



Report No. WI-2017-09
16 December 2017

The Watershed Institute

School of Natural Sciences
California State University
Monterey Bay

<http://ccows.csUMB.edu/pubs/>

100 Campus Center, Seaside, CA
93955-8001

Evaluation of agricultural management practices and water quality in the lower Salinas and Pajaro Valleys

Anna Conlen
Elizabeth Eichorn
Shawnte Greenway
Tiffanee Hutton
Nicole Inglis
Magnolia Morris
Matthew Robinson
Dr. John Olson (instructor)

Senior author contact details:
joolson@csUMB.edu

Acknowledgements

We would like to thank the following individuals for their contributions to this project:

- Pam Krone – Agriculture Water Quality Coordinator, NOAA Office of the National Marine Sanctuaries
- Bridget Hoover – Water Quality Protection Program Director, NOAA Office of the National Marine Sanctuaries
- Sarah Lopez – Technical Program Manager, Central Coast Water Quality Preservation, Inc.

This report primarily represents graduate student work completed within the constraints of a fixed-duration (five-week), limited-verification college class setting.

Cite this report as: CSUMB Class ENVS 660: Conlen A, Eichorn E, Greenway S, Hutton T, Inglis N, Morris M, Robinson M, and Olson J. Evaluation of agricultural management practices and water quality in the lower Salinas and Pajaro Valleys. 2017. Watershed Institute, California State University Monterey Bay, Publication No. WI-2017-09, 40 pages.

Executive Summary

The Central Coast Regional Water Quality Board (CCRWQCB) is currently developing policy to prevent avoidable water pollution from agricultural runoff. Two Agricultural Orders (2012 and 2017) currently require farms that pose the greatest risk to water quality to implement Best Management Practices (BMPs), monitor water quality and report their actions to the CCRWQCB's Irrigated Lands Regulatory Program (ILRP).

Policy makers and stakeholders are interested in how currently applied BMPs are affecting water quality outcomes to inform further management and policy to improve surface water quality. The Water Quality Protection Program administrators at the Monterey Bay National Marine Sanctuary (MBNMS) enlisted the services of CSUMB's ENVS660 Fall 2017 class to examine how reported on-farm management practices relate to water quality monitoring data.

Through geospatial and quantitative analysis, we identified correlations between BMPs and surface water quality. Specifically, we aimed to answer the following questions:

1. Are there particular subwatersheds that are "hot spots" for high nutrient concentrations?
2. Is there a relationship between particular management practices and CMP water quality results?
3. Are there particular geographies where practice adoption is low?
4. Are these the best management practices for improving water quality?

We determined the relationship between 2016 chemical analyte testing results from 22 Cooperative Monitoring Program (CMP) sites on the Central Coast and the 2016 reported nutrient management practices of relevant farms. The analytes considered were total ammonia as N, nitrate + nitrite as N, total nitrogen and total phosphorus as P. We used ArcGIS to delineate catchments for each CMP site, and cleaned and aggregated CMP water quality data and ILRP farm data. We used random forest modelling – a non-parametric empirical modeling technique – to determine which BMP predictors had the strongest correlation with chemical concentration results.

We found that nutrient management practices reported via the ILRP explained up to 46% of the variation in CMP site nutrient concentrations, indicating that BMPs had measurable effects on water quality. In response to our four research questions, we found that:

1. Water quality monitoring sites showed highest surface water nutrient concentrations in the lower Salinas watershed (CMP sites 309BLA, 309TEH, 309MER, 309ASB, 309JON).
2. The BMPs with the strongest correlations to water quality were: classifying practices a trade secret which was associated with increased nitrogen concentrations, and evaluating fertilizer needs and timing, scheduling fertilizer to match crop requirements, and measuring nitrogen and phosphorus content of applied organics which were all associated with decreased nutrient concentrations.
3. Management practice adoption rates varied both spatially and among practices, but were generally lowest in the upper Pajaro and Salinas watersheds (above Gilroy and Salinas respectfully). Evaluating fertilizer needs and timing was the most widely adopted practice, with more than 50% of cultivated acreage in almost every sub-watersheds using this practice. Measuring nitrogen and phosphorus content of applied organics was applied to less than 25% of cultivated acreage in most sub-watersheds.
4. Although some practices were related to nutrient concentrations, considering the limited scope of this study and the complexities of the data, we were unable to determine conclusively which BMPs were the most effective at improving water quality.

ILRP and CMP datasets include several other parameters and variables that should be considered for a more complete understanding of which practices growers can adopt to improve overall water quality on the Central Coast. The methodology used in this study can serve as a trial that informs future research needs, and should not be taken out of context or over-interpreted.

