# Comparing Measured and Predicted Phosphorus Levels in the Continental U.S

Anna Herrera, Savannah Johnson, & John Olson, California State University Monterey Bay


## Abstract

Phosphorous is an essential nutrient for living organisms. Anthropogenic sources of phosphorous include but are not limited to water treatment plants, runoff from agriculture and fertilized lawns, failing septic systems, and animal manure storage runoff. Since phosphorous is often a limiting nutrient, i.e., limits growth in aquatic ecosystems, the element can be used as an ecological indicator to measure anthropogenic effects. In order to determine the appropriate phosphorus quantity for streams we need to define natural background in streams. Natural background is the condition of waters in the absence of human-induced alterations based on the best scientific information available. Determining natural background helps set water quality objectives by defining how much phosphorus came from human sources and how much is supplied naturally. This helps set water quality objectives that do not over or under protect streams.

Due to industrialization, pristine natural background water sites are almost nonexistent. Water quality monitoring was also infrequent prior to industrialization. Given the lack of reference sites undisturbed by anthropogenic effects, estimates of the natural ranges of nutrients in streams are needed. We developed a model to predict natural background total phosphorus (TP) concentrations from watershed environmental factors like geology and climate. Because the model includes temporally dynamic predictors of climate, it predicts monthly TP concentrations for each individual stream segment throughout the continental United States from January 2000-December 2020. This model explained 63% of the variation in natural background concentrations.


## Introduction

Due to human influences, scientific interest in determining water quality in natural background developed over the decades. Phosphorus and nitrogen are of special interest because they are aquatic nutrients increasing in quantity and are connected to population growth (Smith, 2003). Phosphorus was chosen her for a couple reasons:

- Phosphorous is an essential nutrient for living organisms and is naturally found in soil and rocks.
- Phosphorous is considered a limiting nutrient, meaning the element is in the least quantity in and therefore limits growth in aquatic ecosystems, the element can be used as an ecological indicator to measure anthropogenic effects.
- Anthropogenic sources of phosphorous include but are not limited to water treatment plants, runoff from agriculture and fertilized lawns, failing septic systems, and animal manure storage runoff.

Due to industrialization, pristine natural background water sites are almost nonexistent. Water quality monitoring was also infrequent prior to industrialization (Smith, 2003). The lack of undisturbed reference sites means scientists need to estimate the natural ranges of nutrients. Here, static and dynamic data was gathered and organized from monitored stream sites throughout the continental United States (CONUS) to predict water quality in terms of total phosphorus.

## Methods

Two models were run consecutively in RStudio and MySQL to calculate predicted total phosphorus levels from 2001-2019 for COMID's across the continental USA.

### *Model 1*

Model 1 is a random forest model object used to predict if the total phosphorus will be above or below the minimum limit of detection (MDL). The object contains six predictor variables and are used with an out of bag error of 3.23%:

- Total phosphorus reported as milligrams per liter converted to a binary value of 0 (below limit of detection 0.001 mg/L) or 1 (above limit of detection)
- Erodibility of soils on agricultural land in watershed
- Mean Evapotranspiration of last 6 months prior to sample
- Mean % of lithological potassium oxide (K2O) content in surface or near surface geology within catchment
- Mean % of lithological nitrogen (N) content in surface or near surface geology within watershed
- Watershed Soil Permeability

Model 1 predicted a below limit of detection with a class error of 12% and above limit of detection with a class error of 3%. Predictions below MDL are made null and not run through second model.

