

Water Resources Research

RESEARCH ARTICLE

10.1029/2018WR024054

Key Points:

- A flexible random forest approach for salinity yield modeling is presented
- Field-scale distribution characteristics of new soil property maps were the most important catchment salinity yield predictors
- New 30-m soil cover maps highlight areas where salinity control projects may reduce loads by 1.4% or 76,051 Mg

Supporting Information:

- Supporting Information S1

Correspondence to:

T. W. Nauman,
tnauman@usgs.gov

Citation:

Nauman, T. W., Ely, C. P., Miller, M. P., & Duniway, M. C. (2019). Salinity yield modeling of the Upper Colorado River Basin using 30-m resolution soil maps and random forests. *Water Resources Research*, 55, 4954–4973. <https://doi.org/10.1029/2018WR024054>


Received 6 SEP 2018

Accepted 12 MAY 2019

Accepted article online 20 MAY 2019

Published online 22 JUN 2019

Salinity Yield Modeling of the Upper Colorado River Basin Using 30-m Resolution Soil Maps and Random Forests

Travis W. Nauman¹ , Christopher P. Ely² , Matthew P. Miller³ , and Michael C. Duniway¹ 

¹Southwest Biological Science Center, U.S. Geological Survey, Moab, UT, USA, ²California Water Science Center, U.S. Geological Survey, San Diego, CA, USA, ³Utah Water Science Center, U.S. Geological Survey, Salt Lake City, UT, USA

Abstract Salinity loading in the Upper Colorado River Basin (UCRB) costs local economies upward of \$300 million U.S. dollars annually. Salinity source models have generally included coarse spatial data to represent nonagriculture sources. We developed new predictive soil property and cover maps at 30-m resolution to improve source representation in salinity modeling. Salinity loading erosion risk indices were also created based on soil properties, remotely sensed bare ground exposure, and topographic factors to examine potential surface soil erosion drivers. These new maps and data from previous SPARROW models were related to recently updated records of salinity at 309 stream gauges in the UCRB using random forest regressions. Resulting salinity yield predictions indicate more diffuse salinity sources, with slightly higher yields in more arid portions of the UCRB, and less overall load coming from irrigated agricultural sources. Model simulations still indicate irrigation to be the major human source of salinity (661,000 Mg or 12%) and also suggest that 75,000 Mg (1.4%) of annual salinity in the UCRB is coming from areas with excessive exposed bare ground in high-elevation mountain areas. Model inputs allow for field-scale screening of locations that could be targeted for salinity control projects. Results confirm recent studies indicating limited surface erosional influence on salinity loading in UCRB surface waters, but impacts of monsoonal runoff events are still not fully understood, particularly in drylands. The study highlights the utility of new predictive soil maps and machine learning for environmental modeling.

1. Introduction

The Colorado River is the primary source of drinking water for 40 million people and water for 20,000 km² of irrigated agriculture (Flessa, 2004; National Research, 2007; USDO-BOR, 2012). Due to a variety of causes, high concentrations of dissolved solids (salinity) in the Colorado River and its tributaries threaten the usability of this important water source and cause estimated annual economic damages upward of \$300 million per year (U.S. Bureau of USDO-BOR, 2011). While saline soils and geologic units in the basin are a natural source of salinity within the watershed (Anning et al., 2007; Mueller & Osen, 1988; Spahr, 2000), studies have indicated that 30–45% of dissolved solid loads of the Upper Colorado River Basin (UCRB) originate from irrigated agricultural lands, which occupy less than 2% of the area of the basin (Kenney et al., 2009; Miller et al., 2017).

Loads originating from nonagricultural lands have been hypothesized to come from erosion of saline surface soils (Cadaret et al., 2016; Tillman et al., 2018; Tillman & Anning, 2014a; Weltz et al., 2014) and from dissolution processes as water moves through the subsurface and is delivered to streams as baseflow (Rumsey et al., 2017). Empirical analysis of gauges has shown that about a quarter of watersheds evaluated have moderate to strong correlations between suspended sediment and salinity, indicating that erosional processes are likely playing a role in those watersheds in the UCRB (Tillman et al., 2018; Tillman & Anning, 2014a). However, Rumsey et al. (2017) estimated that 89% of salinity loads originate in the baseflow fraction of streamflow in the UCRB. This finding challenges the significance of direct erosional pathways of salinity transport to surface waters via overland flow. However, the exact mechanisms by which salinity is generated in the baseflow fraction is unknown. Possible processes include dissolution of solids in runoff prior to recharge and eventual groundwater discharge to streams, dissolution during vadose-zone transport, or dissolution after discharge into channels with substrates rich in soluble materials in the middle to lower elevations of the UCRB. The finding that baseflow salinity yields were generally greater in basins with

irrigated agriculture (Rumsey et al., 2017) may suggest that groundwater with short residence times can acquire significant salinity and show up in baseflow loads in anthropogenic time frames.

The recent findings of baseflow importance in annual salinity loads do imply a conceptual contradiction to discharge-based forecasting models that utilize the correlation of discharge with salt (Prairie et al., 2005; Prairie & Rajagopalan, 2007). Prairie and colleagues developed gauge specific nonparametric models for the UCRB that allow for spatiotemporal disaggregation of multiple gauges when developing discharge-based forecasting of salt loads. Although discharge models allow for salt forecasting based on a flow for a given gauge, they do not help spatially identify the potential terrestrial sources of salt, which is better done after removing seasonal and interannual variability in discharge (Tillman & Anning, 2014b).

Much of the work that has been done to spatially identify sources of salinity in the UCRB uses the spatially referenced regressions on watershed attributes (SPARROW) model (Schwarz et al., 2006; Smith et al., 1997). The SPARROW model relates spatially referenced stream gauge chemical measurements (e.g., salinity loads—the dependent variable) to attributes of upstream watersheds (e.g., geology, soils, topography, climate, land use, and transport characteristics) using a nonlinear least squares calibration model. The model is a hybrid empirical-mechanistic model that separates a basin into many small stream reach segments, referred to as incremental reaches heretofore, and associated catchments that are referenced by upstream and downstream routing nodes that are used to create a spatial model of chemical constituent transport by water. Once a calibration model is built, predictions of chemical constituent mass for each incremental reach are then routed through all the incremental reaches to estimate loads for each reach and associated catchment. SPARROW also can account for point inputs (e.g., saline springs), irrigation diversions, and attenuation of load due to interactions (e.g., sorption) with the aquatic environment and reservoirs during the routing phase. In the case of salinity, studies have shown that aquatic and reservoir attenuation is negligible and thus does not need to be included in SPARROW salinity modeling (Anning et al., 2007). Studies using SPARROW in the UCRB have utilized a combination of 1:500,000 scale geology maps, detailed agriculture maps, land cover/use maps, rock chemistry, climate and hydrology layers, and the 1:250,000 scale U.S. General Soil Map (also known as STATSGO2) to characterize catchments (Buto et al., 2014; Kenney et al., 2009; Kenney & Buto, 2012; Keum & Kaluarachchi, 2015; Miller et al., 2017; Soil Survey Staff, United States Department of Agriculture - NRCS, Accessed, 2017).

The SPARROW model has helped to identify salinity sources from agriculture and other land uses in the UCRB (e.g., Miller et al., 2017). However, the geologic and soil data employed by prior SPARROW work are mapped at a much coarser scale than the hydrologic incremental reach network at which predictions are made ($n = 10,789$ reaches for UCRB). Geologic source maps used in SPARROW are compilations of geologic formation maps grouped by experts to combine formations by the potential to produce dissolved solids without any prior direct quantitative connection to dissolution chemistry data. The STATSGO2 soil map does have actual soil property data related to salinity and other useful parameters, but each map unit has multiple soil types that each has specific soil property values resulting in high uncertainty when soil properties are summarized for the soil map units (Helmick et al., 2014). These limitations in inputs can mask variation within and between reaches that would be useful to land owners tasked with saline soil management. Additionally, none of the studies that have employed the SPARROW model in the UCRB have employed independent validations of results (e.g., test set withholding or cross-validation), so a robust predictive evaluation of the models has not been documented.

Unfortunately, the finer scale U.S. Department of Agriculture National Resource Conservation Service soil survey geographic (SSURGO) database is not complete in the UCRB. Where complete, SSURGO is still often coarse in scale in remote areas, making it difficult to fully identify saline soil management issues across such a broad area. However, emerging predictive mapping techniques are enabling production of new soil maps at finer scales that still have largely been untested for modeling and land management applications (e.g., Chaney et al., 2019; Hengl et al., 2017; Nauman & Duniway, 2016; Nauman & Duniway, 2019; Ramcharan et al., 2018). The process of producing these products has become more standardized and generally involves relating topographic, climatic, geologic, ecological, land use, and remotely sensed maps to field soil observations in a predictive machine learning framework (e.g., random forest [RF]; Breiman, 2001). This process essentially quantifies the soil forming processes popularized by Hans Jenny (1941, 1961, 1980) into a predictive framework. These new soil maps are generally at 30- to 250-m resolution in

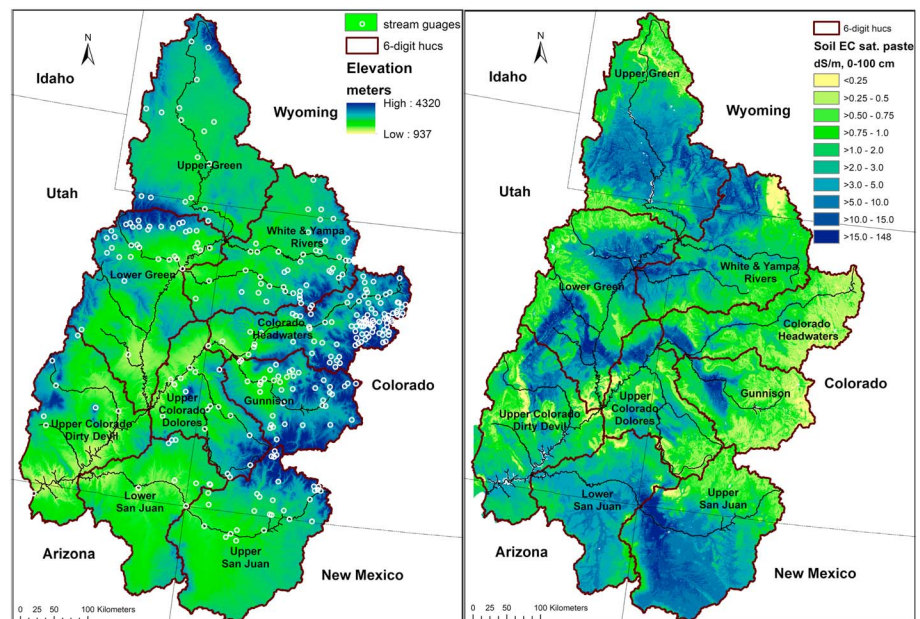


Figure 1. On left, the Upper Colorado River Basin study area extent, elevation, and stream gauge monitoring sites with dissolved solids data used for model building ($n = 309$). Major rivers are shown as black lines, and HUC8 watersheds are outlined in brown and labeled (e.g., Gunnison). On right, the predicted soil electrical conductivity, indicative of salinity, is mapped as an average of predictions at 0-, 30-, 60-, and 100-cm depths (right). This new soil salinity map is utilized as the primary source input for modeling.

raster format, which is a vast improvement over coarser-scale data sets commonly used in modeling frameworks like SPARROW. Application of such fine-resolution data in model development may improve the ability to identify salinity sources at landscape to regional scales, thereby fostering a better understanding of the process by which salts or other contaminants enter surface waters.

