill, R. A.; Leibowitz, S. G.; Olsen, A. R.; Thornbrugh, D. J.; Weber, M. H. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environ. Monit. Assess.* **2017**, *189* (7), 316.
(46) James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*;; Springer: New York, NY, 2013. DOI: 10.1007/978-1-4614-7138-7.
(47) Zambrano-Bigiarini, M. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series; R package version 0.3−10; 2014, http://www.rforge.net/hydroGOF/.
(48) Willmott, C. J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30* (1), 79−82.
(49) Nash, J. E.; Sutcliffe, J. V. River flow forecasting through conceptual models part I-A discussion of principles. *J. Hydrol.* **1970**, *10* (3), 282−290.
(50) Moriasi, D. N.; Arnold, J. G.; Van Liew, M. W.; Bingner, R. L.; Harmel, R. D.; Veith, T. L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50* (3), 885−900.
(51) Olson, J. R. Predicting combined effects of land use and climate change on river and stream salinity. *Philos. Trans. R. Soc., B* **2019**, *374*, 20180005.
(52) Herlihy, A. T.; Paulsen, S. G.; Van Sickle, J.; Stoddard, J. L.; Hawkins, C. P.; Yuan, L. L. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *J. North Am. Benthol. Soc.* **2008**, *27* (4), 860−877.
(53) Bouchez, J.; Moquet, J. S.; Espinoza, J. C.; Martinez, J. M.; Guyot, J. L.; Lagane, C.; Filizola, N.; Noriega, L.; Hidalgo Sanchez, L.; Pombosa, R. River mixing in the amazon as a driver of concentration-discharge relationships. *Water Resour. Res.* **2017**, *53* (11), 8660−8685.
(54) Drever, J. I. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*, 3rd ed.; Prentic-Hall: New Jersey, 1997.
(55) Bishop, K.; Seibert, J.; Köhler, S.; Laudon, H. Resolving the double paradox of rapidly mobilized old water with highly variable responses in runoff chemistry. *Hydrol. Processes* **2004**, *18* (1), 185−189.
(56) Fritz, P., Cherry, J. A., Weyer, K. U., Sklash, M. Storm runoff analysis using environmental isotopes and major ions. In *Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology*; International Atomic Energy Commission: Vienna, Austria, 1976, pp 111−130.
(57) Cormier, S. M.; Zheng, L.; Flaherty, C. M. A field-based model of the relationship between extirpation of salt-intolerant benthic invertebrates and background conductivity. *Sci. Total Environ.* **2018**, *633*, 1629−1636.

DOI: 10.1021/acs.est.8b06777
Environ. Sci. Technol. XXXX, XXX, XXX−XXX