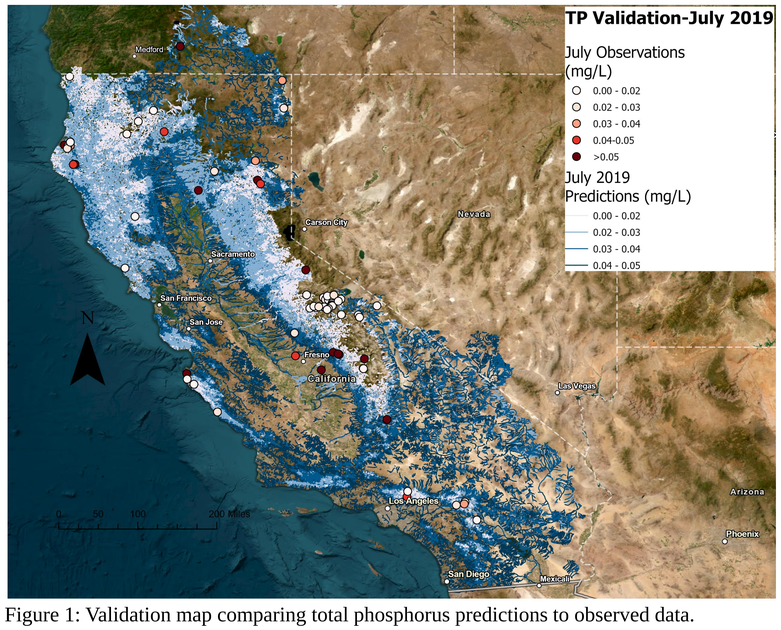


Figure 1: Validation map comparing total phosphorus predictions to observed data.

### *Model 2*

Model 2 is a random forest model object used to predict $\log_{10}$ transformed total phosphorus. Model 2 uncensored phosphorus values utilizes 24 predictor variables with a percent variance explained of 62.83%. The predictors are:

- Total phosphorus reported as mg/L log10 transformed
- Watershed area
- Watershed Percentage clay
- Watershed Elevation
- Mean percent of lithological Phosphorous Oxide ($P_2O_5$) content in surface or near surface geology within a watershed
- Percentage of watershed area classified as lithology type: alluvium and fine-textured coastal zone sediment
- Percentage area in barren land

### *Model 2 predictors continued…*

- Percentage area in coniferous forest
- Percentage area in deciduous forest
- Percentage area in grassland
- Percentage area in mixed forest
- Percentage area in open water
- Percentage area in shrubland
- Mean Precipitation 1, 3, 6, and 12 months prior to sample
- Watershed Runoff
- NHD Elevation slope
- Mean Maximum Temperature the month of sample
- Mean Maximum Temperature of the 6 & 12 months prior to sample
- Mean average Temperature of the 6 & 12 months prior to sample
- Mean Evapotranspiration of 12 months prior to sample

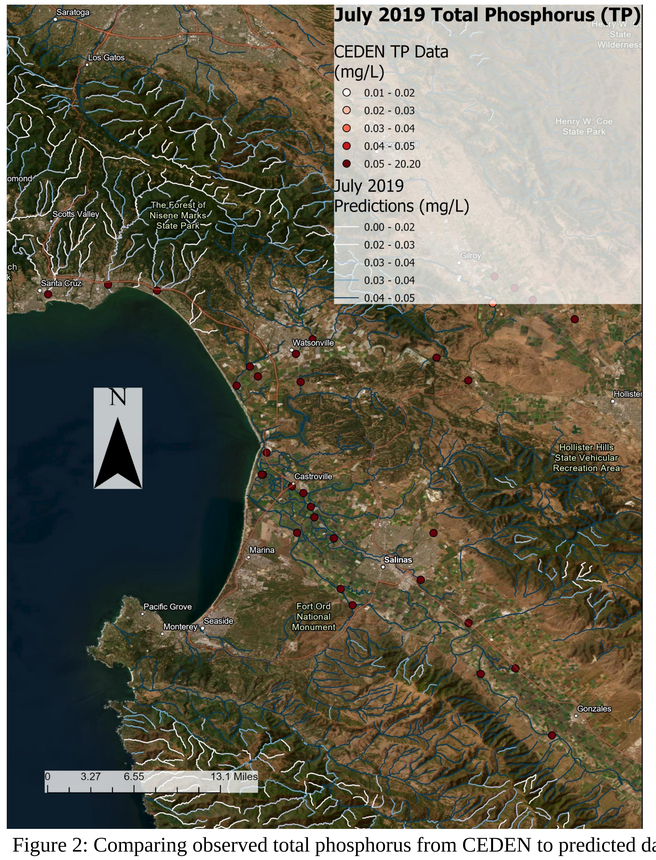


Figure 2: Comparing observed total phosphorus from CEDEN to predicted data locally.

## Results & Discussion

The California TP Validation map shows the predicted total phosphorus is at or below observed levels, showing the predictions are accurate. When comparing the predictions to CEDEN data in the Monterey and Santa Cruz counties we also have a similar comparison. Once the data for all of CONUS is validated, we will have total phosphorus predictions for about 2.5 million steam sites to guide restoration efforts and restore our waters.


Reference:
Richard A. Smith, Richard B. Alexander, and Gregory E. Schwarz (2003) Natural Background Concentrations of Nutrients in Streams and Rivers of the Conterminous United States *Environmental Science & Technology* 37 (14), 3039-3047 DOI: 10.1021/es020663b