This paper aims to utilize both new predictive soil maps and machine learning to help improve spatial predictions of salinity yield in the UCRB. Objectives of this study were to (1) build on prior SPARROW modeling by implementing a RF machine learning approach that utilizes the SPARROW spatial hydrological framework and variables developed by Miller et al. (2017) in addition to improved 30-m soil and topographical maps to estimate UCRB salinity yields and sources, (2) test spatial calibration schemes for salinity load and yield models using RFs, and (3) use new soil cover maps to test if areas with high bare ground exposure are significant erosional salinity sources that may be mitigated. Finally, we (4) examine spatial patterns of important predictor variables within high yield reaches to further illuminate potential mechanisms and processes of salinity transport.

2. Methods

2.1. Study Area

In this study, salinity was assessed for the UCRB above Lake Powell, an area of approximately 280,000 km² (Figure 1). This includes most of the Colorado Plateau, a region of dissected sedimentary lithologic landscapes, and one sensitive to land use changes (Copeland et al., 2017; Schwinning et al., 2008). The Colorado River is sourced from mountainous headwaters, and elevations within the study area range from 940 to 4,300 m. The study area spans parts of six states including Arizona, Colorado, New Mexico, Utah, and Wyoming. Such a large area encompasses a wide range of lithology, topography, and climate, as well as land cover types and uses. Climate varies widely across the watershed with mean annual temperatures ranging from -1.8 to 16.9 °C with an average of 7.7 °C. Annual precipitation averages 368 mm, though some dry areas receive less than 130 mm while higher elevation mountains can receive upwards of 1,400 mm of precipitation (Miller et al., 2017).

2.2. Salinity Modeling Framework

We based most of the spatial structure of our modeling on the SPARROW model as implemented by Miller et al. (2017) and used the same stream gauge monitoring site salinity data and hydrological incremental reach and catchment network (Buto et al., 2017). The gauge data bring together streamflow and different measures of dissolved solids (salinity heretofore) from UCRB gauges in the USGS National Water Information System (<http://waterdata.usgs.gov/nwis>) to estimate a consistent annual salinity load at each gauge. Models were trained and validated with base year 2010 annual load estimates detrended where possible for differences in data record lengths, sample sizes, and variation in discharge over time. The 2010 estimates represent what loads would occur in 2010 under average hydrologic conditions over the period used for detrending. Models to detrend load measurements were created from the 1984–2012 time period (see Tillman & Anning, 2014b for full details of data set development). Of the 312 gauges used by Miller et al. (2017), we used 309 due to slight differences in the extents of our predicted soil property maps along the edge of the study area. Our approach was to replace the explanatory variables used in the SPARROW model, where possible, with finer-resolution variables. While SPARROW is a hybrid statistical-mechanistic model that includes streamflow routing of estimated salinity loads, here we use a purely empirical RF prediction framework to create salinity yield predictions for all incremental reach catchments in the UCRB. We utilized a suite of new soil maps at 30-m resolution created using predictive mapping to replace the geologic source map, as well as some of the other explanatory variables in SPARROW (described in detail below). One key difference between SPARROW and the machine learning approach taken here is that streamflow diversions (e.g., irrigation) were used to adjust the calibration data used to build RF models, whereas SPARROW trains on measured data and then accounts for diversions during load routing. Specifically, we tested whether loads in calibration reaches with diversions should be adjusted upward by assuming that loss of salinity was proportional to loss of flow via diversions within a calibration reach. Since diversions mask load potentially produced from a calibration reach and divert load delivered from upstream reaches, we increased each reach load measured at a downstream gauge in proportion to the amount of flow lost within the reach using the following equation (equation (1)) based on hydrologic reach modeled discharge and diversion data from Miller et al. (2017).

$$L^* = L(Q + D)/Q, \quad (1)$$

where L^* = adjusted load, L = measured load, Q = reach discharge, and D = diversion discharge loss.

Using equation (1) for diversion adjustment assumes that there is no return flow from diversion, which is a potential source of error. Since over 98% of surface water diversions in the UCRB are used for irrigation (Maupin et al., 2018) and UCRB baseflow load appears to be connected to irrigation (Rumsey et al., 2017), we tested models with and without diversion corrections to see which approach would perform best with the pseudo-independent out-of-bag error metrics produced by RFs (discussed later). Overall summed basin loads from catchment predictions were also assessed against the Lees Ferry gauge load measurement to see if diversion adjustments affect those estimates.

Two spatial frameworks were tested for calibration of models by aggregating independent variables for each gauge to these extents: (1) a between-gauge nested-basin defined for each gauge as including all upstream incremental reaches (from Miller et al., 2017, $n = 10,789$) moving through all potential upstream flow pathways from the gauge until all pathways (a) hit another gauge or (b) hit a headwater endpoint and (2) a whole-watershed approach that includes all incremental reaches upstream of each gauge through headwaters for aggregation. The nested-basin method estimated the calibration loads by subtracting salinity loads of gauge(s) at the upstream end of the nested basin from the load at the gauge defining the downstream pour point of the nested basin. These “nested-basin” calibration instances conceptually represent the load produced between gauges with each instance defined by the downstream gauge. The whole-watershed approach characterizes the entire watershed above each gauge as a calibration basin for independent variable aggregation and then uses the salinity load value at each gauge for calibration dependent variable definition. Once models were trained with these calibration data sets, the full set of smaller incremental reach catchments ($n = 10,789$) were also attributed with aggregated values of the independent variables to match the aggregation calculations used for the calibration reaches. This enabled models created from these differing spatial calibration frameworks to then be predicted for these smaller incremental reaches. This differs from SPARROW, which calibrates at the incremental reach catchment scale. To handle diversions in

these frameworks, equation (1) is applied to the entire calibration catchment (nested basin or “whole watershed”). In computing equation (1), D would be computed as a sum of all diversions within the calibration catchment, and Q and L would come from the values at the pour point gauge. Diversions were tested solely for trying to avoid bias in training data, and predictions are only based upon RF model relationship with independent data values in incremental reach catchments (no diversion adjustment at prediction stage).

2.3. Spatial Data Sets Used in Salinity Modeling

Development of the salinity model began by compiling a database of GIS independent variable layers with spatial extent covering the Colorado River watershed above Lees Ferry. These data were then summarized to various catchments used in model fitting and predictions as documented in Table S1 and Text S2 in the supporting information. The various layers were summarized in a variety of ways to look at statistical summaries of each catchment (e.g., mean and quantile values) or by percent of a catchment that falls into a certain class (e.g., percent of catchment with flooded irrigation in high salinity soils). Soil data were queried from the National Cooperative Soil Survey-National Characterization database (NCSS, 2017) and other soil survey data sets to make new soil predictive maps (see Text S2). Topography variables were derived from the National Elevation Dataset (Gesch, 2007; Gesch et al., 2002) and processed using the terrain analysis functions available in SAGA GIS (Conrad & Wichmann, 2011). The U.S. Geological Survey National Gap Analysis Program National Inventory of Vegetation and Land Use (Gergely & McKerron, 2013) were combined with Landsat 8 models of bare ground exposure created from Bureau of Land Management (BLM) vegetation monitoring data to map areas with high amounts of bare ground exposure for each vegetation type. Hydrology and climate variables were computed from Flint and Flint (2007) and the PRISM climate data set (PRISM Climate Group, 2010). Soil erosion-related maps including bare ground exposure, soil properties, and terrain maps were combined together in a variety of ways to create variables that represent potential erosion salinity loading risk dummy variables, which we refer to as “erosion risk indices” heretofore (more details in Text S1 and Table S1). The following subsections and Text S1 document how these data sources were utilized to create predictor variables for salinity modeling, and Table S1 documents the 50 specific variables resulting from that preparation.

2.4. Soil Property Variables

Soil property raster maps with 30-m resolution were created for this study using predictive soil mapping approaches (e.g., Hengl et al., 2017; Ramcharan et al., 2018; Nauman & Duniway, 2019; see development details in Text S2 and Table S2). These maps were summarized for calibration reach catchments (both whole-watershed and nested-basin approaches separately) and for the UCRB incremental reach catchments (for prediction; Soil Property Variables section of Table S1). Maps for 0-, 30-, 60-, and 100-cm depths were averaged for soil electrical conductivity (ec) due to its hypothesized role as the primary salinity source indicator. Other soil property maps were only included for the 0-cm depth (surface) due to hypothesized influence on watershed runoff characteristics. These maps included available water capacity between 0.033 and 1.5 MPa (awc —indicator of water storage potential), soil erodibility (kw ; Sharpley & Williams, 1990, equation 4), sodium adsorption ratio (sar —high sar can limit infiltration), percent rock content by mass ($rock$ —can limit infiltration and armor against erosion), and percent content of fine sands by mass of <2-mm particles (fs —indicator of sandy aeolian soils with high infiltration rates). Soil kw was summarized in other variables as part of testing the potential of soil erosion to contribute to salinity loads in erosion risk indices. We also included a 250-m resolution prediction of the probability of bedrock occurrence ($Brock$) at less than 2-m depth (Shangguan et al., 2017) as total soil depth to bedrock has an influence both on runoff dynamics and on the pool of salts more easily accessible to groundwater in the vadose zone (as opposed to bedrock interactions).

2.5. Topographical Data

Several topographical parameters were created for use directly in salinity modeling as well as for creating risk indices. Details of processing and definitions of all topographic independent variables are available in the supporting information (Text S1 and Table S1). Upslope flow accumulation catchment area ($facc$) and flow length distance to nearest channel ($flen$; calculation details in Text S1) were included as covariates to see if these would help detect regional gully erosion (Bailey, 1935; McFadden & McAuliffe, 1997) related load. Since gully erosion works upstream from channels, areas with lower $flen$ values would have higher

probability of gully erosion issues. We also included the digital elevation model itself (*elev*, meters above sea level), slope gradient in degrees (*slp*), aspect “southness” (*sness*), a -1 to 1 index ($\cos[\text{aspect}-180]$) of how south or north facing any location is, and a topographic protection index (*protind*) to better characterize runoff and water balance dynamics in catchments. Both the *flen* and *facc* were included both as landscape transport characteristics (catchment averages) as well as in creation of erosion risk indices in combination with other layers that together identified pixels with hypothesized higher probability of contributing salinity to surface waters due to surface runoff and erosion processes (e.g., a pixel with high soil salinity, high erodibility, large bare ground exposure, near to a channel, with a large upslope drainage area).

2.6. Bare Ground Exposure Mapping

Exposure of bare ground to the erosive elements of rain drop impact and surface runoff is a primary driver of water erosion (Branson & Owen, 1970; Hernandez et al., 2017; Nearing et al., 1989) and thus a risk factor for salinity loading by route of water erosion. We address this risk factor by mapping percent of bare ground exposure for the year 2013 (details in Text S1), the closest year to 2010 available using Landsat 8. We developed bare ground binary indices from quantile cutoffs within vegetative macrogroup strata in the U.S. Geological Survey National Gap Analysis Program National Inventory of Vegetation and Land Use data set. The aim of this approach was to highlight pixels with high bare ground exposure relative to similar areas around the basin that are likely to be in a persistent bare ground ecological state (Duniway et al., 2016; Miller et al., 2011), with the acknowledgment that some additional error will be introduced due to the time between the 2010 estimates of salinity and the date of the bare ground indices used here (2013). However, we expect this temporal mismatch likely introduces very little error given variables are summarized by the incremental reach catchments and put into relative terms (i.e., percent of catchments with bare ground above a certain quantile with vegetation classes). Areas of elevated bare ground tend to be spatially clustered over time where localized human access (e.g., roads) causes disturbance, whereas broader regional variability in climate patterns might cause most of the basin to have similar fluctuations with climate. Since the 2010–2013 period marked a decline in development and related potential new human disturbance due to the 2009 recession, which particularly stunted both new housing and energy development (Stock & Watson, 2012), we expect human-related bare ground patterns to be relatively consistent over that period. Given that 2010–2013 was a period of drought with both 2010 and 2013 both being less droughty years following extreme drought years (see Nauman et al., 2018 for trends during that period), and disturbance tends to persist for long periods of time in this region (Fick et al., 2016; Miller et al., 2011), we assume that the differences in bare ground between 2010 and 2013 will be minimal, but we cannot account for any associated error within our modeling. There were overlaps between our 90th percentile bare ground (*bgm90q.pct*) and 75th percentile bare ground (*bgm75q.pct*) erosion risk index areas and irrigated agriculture areas (5.9% and 15% of the irrigated areas, respectively) that could be due to mismatch between the dates. However, the irrigated areas may truly have elevated bare ground during fallow periods depending on crop rotations and management of farms.

2.7. Hydrological and Climate Variables

We selected several variables that were used by Miller et al. (2017) to characterize basin hydrological variables in addition to some PRISM-based climate layers (PRISM Climate Group, 2010) to characterize potential impacts of runoff on salinity yields (Table S1, Hydrological and Climate Variables section). The climate variables from Miller et al. (2017) included mean total annual excess of water (*exc*), mean total annual climatic water deficit (*cwd*), mean total annual snowmelt (*melt*), and mean total annual recharge (*rch*), which originated from Flint and Flint (2007). From the PRISM 30-year normal data set (1981–2010), we included mean annual precipitation (*ppt*) and the ratio of summer (June–September) to annual precipitation (*ppratio*) in the salinity model to include overall precipitation patterns and potential influence of regional summer monsoons. All nontopography raster variables were projected and snapped to the original EDNA DEM grid in R with the raster package using a parallelized bilinear interpolation (Hijmans et al., 2016) and then averaged for catchments used in modeling.

2.8. Salinity Source Variables

Following the structure of the SPARROW model in Miller et al. (2017), we used our 0- to 100-cm soil *ec* map as our “geologic source” data set in replacement of the 1:500,000 geologic map. We also included agricultural sources by using the Buto et al. (2014) agriculture map that delineates agriculture by irrigation method.

Sprinkler and flood irrigated lands were separated into saline (>75th percentile of *ec* in basin) and nonsaline (<75th percentile *ec*), resulting in four agricultural irrigation source classes. Our “geologic” sources were delineated by splitting up the soil *ec* map into six distribution-based classes (0-10th, 10-25th, 25-50th, 50-75th, 75-90th, and 90-100th percentile classes; Table S1, Source Variables section). We also included saline spring point sources used in Miller et al. (2017), with each point being included in a springs raster where the pixel coinciding with each spring is attributed with the salinity load. Every pixel received only one source type between the *ec* percentiles or the irrigation class separated by *ec* percentiles. The saline spring sources were included in addition to the *ec* and agriculture sources. Since RF models were trained separately for load (Mg) and yield (Mg/square km) estimates to allow for evaluation of the two approaches, source variables were summarized for each approach within each catchment separately with load prediction models using total area of source types and yield prediction models using percent of catchment occupied by each source type (details in Table S1).

2.9. Erosion Risk Index Maps

We created erosion risk indices for the entire UCRB (including agricultural areas) that highlight pixels that may have a disproportionate impact on loading due to multiple erosion risk factors. Index selection was driven by a priori hypotheses regarding drivers of salinity “hot spots.” We hypothesized greater salinity contributions to surface waters due to surface erosional processes from pixels with some combination of higher soil *ec*, short distance to drainage (*flen*), high soil erodibility (*kw*), high upslope catchment area that may contribute more runoff (*facc*), and/or high bare ground exposure (*bg*). Several thresholds and combinations of the variables were compiled, from which proportions of each catchment that met those criteria were calculated for use as explanatory variables (Table S1, Erosion risk index variables section). Many combinations were tested in initial models to help distinguish important variable thresholds and combinations. The final model chosen was pruned to include only the indices and other variables that were most important to model performance to better understand potential salinity yield drivers.

2.10. Model Assessment and Predictions

Four initial RF models were created based a factorial of the two calibration reach definitions (nested basin and whole watershed) and the two approaches for characterizing instream salinity (yield and load): (1) nested-basin load, (2) nested-basin yield, (3) whole-watershed load, and (4) whole-watershed yield. For these modeling frameworks, predictions made for a catchment can be interpreted as the load or yield produced in that catchment that one could expect to measure in surface water at the catchment outlet.

For initial framework testing, five iterations of RFs were run with 500 trees (*ntree*), and default settings for number of variables selected (*mtry* = # variables/3), and node size of one. Initial RF models included all source and watershed characterization variables in Table S1, totaling 50 variables in all. Load models included source variables in terms of area of each source type within a catchment, which is consistent with the SPARROW approach. Yield models included source variables in terms of percent of area within a catchment occupied by source types. The initial overall model accuracies were evaluated based on model out-of-bag fit and by comparing basin-level summed load predictions to those measured at Lees Ferry, Arizona. Since most complex empirical models will generally produce overoptimistic accuracy results when assessed with training data, RFs produce out-of-bag fit metrics by predicting results onto instances left out of model building for each RF member ensemble tree and summarizing their accuracy (Breiman, 2001). This produces similar assessments to many withholding or cross validation approaches, but the workflow of RF building allows it to occur concurrent to model building. Out-of-bag fit was quantified by calculating the coefficient of determination (R^2) and root mean square error (RMSE). The total basin salinity load predicted from each model was estimated by adding the loads predicted for all incremental reach catchments in the UCRB (for yield models this is done by multiplying the predicted yield times area of the reach catchment). The loads should theoretically match the total load measured at the outlet gauge at Lees Ferry, Arizona. For calibration data adjusted for water diversions, we adjusted the measured Lees Ferry gauge load to account for diversion load loss by increasing load proportional to relative flow loss due to diversions (increased the measured load at Lees Ferry from 5.45 million Mg/year to 5.81 million Mg/year).

The selected model framework was then evaluated in detail to improve upon and further validate the model. The initial RF model was rerun with varying numbers of sampled variables (*mtry*), numbers of trees (*ntree*),

and node size to manually tune model parameters to minimize out-of-bag error. The new model was then pruned iteratively by eliminating variables with low variable importance as calculated by RFs as the out-of-bag percent increase in mean square error (%IncMSE) when a given variable is randomized (details in Text S3). The pruned RF model was then manually tuned to optimize out-of-bag error by adjusting *n*tree, *m*try, and node size.

Regression bias can be produced in RFs due to predictions being solely based on summarizing actual training data instances (Nguyen et al., 2015; Zhang & Lu, 2012). If a model has any error (i.e., all models we are aware of) summarization of these training instances for values at the extremes of the training distribution will always be skewed back toward the center of the distribution due to averaging of any erroneous predictions through the full RF ensemble. We addressed this by creating a linear model correction that fit the RF training set predicted values to the observed values conceptually following Zhang and Lu (2012). The linear adjustment is simply a linear least squares model relating the RF predicted values of training instances as the dependent variable to the observed values. To test final model predictive ability and the approach of linearly correcting model bias, we ran a tenfold cross-validation of the two-step approach of (1) fitting the RF with selected explanatory variables, then (2) creating a prediction bias adjustment linear model. For each fold, the training set was fit using the two-step model approach, and then the fitted model was predicted with the withheld data. This was repeated for all tenfolds such that cross-validated predictions were produced for every stream gauge used in the model. Final selected model performance reporting is documented by plotting and reporting R^2 and RMSE of RF model fit, linear bias-adjusted model fit, RF out-of-bag fit, and tenfold cross-validation fit of the RF bias-adjusted two-step model. We also looked for patterns in residuals across predicted values in plots and across space for spatial autocorrelation using Geary's C test statistic in R (with neighbors weighted by inverse distances between all points; Bivand et al., 2011; Geary, 1954). Because RFs use squared errors to determine tree breaks, they are sensitive to skewed dependent variables. We used a natural log transformation of load and yields to create normal distributions. Since RFs do not produce mean response estimates and simply split data into mostly homogenous subsets, retransformation of predictions to original units was done directly without any bias correction. Multiple predictive soil mapping studies have not included retransformation bias corrections for tree-based models without apparent issues (Hengl et al., 2017; Mulder et al., 2016; Ramcharan et al., 2018) including work that directly assessed both error and prediction intervals after retransformation without bias correction (Nauman & Duniway, 2019).

2.11. Land Cover and Land Use Simulations

Land cover- and land use-related variables were also evaluated for impacts on overall UCRB loads. To do this, we adjusted management related erosion risk indices (e.g., risk indices with bare ground component) and agricultural source extents to reflect removal of these areas in the prediction matrix of incremental reaches to produce simulation predictions. This creates a scenario where exposure of bare ground is decreased (e.g., by successful restoration action) or the agricultural irrigation type is eliminated from a watershed. This allows a new prediction of salinity production from the original RF model matrix in these scenarios. This type of approach has been used to back-cast estimated soil carbon levels to estimate changes globally due to human land use over the last 12,000 years (Sanderman et al., 2017). By subtracting simulated prediction scenarios of land cover or land use changes from the original model prediction of salinity loads, we can get an estimate of how management changes might impact total basin load and load within each incremental reach catchment. To ensure that simulations do not extend beyond the inferential domain learned by models from the training space, we compared simulated conditions to the distributions of nonsimulation variables for catchments not modified in simulations to ensure similarity to the overall environmental variability of the full training set.

3. Results

3.1. Model Selection and Performance

The whole-watershed yield modeling framework without diversion adjustments was selected as the best modeling framework for RF implementation. Models trained with data not accounting for diversions generally had similar out-of-bag accuracy to diversion-adjusted models except for the chosen nested-basin yield model where the training data without diversion improved R^2 values by 0.02 when comparing an average of five model runs. Diversion adjusted yield models also tended to underestimate total UCRB loads when

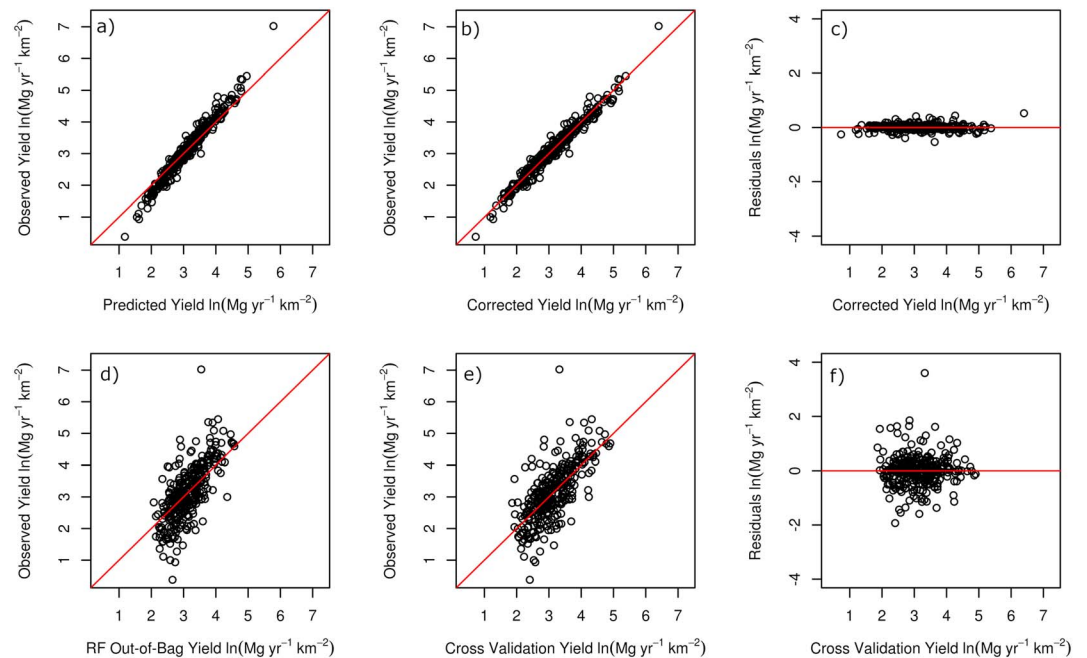


Figure 2. Predicted versus observed and residuals plots illustrating model performance for calibration reaches with 1:1 reference line (red). Part (a) shows the pruned random forest fit before bias correction. Part (b) shows fit after bias correction. Part (c) shows residuals of fit after bias correction. Part (d) shows the pruned random forest out-of-bag model fit. Part (e) shows cross-validation fit of the bias-corrected random forest modeling approach. Part (f) shows the residuals of the cross-validation fit in part (e).

predicted for all incremental reach catchments relative to models without diversion adjustments. Although initial load RF models trained with nested-basins and whole-watershed catchments had higher out-of-bag accuracies than the whole-watershed yield model (out-of-bag $R^2 = 0.66$, 0.88 , and 0.44 , respectively), load models over predicted total UCRB load considerably. The nested-basin reach load model runs predicted more than ~ 19.9 million Mg/year for the UCRB, while the whole-watershed load model predicted ~ 15.8 million Mg/year. The initial RF yield models for both the nested-basin approach and the whole-watershed approach had initial model runs that predicted 5.2 – 6.1 million Mg/year for the UCRB, comparable to the gauge at Lees Ferry (5.46 million Mg/year). The whole-watershed yield initial RF model had considerably better out-of-bag accuracy ($R^2 = 0.44$) compared to the nested-basin yield model ($R^2 = 0.24$); thus, the whole-watershed yield model was chosen for pruning, tuning, and final bias adjustment and cross-validation. The pruning of the whole-watershed yield model narrowed explanatory variables to 17 and improved out-of-bag accuracy from an average R^2 of 0.44 to an average R^2 of 0.49 over five model runs.

The whole-watershed yield RF model fit was very strong ($R^2 = 0.93$, $\text{RMSE} = 0.23$ [in ln units]) but systematically slightly overestimated low values and underestimated high values creating a bias in the model fit. This bias in Figure 2a is likely due to the averaging of potentially skewed extreme yield values across terminal tree nodes in the RF model (Nguyen et al., 2015; Zhang & Lu, 2012). However, the linear correction employed creates an unbiased fit (Figure 2b; $R^2 = 0.96$, $\text{RMSE} = 0.17$ [in ln units]). The improvement in fit from the linear adjustment also is evident in comparing the out-of-bag predictions of the RF model (Figure 2d) to the cross-validation of the RF with linear adjustment approach shown in Figure 2e. Overall model fit strength is slightly lower in out-of-bag ($R^2 = 0.49$, $\text{RMSE} = 0.64$ [in ln units]) than the cross validation ($R^2 = 0.50$, $\text{RMSE} = 0.63$ [in ln units]) assessment that includes bias adjustment. Although considerably lower than the training fit, cross-validation results indicate that the RF with linear adjustment has moderate predictive ability and visibly improves upon prediction bias observable in the out-of-bag (Figure 2d) and training fit in RF predictions (Figure 2a). The highest and lowest values are among the worst predicted in the cross validation, suggesting that these outlying yield measurements are important in the full model to characterize the entire range of basin salinity yields. The final RF model (with linear adjustment) predicted a total UCRB load (by summing incremental reach catchment predictions) of 5.37 million Mg, within 2% of

fs.ave - catchment average fine sand content.
awc.ave - catchment average surface soil available water capacity (15 - 1/3 bar)
sar.ave - catchment average surface sodium adsorption ratio
ec0_75N.pct - % catchment with sprinkler irrigation and soil *ec* between 0th and 75th %tiles.
ec75_100F.pct - % catchment with flood irrigated and soil *ec* between the 75th and 100th %tiles
ec.75q - catchment 75% quantile value of soil *EC*
ec25_50.pct - % catchment with soil *EC* values between 25th and 50 percentiles
slp.ave - catchment average slope
bgm90q.pct - % catchment with bareground values higher than 90th %tile of vegetation macrogroup
rock.ave - catchment average surface rock content
protind.ave - catchment average topographic protection index
kw.ave - catchment average surface soil erodibility
ec0_10.pct - % catchment with soil *EC* between 0th and 10th percentiles
ec.ave - catchment average soil *EC*
kw.75q - catchment 75% quantile value of surface soil erodibility
ec10_25.pct - % catchment with soil *ec* between 10th and 25th percentiles
sness.ave - catchment average south (1) to north (-1) aspect index (cosine[aspect-180])

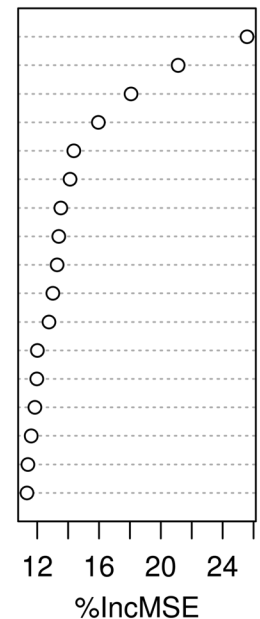


Figure 3. Random forest variable importance plot highlighting variables that cause the highest increase in mean square error (MSE) when randomly permuted in a model iteration. Higher %IncMSE indicates the variable has a larger role in creating an accurate model.

the gauged amount of 5.46 million Mg measured at Lees Ferry. Model residuals did not exhibit evidence of spatial autocorrelation for the RF training fit (Geary $C = 0.908$, $p = 0.156$, two-sided, weighted by inverse distance to all other points), cross validation fit (Geary $C = 0.898$, $p = 0.164$), and out-of-bag fit (Geary $C = 0.897$, $p = 0.1167$), indicating spatial dependence is not significantly influencing predictions. Maps of residuals also did not exhibit any discernable spatial patterns (Figure S1).

3.2. Variable Importance in Model

The final pruned model consisted of 17 variables that included mostly soil properties including multiple summarizations of soil *ec* within catchments. Other selected variables included sprinkler agriculture in less saline areas (*ec0_75.pct*), flooded saline agriculture (*ec75_100F.pct*), bare ground areas above 90th percentiles within vegetation macrogroups (*bgm90q.pct*), and topography layers (Figure 3). Interestingly, soil fine sand content (*fs.ave*), available water capacity (*awc.ave*), and soil sodium adsorption ratio (*sar.ave*) ended up being the most important predictor variables. Areas with high fine sand tend to be aeolian soils with lower salts and high infiltration, and partial variable importance plots show that salinity yields decrease with more fine sands (partial importance plots not published). Higher soil *awc.ave* highlights areas with at least moderate clay content, higher soil organic carbon, and specific higher activity mineralogy types (NCSS, 2017). The soil *awc* map also shows a trend with elevation and may be acting as a surrogate for variation in climate and runoff (no climate or runoff parameters were selected in model pruning). Partial importance plots show that higher values of *awc.ave* and *sar.ave* tend to be associated with higher yields. The fourth and fifth most important variables had very similar %IncMSE values and represented the percent of area in catchments with sprinkler irrigated agriculture in lands with soil *EC* values lower than the 75th quantile (*ec0_100N.pct*), and the percent of area in reaches with flood irrigated agriculture in lands with soil *ec* values higher than the 75th quantile (*ec75_100F.pct*).

3.3. Land Cover and Land Use Simulations

The selected irrigated agricultural variables were adjusted to 0% in all reaches in separate and combined model simulations to estimate impacts on UCRB loads by subtracting simulated basin load predictions from the original RF model predictions. When irrigation types were simulated separately, flood irrigated agriculture (*ec75_100F.pct*) in saline soils accounted for 501,211 Mg of predicted annual load, and sprinkler irrigated agriculture (*ec0_75N.pct*) accounted for 189,291 Mg of load. When both irrigation types were removed in a combined simulation, this resulted in 661,375 Mg being attributed to irrigated agriculture

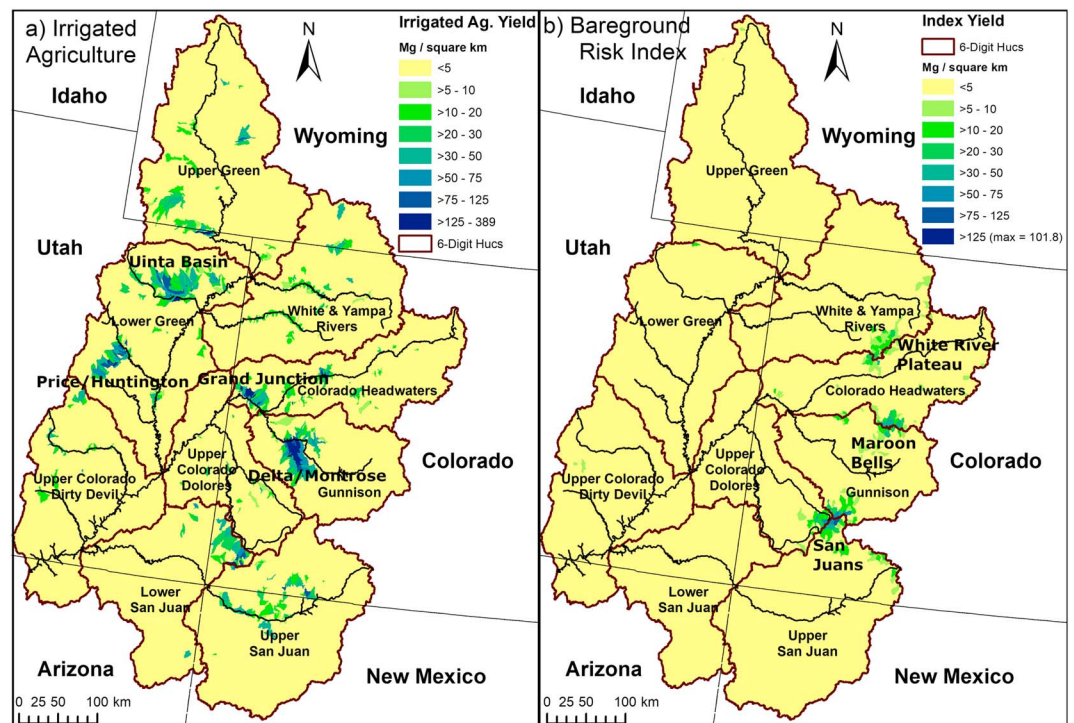


Figure 4. Simulated annual yields in Upper Colorado River Basin incremental reach catchments ($n = 10,879$) from (a) irrigated agriculture and (b) elevated bare ground exposure.

for the model (Figure 4a). Although this is 12.1% of the total predicted UCRB yield, it is lower than previous estimates. Simulated results indicate that in catchments where irrigation was eliminated this resulted in an average load decrease of 0.54 Mg per 30-m pixel, with *ec75_100F.pct* area representing 0.64 Mg per 30-m pixel and *ec0_75N.pct* areas representing 0.38 Mg per 30-m pixel. Irrigated agricultural yield hot spots were primarily observed in the Uinta Basin, Price/Huntington area, Grand Junction valley, and Delta/Montrose areas (Figure 4a). Smaller but significant irrigation yields are also visible in the Lower San Juan, Upper San Juan, and Upper Green River basins (Figure 4a).

The impact of restoration of areas identified by the model as being important erosion risk index predictors of catchment salinity loading (bare ground greater than the 90th percentile within vegetation macrogroups; *bgm90q.pct*) was simulated by reducing the area of these sources to no more than 5% of a catchment area. We did not reduce these source areas to zero because all training catchments had at least a small amount of these areas (minimum of 0.05%) and the ~5% *bgm90q.pct* retained was roughly equivalent to the 25th percentile of the distribution of *bgm90.pct* in training catchments. Simulation results indicate that these areas were associated with 76,051 Mg of load annually, equivalent to roughly 1.4% of predicted UCRB annual load. Simulated results indicate that in catchments where restoration of bare ground was simulated, resulting yields decreased an average of 0.0013 Mg per 30-m pixel. However, despite the broad identification of these areas of elevated bare ground across the UCRB, the *bgm90q.pct* simulation only noticeably decreased yield in high elevation areas in the alpine-subalpine transition zones of the San Juan Mountains, Maroon Bell Mountains, and the White River Plateau (Figure 4b). This seems to indicate that the role of the *bgm90q.pct* variable is likely part of a nested relationship within the RF model trees and localized efforts within the high simulated yield areas in Figure 4b would be more impactful than the 0.0013 Mg per pixel estimate for the broader set of simulated yield decreases in *bgm90q.pct* implemented across the UCRB. The areas exhibiting high yields in the restoration simulation (Figure 4b) all share the attributes of having both higher levels of *bgm90q.pct* and higher soil *ec* than other similar mountain zones around the UCRB.

To ensure simulations did not extend beyond the learning inference of the RF model, five number statistical summaries and the mean of explanatory variables were compared between training catchments not modified in simulations and the overall UCRB training catchment distributions. These comparisons showed

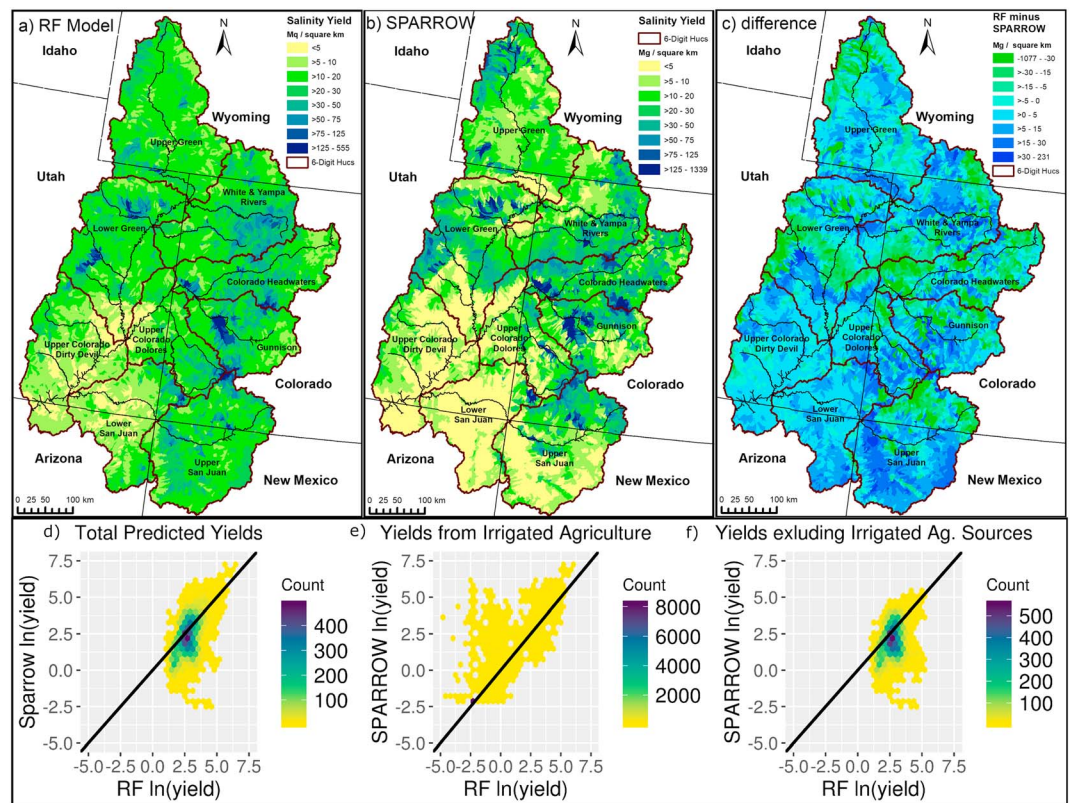


Figure 5. Maps (a–c) comparing predicted 2010 detrended annual salinity yields for all incremental reach catchments in the Upper Colorado River Basin ($n = 10,879$) with (a) the new two-step random forest (RF) model, (b) Miller et al. (2017) SPARROW model, and (c) showing the difference between the RF and SPARROW (RF–SPARROW) highlighting areas where the RF predictions are higher (darker blues) and where SPARROW predictions are higher (green areas). Graphs (d–f) of predicted yields ($\text{Mg} \cdot \text{km}^{-2} \cdot \text{year}^{-1}$) from RF model in this paper and SPARROW based on (d) overall yield, (e) yields associated with agricultural flood irrigation, and (f) yields excluding flood irrigated sources. Graphs are presented as hexagon bins of point density due to large number of reach catchment predictions being compared.

little difference indicating simulations were operating in an interpolative manner (exhaustive statistical summaries available in code comments on Github in “FullWatershedModels.R”). For all simulations there were >87 training catchments not modified that would serve as a contrast in simulations from the original model learning matrix ruleset.

3.4. Patterns in Yield and Comparisons to SPARROW

Patterns in predicted RF yield form several discrete clusters, the largest of which are located in the Gunnison Basin watershed, the lower Green River Basin, the Colorado Headwaters watershed, the White and Yampa Rivers Basin, and the Upper San Juan Basin (Figure 5a). Other isolated areas of high yield were detected in most of the major watersheds, but the Upper Colorado Dirty Devil and Lower San Juan and lower portions of the Upper Colorado Dolores basins had clearly lower yields (Figure 5a). SPARROW model results (Miller et al., 2017) show very similar distribution of yield hot spots to the RF model except in the Upper Green Basin and portions of the Lower Green (Figure 5b). Areas of lower yield are clearly more prevalent in the SPARROW model in the more arid south and southwestern parts of the UCRB as well as in portions of the Upper Green and Gunnison basins (Figures 5a–d and 5f). The SPARROW model predicts yields of zero or close to zero in many semiarid drainages, whereas the RF model tends to predict more yields in the $5\text{--}20 \text{ Mg} \cdot \text{year}^{-1} \cdot \text{km}^{-2}$ range in those same areas. In general, the new RF model predicts more diffuse salinity sources across the UCRB and with greater emphasis on the western slope of the Rockies in Colorado and the Upper San Juan Basin. Although there is correlation between irrigated agricultural yield in RF predictions compared to SPARROW (Figure 5e), the SPARROW model predicts much more load from irrigated agriculture than the RF (1.82×10^6 vs. $6.6 \times 10^5 \text{ Mg}$, Figure 5e).

Comparing model performance between the RF model and SPARROW is difficult due to the lack of independent validation in the UCRB SPARROW studies. Miller et al. (2017) reported a SPARROW yield model training fit with an R^2 of 0.73 and RMSE of 0.47 (ln units), whereas the RF yield model fit here had an R^2 of 0.96 and RMSE of 0.23 (ln units). However, our cross-validation results indicated a model R^2 of 0.50 and RMSE of 0.63 (ln units). Although these cross-validation statistics indicate considerable error sources that still need to be accounted for, it is common to see overfitting in RF and many statistical modeling techniques. Machine learning algorithms simply look for patterns in data regardless of causality, which is why the default evaluation of RFs are based on pseudo-independent out-of-bag errors and not training data fit (Breiman, 2001), and it is standard to use cross-validation or withholding validations to evaluate machine learning models' predictive capability (e.g., Caruana & Niculescu-Mizil, 2006; Hengl et al., 2017). It is difficult to say how much overfitting may be reflected in reported SPARROW fit statistics, but large complicated nonlinear least squares regressions are by no means immune to the phenomenon (Hawkins, 2004). A bootstrapping approach is used in SPARROW to estimate 90% confidence intervals (CIs) to describe uncertainty in regression parameters, and the CIs reported in previous SPARROW runs exhibit quite wide intervals, often with larger ranges than the actual coefficient value used in predictions (Kenney et al., 2009; Miller et al., 2017) indicating relatively high uncertainty. These wide CIs provide evidence that a similar cross-validation approach to ours would likely reveal model overfitting by some unknown magnitude. This unknown potential for overfitting in SPARROW models makes it speculative to categorically discern if SPARROW or the RF model here are more accurate without more comparable validation of SPARROW.

4. Discussion and Conclusions

This study presents a data-driven RF approach to modeling salinity yields into surface waters of the UCRB. The 30-m resolution soil property maps, maps highlighting areas with high amounts of bare ground exposure, and certain topographical layers (Figure 3) were able to represent half of the variation in observed 2010 detrended salinity yield measured at UCRB stream gauges. The RF model also accurately predicted total basin load within 2% when used to predict yield at 10,789 incremental reach catchments prepared in recent SPARROW modeling work. The new 30-m input maps can be used to look within high-yield catchments to evaluate field-scale variation in important predictor variables for planning potential salinity control projects. Simulations based on the developed RF model demonstrated scenarios for how reductions in irrigated agriculture and mitigating bare ground exposure may reduce salinity loading into UCRB surface drainages. Although much of the trend in UCRB yield is similar between RF results and previous SPARROW modeling, the RF results suggest that irrigated agriculture is a significantly smaller source of load and that salinity yields through the UCRB are more diffuse than previously thought.

Like previous studies, we found that irrigated agriculture was a major contributor to UCRB salinity load (12%) and that flooded agriculture in more saline areas was the biggest contributor (9.3%). However, these are much lower contributions than those found by Miller et al. (2017) where irrigated agricultural sources contributed 31% of UCRB load and high yield areas with flooded irrigation contributed 17%. This difference in agricultural yield seems likely to be a result of the new input variables, but it could also be related to how diversions are dealt with in modeling, and inclusion of only a fraction of potential irrigated agriculture sources in the basin during model pruning. In testing RF training strategies, diversions caused underestimates of total basin load in our models, lowered model accuracy slightly, and were not used in the final selected model. Given that 98% of UCRB surface water diversions are irrigation related (Maupin et al., 2018) and associated load is likely to return to surface waters via groundwater (Rumsey et al., 2017), adjusting for diversion seems likely to add error to modeling in this scenario. The differing irrigation results could also be influenced by the differing RF training and prediction framework applied. The RF results also suggest that more load is produced in lower-elevation areas than previous SPARROW models, which also seems likely to be driven by the influence of the soil maps. Finally, the RF approach only identified two of four irrigated agriculture categories as predictors in the model (sprinkler 0-75 and flood 75-100), whereas the SPARROW approach included all irrigated agricultural lands.

Although the RF approach is more empirical in nature than SPARROW, it utilizes the same spatial hydrological topology of previous UCRB SPARROW modeling. The SPARROW approach differs from the RF mainly in using smaller calibration catchments which, while providing more detailed information about

salinity source and transport requires more assumptions about isolating load specific to those smaller catchments. SPARROW also uses a more process-driven flow routing in the salinity prediction process that maintains a mass balance and can deal with other factors mitigating load transport that may come up (e.g., reservoir attenuation). However, despite not having the same theoretical constraints on the prediction and routing of loads, the RF still renders overall summed load of the UCRB very close to the measured amount at Lees Ferry, Arizona. Due to lack of comparable validations, it is not possible to conclude which model is more accurate, but a clear advantage of RFs is that they are highly parallelizable for high-performance computing and thus can be validated robustly. The importance of the new soil variables in the RF model suggests that rerunning SPARROW with these new inputs maps could be helpful, and we feel this should be pursued in future work.

The 30-m soil, topographical, and risk index maps add much more detailed information about the characteristics of the catchments. In particular, the geologic source maps used in prior regional SPARROW salinity modeling are quite coarse (1:500,000)—often having much larger polygons than the incremental reach catchments being modeled. The SPARROW model also assumes that a lithology class will have similar solids available for dissolution in wetter mountain areas as areas of the same formation in the lower elevation semiarid portions. Although this may be true in unweathered bedrock if the geologic groups used are relatively consistent in composition (a rather broad assumption), weathered regolith and soils will probably have much fewer soluble constituents for dissolution in wetter mountain regions as evidenced by the new soil maps used in this paper. The new soil maps used in this study show that 0- to 100-cm soil salinity levels are much lower in the mountains than in the semiarid areas of UCRB and that spatial patterns in soil salinity are quite complex. Although previous SPARROW runs did include STATSGO2 soil data to help discern differences between lithologic sources and soil variability, STATSGO2 is also quite coarse in scale (1:250,000) and is likely to be masking variability due to high uncertainty in associated property estimates (Helmick et al., 2014).

4.1. New Soil Maps

Predictive soil maps produced for this study had moderate to moderately strong accuracies (R^2 between 0.41 and 0.67, mean of 0.55) that are generally more accurate than other soil mapping projects of similar size and scope (e.g., Adhikari et al., 2013; Chaney et al., 2019; Mulder et al., 2016; Ramcharan et al., 2018; Viscarra Rossel et al., 2015). The SoilGrids 250-m project is notable in having higher validation accuracy, but issues with the use of a horizon-based cross-validation (as opposed to a location-based) and uncertainty issues with the 3-D approach used (Nauman & Duniway, 2019) make the true accuracy of SoilGrids uncertain. Other soil salinity mapping case studies have had encouraging results in mapping smaller extents (<100,000 ha) but have relied on remote sensing indices (Nield et al., 2007) and electromagnetic induction measurements (Guo et al., 2013; Taghizadeh-Mehrjardi et al., 2014) that limit the feasibility across large areas and in more vegetated systems. Soil erodibility has been mapped in a very similar fashion across Europe at 500-m resolution with similar validation accuracy ($R^2 = 0.40$; Panagos et al., 2014). However, in the complicated topography of the UCRB, variability within a 500-m pixel can be considerable (e.g., Figures 1 and 6). The bare ground exposure model accuracy (cross validation $R^2 = 0.44$) also compared favorably with recent regional soil cover remote sensing work, which stratified models within vegetation types to reduce noise (cross-validation R^2 between 0.35 and 0.58; Poitras et al., 2018). However, all these predictive map inputs have considerable error, which may have limited salinity model performance.

4.2. Scaling Questions

Although the RF yield approach in this paper used the same gauge and hydrologic network as the SPARROW model implemented by Miller et al. (2017), there are key differences in the model training and prediction spatial frameworks. The chosen RF model framework looks at the entire watershed upstream of each gauge in the model training step, whereas all incremental reaches in the network are considered during the SPARROW calibration process. Conceptually, the RF modeling of gauges further downstream might be unduly influenced by the headwater portions, and considering that the incremental reach catchments used for the prediction step are much smaller, this prompts questions about scaling issues. However, if scaling was the main driver of differences between SPARROW and the RF model, then the mountain reach predictions should be more consistent between the two models, since RF training catchment sizes would be much more similar to those in SPARROW in these areas. SPARROW predicted higher yields within the

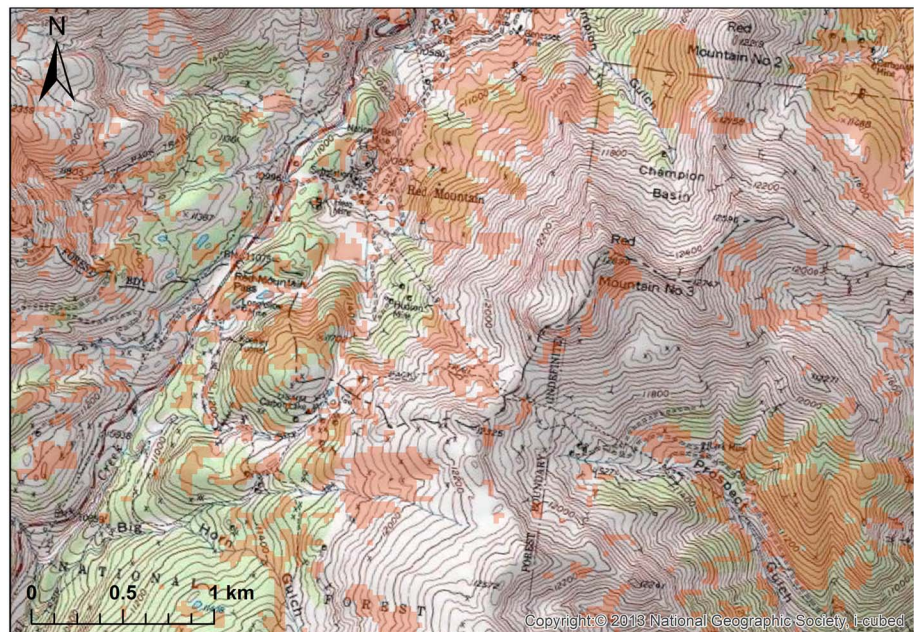


Figure 6. Area in the San Juan Mountains near Red Mountain Pass, Colorado, identified by risk index as having elevated salinity yield in simulations. This area has been heavily mined as the USGS 7.5-min topographical map background indicates. Transparent orange areas indicate pixels identified as having >90th percentile of exposed bare ground relative to areas within the same vegetation macrogroup in the national U.S. Geological Survey National Gap Analysis Program National Inventory of Vegetation and Land Use data set (Gergely & McKerrow, 2013).

mountains that show substantial spatial difference from the RF model (Figure 5c), which suggests other factors are driving model differences. Residuals also do not indicate bias or downstream gradients (Figure S1), and no spatial autocorrelation was detected in residuals produced from model training, cross-validation, or in out-of-bag predictions. However, there still may be some influence of calibration catchment scale that these analyses were unable to detect.

4.3. Erosion Risk Index Insights

Of all the erosion risk indices tested for influence in the RF model, only the percentage of catchments with bare ground exposure higher than the 90th percentile of bare ground within a vegetation macrogroup (bgm90q.pct) was selected in the final model. A simulated restoration action where these areas were reduced to just 5% of catchments (if they had been higher than that before) resulted in a decrease of 76,051 Mg of load annually in the UCRB, or about 1.4% of the total UCRB load. The vast majority of simulated load decrease came in high-elevation subalpine-alpine areas of mountain ranges with high bgm90q.pct values and higher soil salinity than other mountain ranges in the UCRB, which are quite numerous in the basin. The majority of this load is attributed to the San Juan and Maroon Bells mountains in areas with documented legacies of mining (Figures 4b and 6; also Freeman & Weisner, 1984; Steven & Eaton, 1975). This includes areas near Ouray, Red Mountain Pass, and Silverton that have extensive areas of land impacted by historic mining activities that are likely to be linked with these simulated results (Figure 6). Areas of the White River Plateau were also highlighted by the risk index simulation as having elevated yields north of Glenwood Springs, CO (Figure 4b). The White River Plateau area is a much flatter high-elevation area with significant road networks, which may suggest more distributed land uses (e.g., grazing and recreation) may play a role in higher predicted salinity yields. However, these qualitative associations between land use and erosion risk are highly speculative, and more definitive causality requires further investigation.

4.4. Exploring Source Dynamics

Although this and past studies have been able to spatially model salinity sources at the incremental reach scale and have clearly highlighted flood irrigation contributions, processes of natural and more diffuse source propagation into surface waters are less clear. Contributions from erosion has been a focus of research and mitigation efforts (e.g., Cadaret et al., 2016; Tillman et al., 2018), but Rumsey et al. (2017)

provide evidence that diffuse groundwater discharge-driven baseflow accounts for 80-90% of UCRB loads, although it is unclear exactly where in the groundwater recharge cycle this happens, or how long it takes. Several studies indicate that about a quarter of UCRB gauges have relationships between suspended sediment and salinity that indicate surface erosion influence on salinity in those catchments (Tillman et al., 2018; Tillman & Anning, 2014a). Our results are generally consistent with these studies, as we could relate only 1.4% of basin load to one of our erosion risk index variables (bgm90q.pct). Rumsey et al. (2017) also showed that agricultural irrigation areas were associated with increased baseflow loads. This finding is also consistent with our experience in selecting a model strategy that avoided correcting for diversions, which are likely to be moving through short residence time groundwater pathways back into surface waters after being applied for irrigation. Most of the high load baseflow contributions observed by Rumsey et al. (2017) were at middle elevations, but lower-elevation reaches were still highly influenced by high yield baseflow inputs due to more saline geologic materials, which also is consistent with RF model overall yield predictions (Figure 5a). The pattern of midelevation yields with tapering, but still significant yields at lower elevations would also be consistent with our observation of slightly higher yields further into low-elevation areas than previous SPARROW models (Figures 5–5c). This marginally greater significance of yield in lower elevation reaches is potentially due in part to groundwater discharge interacting with more saline materials at these elevation ranges that have more dissolvable solids in soils and geologic formation exposures.

Stepping back to basin-wide salinity sources, if we assume that 80-90% of UCRB load is from baseflow (Rumsey et al., 2017), which includes much of the agricultural inputs, and that ~1-2% comes from elevated bare ground exposure as shown by this study, the remaining 8-18% of loads are thus likely to be originating from uncharacterized surface runoff dynamics. This fraction could be due to error in the models. However, it is possible that the semiarid areas of the UCRB predicted to have slightly higher yields than past studies may be contributing salts sourced in surface runoff due to bare ground exposure not picked up in our relative bare ground indices. Increased bare ground exposure in the region has been documented in widespread reductions in vegetation and biological soil crust cover due to a combinations of postcolonial disturbance that include domestic livestock, proliferation of roads and off-highway vehicle activity, energy exploration and development, urban and ex-urban development, and increasing aridity linked to climate change (Belnap, 1995; Belnap et al., 2004; Copeland et al., 2017; Duniway et al., 2018; Li et al., 2013; Miller, 2005; Miller et al., 2011; Munson et al., 2011; Munson et al., 2011; Nauman et al., 2017; Nauman et al., 2018; Neff et al., 2005; Neff et al., 2008; Poitras et al., 2018; Schwinning et al., 2008). These disturbed semiarid areas could be sources for some of this unaccounted-for load, particularly during monsoon events, which can move substantial surface and near-surface salts in just one event (Laronne & Schumm, 1977). This monsoonal transport mechanism in semiarid areas could explain why our lower-elevation yields are higher than previous models due to better representation of the surface salts available for transport in these events. Rumsey et al. (2017) also acknowledge that the monsoonal pulses in discharge and associated load are difficult to quantify.

4.5. Future Application

The presented salinity yield maps and input layers have the potential to inform land management and research in a variety of ways. The updated salinity yield maps can be used to prioritize catchments for salinity control projects and more targeted studies. The soil maps, risk index, and updated agriculture irrigation salinity maps can help identify areas within these catchments to target and plan projects, which can then be simulated by repredicting from the original RF model with new input data to assess potential outcomes. The detail afforded by the 30-m mapping allows much more specificity as to which landscapes to target mitigation efforts relative to previous coarser spatial resolution modeling. In particular, areas highlighted by the bare soils erosion risk index can be specifically evaluated and targeted for salinity control projects and restoration activities. Other regional work using similar 30-m resolution spatial modeling has highlighted areas in the more arid parts of the UCRB as potential dust sources (Nauman et al., 2018). By combining the utility of remotely sensed vegetation cover maps with soils, climate, hydrological, and topographical data available at 30 m and finer in similar approaches to this paper, a variety of sensitive and vulnerable areas can be identified for land managers to direct resources.

Data generated during this study are available at USGS ScienceBase-Catalog (Nauman, 2019). Code developed for data preparation and modeling is available online on Github (https://github.com/usgs/UCRB_

Salinity). We expect these new data sets, particularly the soil maps, to have utility for several types of users, including regional land managers, water planners, and policy makers. Concurrent work in developing the predictive soil maps used in this study includes distribution of uncertainty maps (following Meinshausen, 2006; Vaysse & Lagacherie, 2017; Nauman & Duniway, 2019) that help better depict spatial variability in the accuracy of maps and also highlights areas where soil variability may exist even within a 30-m pixel. It should be stressed that cross-validation results reported in this study indicate considerable error in the maps and models presented, and we recommend using these data as an initial reconnaissance to direct field visits and further investigation into sites identified in this model before making management or policy decisions. We see potential for further improvements in salinity modeling as soil maps and other data sources improve. This study also suggests that the SPARROW modeling framework could be updated not only with better soils and covariate data, but possibly with machine learning to improve the calibration process.

Acknowledgments

Funding for this work was provided by the Bureau of Land Management and the U.S. Geological Survey Ecosystem Mission Area. Use of trade, product, or firm names is for information purposes only and does not constitute an endorsement by the U.S. Government. We thank Fred Tillman of the U.S. Geological Survey Arizona Water Science Center for helpful comments and suggestions that improved the paper considerably. Data associated with this paper are available online (at <https://doi.org/10.5066/P9QSFJDJN> and code at https://github.com/usgs/UCRB_Salinity).

References

- Adhikari, K., Kheir, R. B., Greve, M. B., Bocher, P. K., Malone, B. P., Minasny, B., et al. (2013). High-resolution 3-D mapping of soil texture in Denmark. *Soil Science Society of America Journal*, 77(3), 860–876. <https://doi.org/10.2136/sssaj2012.0275>
- Anning, D. W., Bauch, N. J., Gerner, S. J., Flynn, M. E., Hamlin, S. N., Moore, S. J., et al. (2007). Dissolved solids in basin-fill aquifers and streams in the southwestern United States: US Geological Survey Report, 2328-0328.
- Bailey, R. W. (1935). Epicycles of erosion in the valleys of the Colorado Plateau Province. *The Journal of Geology*, 43(4), 337–355. <https://doi.org/10.1086/624315>
- Belnap, J. (1995). *Surface disturbances: Their role in accelerating desertification*. *Desertification in Developed Countries* (pp. 39–57). Springer. https://doi.org/10.1007/978-94-009-1635-7_4
- Belnap, J., Phillips, S. L., & Miller, M. E. (2004). Response of desert biological soil crusts to alterations in precipitation frequency. *Oecologia*, 141(2), 306–316. <https://doi.org/10.1007/s00442-003-1438-6>
- Bivand, R., Anselin, L., Berke, O., Bernat, A., Carvalho, M., Chun, Y., et al. (2011). spdep: Spatial dependence: Weighting schemes, statistics and models, R package version 0.5-31, URL <http://CRAN.R-project.org/package=spdep>.
- Branson, F. A., & Owen, J. B. (1970). Plant cover, runoff, and sediment yield relationships on Mancos Shale in western Colorado. *Water Resources Research*, 6(3), 783–790. <https://doi.org/10.1029/WR006i003p00783>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Buto, S. G., Gold, B. L., & Jones, K. A. (2014). Development of a regionally consistent geospatial dataset of agricultural lands in the Upper Colorado River Basin, 2007-10: US Geological Survey, 2328-0328.
- Buto, S. G., Spangler, L. E., Flint, A. L., & Flint, L. E. (2017). Catchment-flowline network and selected model inputs for an enhanced and updated spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin: U.S. Geological Survey data release, <https://doi.org/10.5066/F76T0JT4>
- Cadaret, E. M., Nouwakpo, S. K., McGwire, K. C., Weltz, M. A., & Blank, R. R. (2016). Experimental investigation of the effect of vegetation on soil, sediment erosion, and salt transport processes in the Upper Colorado River Basin Mancos Shale formation, Price, Utah, USA. *Catena*, 147, 650–662. <https://doi.org/10.1016/j.catena.2016.08.024>
- Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms, in Proceedings Proceedings of the 23rd international conference on Machine learning2006, ACM, p. 161-168.
- Chaney, N. W., Minasny, B., Herman, J. D., Nauman, T. W., Brungard, C. W., Morgan, C. L. S., et al. (2019). POLARIS soil properties: 30-m probabilistic maps of soil properties over the Contiguous United States. *Water Resources Research*, 55, 2916–2938. <https://doi.org/10.1029/2018WR022797>
- Conrad, O., and Wichmann, V., 2011, SAGA GIS (www.saga-gis.org): Hamburg, Germany.
- Copeland, S. M., Bradford, J. B., Duniway, M. C., & Schuster, R. M. (2017). Potential impacts of overlapping land-use and climate in a sensitive dryland: a case study of the Colorado Plateau, USA. *Ecosphere*, 8(5). <https://doi.org/10.1002/ecs2.1823>
- Duniway, M. C., Geiger, E. L., Minnick, T. J., Phillips, S. L., & Belnap, J. (2018). Insights from long-term ungrazed and grazed watersheds in a Salt Desert Colorado Plateau Ecosystem. *Rangeland Ecology & Management*, 71(4), 492–505. <https://doi.org/10.1016/j.rama.2018.02.007>
- Duniway, M. C., Nauman, T. W., Johanson, J. K., Green, S., Miller, M. E., Williamson, J. C., & Bestelmeyer, B. T. (2016). Generalizing ecological site concepts of the Colorado Plateau for landscape-level applications. *Rangelands*, 38(6), 342–349. <https://doi.org/10.1016/j.rala.2016.10.010>
- Fick, S. E., Decker, C., Duniway, M. C., & Miller, M. E. (2016). Small-scale barriers mitigate desertification processes and enhance plant recruitment in a degraded semiarid grassland. *Ecosphere*, 7(6). <https://doi.org/10.1002/ecs2.1354>
- Flessa, K. W. (2004). Ecosystem services and the value of water in the Colorado River delta and Estuary, USA and Mexico: Guidelines for mitigation and restoration, *International seminar on restoration of damaged lagoon environments* (pp. 79–86).
- Flint, A. L., & Flint, L. E. (2007). Application of the basin characterization model to estimate in-place recharge and runoff potential in the basin and range carbonate-rock aquifer system, White Pine County, Nevada, and adjacent areas in Nevada and Utah: Geological Survey (US), 2328-0328.
- Freeman, V. L., & Weisner, R. C. (1984). Maroon bells-snowmass wilderness and additions, Colorado: US Geological Survey Professional Paper, no. 1300, p. 459.
- Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The incorporated statistician*, 5(3), 115–146. <https://doi.org/10.2307/2986645>
- Gergely, K. J., & McKerrrow, A. (2013). Terrestrial ecosystems: National inventory of vegetation and land use: US Geological Survey, 2327-6932.
- Gesch, D. B. (2007). The National Elevation Dataset. In D. Maune (Ed.), *Digital elevation model technologies and applications: The DEM users manual*, American Society for Photogrammetry and Remote Sensing, (pp. 99–118).

- Gesch, D. B., Oimoen, M., Greenless, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The national elevation dataset. *Photogrammetric engineering and remote sensing*, 68(1), 5–11.
- Guo, Y., Shi, Z., Li, H. Y., & Triantafyllis, J. (2013). Application of digital soil mapping methods for identifying salinity management classes based on a study on coastal central China. *Soil use and management*, 29(3), 445–456. <https://doi.org/10.1111/sum.12059>
- Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, 44(1), 1–12. <https://doi.org/10.1021/ci0342472>
- Helmick, J. L., Nauman, T., & Thompson, J. (2014). Developing and assessing prediction intervals for soil property maps derived from legacy databases: In GlobalSoilMap Proceedings Orlean, France Arrouays et al.,(eds).
- Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, 12(2). <https://doi.org/10.1371/journal.pone.0169748>
- Hernandez, M., Nearing, M. A., Al-Hamdan, O. Z., Pierson, F. B., Armendariz, G., Weltz, M. A., et al. (2017). The rangeland hydrology and erosion model: A dynamic approach for predicting soil loss on rangelands. *Water Resources Research*, 53, 9368–9391. <https://doi.org/10.1002/2017WR020651>
- Hijmans, R. J., van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., et al. (2016). Package “raster”: R package. <https://cran.r-project.org/web/packages/raster/index.html> (accessed 1 October 2016).
- Jenny, H. (1941). *Factors of Soil Formation*, (Vol. 52, p. 415). New York: New York, McGraw-Hill. <https://doi.org/10.1097/00010694-194111000-00009>
- Jenny, H. (1961). Derivation of state factor equations of soils and ecosystems. *Soil Science Society of America Journal*, 25(5), 385–388. <https://doi.org/10.2136/sssaj1961.03615995002500050023x>
- Jenny, H. (1980). Ecological studies analysis and synthesis vol. 37. The soil resource origin and behavior, Jenny, H. Ecological Studies: Analysis and synthesis. In *The Soil Resource: Origin and Behavior*, (Vol. 37, p. 377). Springer-Verlag: New York, N.Y., USA; Berlin, West Germany.
- Kenney, T. A., & Buto, S. G. (2012). Evaluation of the temporal transferability of a model describing dissolved solids in streams of the Upper Colorado River Basin. *JAWRA Journal of the American Water Resources Association*, 48(5), 1041–1053.
- Kenney, T. A., Gerner, S. J., Buto, S. G., & Spangler, L. E. (2009). Spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin: US Geological Survey, 2328-0328.
- Keum, J., & Kaluarachchi, J. J. (2015). Calibration and uncertainty analysis using the SPARROW model for dissolved-solids transport in the Upper Colorado River Basin. *JAWRA Journal of the American Water Resources Association*, 51(5), 1192–1210. <https://doi.org/10.1111/1752-1688.12302>
- Laronne, J. B., & Schumm, S. A. (1977). Evaluation of the storage of diffuse sources of salinity in the Upper Colorado River Basin: Completion report (Colorado Water Resources Research Institute); no. 79.
- Li, J., Okin, G. S., Skiles, S. M., & Painter, T. H. (2013). Relating variation of dust on snow to bare soil dynamics in the western United States. *Environmental Research Letters*, 8(4). <https://doi.org/10.1088/1748-9326/8/4/044054>
- Maupin, M. A., Ivahnenko, T. I., & Bruce, B. (2018). Estimates of water use and trends in the Colorado River Basin, Southwestern United States, 1985–2010, 2018-5049.
- McFadden, L. D., & McAuliffe, J. R. (1997). Lithologically influenced geomorphic responses to Holocene climatic changes in the Southern Colorado Plateau, Arizona: A soil-geomorphic and ecologic perspective. *Geomorphology*, 19(3), 303–332.
- Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7, 983–999.
- Miller, M. E., 2005, The structure and functioning of dryland ecosystems—Conceptual models to inform long-term ecological monitoring: US Geological Survey Scientific Investigations Report 2005-5197, v. 2005, no. 5197, p. 1.
- Miller, M. E., Belote, R. T., Bowker, M. A., & Garman, S. L. (2011). Alternative states of a semiarid grassland ecosystem: Implications for ecosystem services. *Ecosphere*, 2(5), 1–18.
- Miller, M. P., Buto, S. G., Lambert, P. M., & Rumsey, C. A. (2017). Enhanced and updated spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin: US Geological Survey, 2328-0328.
- Mueller, D. K., & Osen, L. L. (1988). Estimation of natural dissolved-solids discharge in the upper Colorado River basin, western United States, Volume WRIR 87-4069, Department of the Interior, US Geological Survey.
- Mulder, V. L., Lacoste, M., Richer-de-Forges, A. C., & Arrouays, D. (2016). GlobalSoilMap France: High-resolution spatial modelling the soils of France up to two meter depth. *Science of The Total Environment*, 573, 1352–1369. <https://doi.org/10.1016/j.scitotenv.2016.07.066>
- Munson, S. M., Belnap, J., & Okin, G. S. (2011). Responses of wind erosion to climate-induced vegetation changes on the Colorado Plateau. *Proceedings of the National Academy of Sciences*, 108(10), 3854–3859. <https://doi.org/10.1073/pnas.1014947108>
- Munson, S. M., Belnap, J., Schelz, C. D., Moran, M., & Carolin, T. W. (2011). On the brink of change: plant responses to climate on the Colorado Plateau. *Ecosphere*, 2(6), art68. <https://doi.org/10.1890/ES11-00059.1>
- National Research, C. (2007). *Colorado River Basin water management: Evaluating and adjusting to hydroclimatic variability*. National Academies Press.
- Nauman, T. W. (2019). *Salinity yield modeling spatial data for the Upper Colorado River Basin*. USA: U.S. Geological Survey data release. <https://doi.org/10.5066/P9QSFJDJN>
- Nauman, T. W., & Duniway, M. C. (2016). The automated reference toolset: A soil-geomorphic ecological potential matching algorithm. *Soil Science Society of America Journal*, 80(5), 1317–1328. <https://doi.org/10.2136/sssaj2016.05.0151>
- Nauman, T. W., & Duniway, M. C. (2019). Relative prediction intervals reveal larger uncertainty in 3D approaches to predictive digital soil mapping of soil properties with legacy data. *Geoderma*, 347, 170–184. <https://doi.org/10.1016/j.geoderma.2019.03.037>
- Nauman, T. W., Duniway, M. C., Villarreal, M. L., & Poitras, T. B. (2017). Disturbance automated reference toolset (DART): Assessing patterns in ecological recovery from energy development on the Colorado Plateau. *Science of The Total Environment*, 584, 476–488. <https://doi.org/10.1016/j.scitotenv.2017.01.034>
- Nauman, T. W., Duniway, M. C., Webb Nicholas, P., & Belnap, J. (2018). Elevated aeolian sediment transport on the Colorado Plateau, USA: The role of grazing, vehicle disturbance, and increasing aridity. *Earth Surface Processes and Landforms*, 43(14), 2897–2914. <https://doi.org/10.1002/esp.4457>
- NCSS. (2017). National Cooperative Soil Survey Characterization Database.
- Nearing, M., Foster, G., Lane, L., & Finkner, S. (1989). A process-based soil erosion model for USDA-Water Erosion Prediction Project technology. *Transactions of the ASAE*, 32(5), 1587–1593. <https://doi.org/10.13031/2013.31195>
- Neff, J. C., Ballantyne, A. P., Farmer, G. L., Mahowald, N. M., Conroy, J. L., Landry, C. C., et al. (2008). Increasing eolian dust deposition in the western United States linked to human activity. *Nature Geoscience*, 1(3), 189–195. <https://doi.org/10.1038/ngeo133>

- Neff, J. C., Reynolds, R. L., Belnap, J., & Lamothe, P. (2005). Multi-decadal impacts of grazing on soil physical and biogeochemical properties in southeast Utah. *Ecological Applications*, 15(1), 87–95. <https://doi.org/10.1890/04-0268>
- Nguyen, T.-T., Huang, J. Z., & Nguyen, T. T. (2015). Two-level quantile regression forests for bias correction in range prediction. *Machine Learning*, 101(1-3), 325–343. <https://doi.org/10.1007/s10994-014-5452-1>
- Nield, S. J., Boettinger, J. L., & Ramsey, R. D. (2007). Digitally mapping gypsic and natric soil areas using Landsat ETM data. *Soil Science Society of America Journal*, 71(1), 245–252. <https://doi.org/10.2136/sssaj2006-0049>
- Panagos, P., Meusburger, K., Ballabio, C., Borrelli, P., & Alewell, C. (2014). Soil erodibility in Europe: A high-resolution dataset based on LUCAS. *Science of The Total Environment*, 479–480, 189–200. <https://doi.org/10.1016/j.scitotenv.2014.02.010>
- Poitras, T. B., Villarreal, M. L., Waller, E. K., Nauman, T. W., Miller, M. E., & Duniway, M. C. (2018). Identifying optimal remotely-sensed variables for ecosystem monitoring in Colorado Plateau drylands. *Journal of Arid Environments*, 153, 76–87. <https://doi.org/10.1016/j.jaridenv.2017.12.008>
- Prairie, J. R., & Rajagopalan, B. (2007). A basin wide stochastic salinity model. *Journal of Hydrology*, 344(1-2), 43–54. <https://doi.org/10.1016/j.jhydrol.2007.06.029>
- Prairie, J. R., Rajagopalan, B., Fulp, T. J., & Zagana, E. A. (2005). Statistical nonparametric model for natural salt estimation. *Journal of environmental engineering*, 131(1), 130–138. [https://doi.org/10.1061/\(ASCE\)0733-9372\(2005\)131:1\(130\)](https://doi.org/10.1061/(ASCE)0733-9372(2005)131:1(130))
- PRISM Climate Group, 2010, 30-yr climate normals. <http://www.prism.oregonstate.edu/normals/>
- Ramcharan, A., Hengl, T., Nauman, T., Brungard, C., Waltman, S., Wills, S., & Thompson, J. (2018). Soil property and class maps of the Conterminous United States at 100-meter spatial resolution. *Soil Science Society of America Journal*, 82(1), 186–201. <https://doi.org/10.2136/sssaj2017.04.0122>
- Rumsey, C. A., Miller, M. P., Schwarz, G. E., Hirsch, R. M., & Susong, D. D. (2017). The role of baseflow in dissolved solids delivery to streams in the Upper Colorado River Basin. *Hydrological Processes*, 31(26), 4705–4718. <https://doi.org/10.1002/hyp.11390>
- Sanderman, J., Hengl, T., & Fiske, G. J. (2017). Soil carbon debt of 12,000 years of human land use. *Proceedings of the National Academy of Sciences*, 114(36), 9575–9580. <https://doi.org/10.1073/pnas.1706103114>
- Schwarz, G. E., Hoos, A. B., Alexander, R. B., & Smith, R. A. (2006). The SPARROW surface water-quality model: Theory, application and user documentation: US geological survey techniques and methods report, book, v. 6, no. 10, p. 248.
- Schwinning, S., Belnap, J., Bowling, D. R., & Ehleringer, J. R. (2008). Sensitivity of the Colorado Plateau to change: Climate, ecosystems, and society. *Ecology and Society*, 13(2), 28.
- Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., & Dai, Y. (2017). Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems*, 9, 65–88. <https://doi.org/10.1002/2016MS000686>
- Sharpley, A. N., & Williams, J. R. (1990). EPIC-erosion/productivity impact calculator: 1. Model documentation, USDA.
- Smith, R. A., Schwarz, G. E., & Alexander, R. B. (1997). Regional interpretation of water-quality monitoring data. *Water Resources Research*, 33(12), 2781–2798. <https://doi.org/10.1029/97WR02171>
- Soil Survey Staff, United States Department of Agriculture - NRCS, Accessed 2017, U.S. General Soil Map (STATSGO2). Available online at <http://soildatamart.nrcs.usda.gov>.
- Spahr, N. E. (2000). Water quality in the Upper Colorado River Basin, Colorado, 1996–98, US Department of the Interior, US Geological Survey.
- Steven, T. A., & Eaton, G. P. (1975). Environment of ore deposition in the Creede mining district, San Juan Mountains, Colorado; I, Geologic, hydrologic, and geophysical setting. *Economic Geology*, 70(6), 1023–1037. <https://doi.org/10.2113/gsecongeo.70.6.1023>
- Stock, J. H., & Watson, M. W. (2012). Disentangling the Channels of the 2007–2009 Recession: National Bureau of Economic Research.
- Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., & Malone, B. P. (2014). Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*, 213, 15–28. <https://doi.org/10.1016/j.geoderma.2013.07.020>
- Tillman, F., Anning, D., Heilman, J., Buto, S., & Miller, M. (2018). Managing salinity in upper Colorado River basin streams: Selecting catchments for sediment control efforts using watershed characteristics and random forests models. *Water*, 10(6). <https://doi.org/10.3390/w10060676>
- Tillman, F. D., & Anning, D. W. (2014a). A data reconnaissance on the effect of suspended-sediment concentrations on dissolved-solids concentrations in rivers and tributaries in the Upper Colorado River Basin. *Journal of hydrology*, 519, 1020–1030. <https://doi.org/10.1016/j.jhydrol.2014.08.020>
- Tillman, F. D., & Anning, D. W. (2014b). Updated estimates of long-term average dissolved-solids loading in streams and rivers of the Upper Colorado River Basin: US Geological Survey, 2331–1258.
- USDOI-BOR (2011). Quality of Water – Colorado River Basin Progress Report No. 23. U.S. Department of Interior 76 pages, accessed May 7, 2018 at <http://www.usbr.gov/uc/progact/salinity/pdfs/PR23final.pdf>
- USDOI-BOR (2012). Colorado River Basin water supply and demand study. <https://www.usbr.gov/lc/region/programs/crbstudy/final-report/studyreport.html>.
- Vaysse, K., & Lagacherie, P. (2017). Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma*, 291, 55–64. <https://doi.org/10.1016/j.geoderma.2016.12.017>
- Viscarra Rossel, R. A. V., Chen, C., Grundy, M. J., Searle, R., Clifford, D., & Campbell, P. H. (2015). The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. *Soil Research*, 53(8), 845–864. <https://doi.org/10.1071/SR14366>
- Weltz, M., Nouwakpo, S. K., Rossi, C., Jolley, L., & Frasier, G. (2014). Salinity mobilization and transport from rangelands: Assessment, recommendations, and knowledge gaps: General Technical Report 1. Reno, Nevada., 61.
- Zhang, G., & Lu, Y. (2012). Bias-corrected random forests in regression. *Journal of Applied Statistics*, 39(1), 151–160. <https://doi.org/10.1080/02664763.2011.578621>

References From the Supporting Information

- Freeman, T. G. (1991). Calculating catchment area with divergent flow based on a regular grid. *Computers & Geosciences*, 17(3), 413–422. [https://doi.org/10.1016/0098-3004\(91\)90048-I](https://doi.org/10.1016/0098-3004(91)90048-I)
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., et al. (2015). Completion of the 2011 National Land Cover Database for the Conterminous United States—Representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5), 345–354.

- Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe Version 4: Available from the CGIAR-CSI SRTM 90m Database (<http://srtm.csi.cgiar.org>), v. 15.
- Marsett, R. C., Qi, J., Heilman, P., Biedenbender, S. H., Watson, M. C., Amer, S., et al. (2006). Remote sensing for grassland management in the arid southwest. *Rangeland Ecology & Management*, 59(5), 530–540. <https://doi.org/10.2111/05-201R.1>
- R Development Core Team (2008). R: A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing.
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'MASS': Cran R.
- Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), 309–319. <https://doi.org/10.1029/96WR03137>
- U.S. Department of Agriculture (2013). National soil survey handbook, title 430-VI.
- USGS (2016). Using the USGS Landsat 8 Product (http://landsat.usgs.gov/Landsat8_Using_Product.php).