Table of Contents

Acknowledgements	2
Executive Summary	3
Table of Contents	5
1 Introduction	6
1.1 Background	6
1.2 Study Site.....	7
1.3 Project Goals.....	9
2 Methods	10
2.1 Overview.....	10
2.2 Subwatershed delineation.....	10
2.3 Data management	14
2.3.1 Data sources	14
2.3.2 CMP Data Cleaning.....	14
2.3.3 ILRP Data Cleaning	16
2.3.4 Data matrix creation	18
2.4 Random forest statistical analysis.....	18
3 Results	20
3.1 Model results.....	20
3.2 Are there particular subwatersheds that are “hot spots” for high nutrient concentrations?	27
3.3 Is there a relationship between particular management practices and CMP water quality results?	30
3.4 Are there particular geographies where practice adoption is low?	30
3.5 Are the ILRP BMPs the best management practices for improving water quality?.....	31
4 Discussion.....	32
4.1 Study limitations.....	32
4.2 Areas for Future Study	35
5 Conclusion	36
7 References	37
8 Appendix	39

1 Introduction

1.1 Background

The Central Coast Region of California extends approximately 250 miles from San Mateo County to Santa Barbara County and contains roughly 435,000 acres of irrigated land, over 3,000 agricultural operations, more than 17,000 miles of surface waters, and approximately 4,000 square miles of groundwater basins (CCRWQCB 2012). The Porter–Cologne Water Quality Control Act provides California with some of the most stringent water pollution control regulations in the country; yet due to years of unchecked nonpoint source pollution from agricultural runoff, the Central Coast Region routinely reports high rates of nutrient loading and toxicity (Starner et al. 2006, CCRWQCB 2011).

Agricultural Order 2.0, approved by the Central Coast Regional Water Quality Control Board (CCRWQCB) on March 15, 2012, required growers whose farms pose the greatest risk to water quality in the Central Coast Region of California to implement Best Management Practices (BMPs) and to monitor the quality of water discharged from irrigated land they own or operate in order to prevent avoidable water pollution (CCRWQCB 2012). On March 8, 2017, CCRWQCB approved Ag Order 3.0, which added a groundwater monitoring component and expanded the Total Nitrogen Applied reporting obligation (CCRWQCB 2017). Agricultural Order 1.0, implemented in 2004 by the Central Coast Regional Water Quality Control Board (CCRWQCB), and the iterations that followed, created a need for a new regulatory framework to ensure grower compliance and a means to monitor regional changes in water quality.

The Irrigated Lands Regulatory Program (ILRP) is administered by the CCRWQCB and serves as the means through which the board regulates discharges from irrigated agricultural lands used for commercial crop production through issuing Conditional Waivers of Waste Discharge Requirements. Depending on the size of a grower's operation, the type of crops being grown and a ranch's proximity to impaired waterways, an agricultural operation will be placed into one of three tiers through the ILRP: Tier 1, 2, or 3. Tier 3 growers potentially pose a larger risk to water quality and are thus subjected to more extensive regulations while Tier 1 and Tier 2 growers have fewer obligations under the most recent ag orders.

Complying with these regulations can require a substantial amount of effort and fiscal expenditure on the part of growers. The Cooperative Monitoring Program (CMP) evolved in accordance with Ag Order 2.0 as a means of lessening the burden of

monitoring requirements on growers. Over 99% of all farms participating in the ILRP have elected to take part in the CMP and thus pay Preservation Inc., a local nonprofit, an annual fee to perform mandatory surface water quality monitoring on their behalf. Preservation Inc. monitors 50 sites throughout the Central Coast Region.

While growers have been tasked with re-thinking and altering many of their on-farm practices, they have yet to be presented with concrete evidence that their efforts over the last few years have improved regional water quality. Our clients, the Water Quality Protection Program administrators at the Monterey Bay National Marine Sanctuary (MBNMS), have therefore requested an analysis of how the water quality monitoring data of the lower Salinas and Pajaro Valleys relates to reported on-farm practices in order to help guide the development of future agricultural orders.

1.2 Study Site

The Salinas and Pajaro watersheds, on the Central Coast of California, span parts of Monterey, Santa Cruz, San Benito, San Luis Obispo and Santa Clara counties and share boundaries with Elkhorn Slough, Moro Cojo and Alisal watersheds on the coast. All drain into the Monterey Bay National Marine Sanctuary, comprising the majority of the bay's freshwater inputs (Fig 1). Land use across the upper watersheds varies while the lower watersheds are heavily used for agricultural production where a mild, Mediterranean climate makes year-round cultivation possible. Primary crops in the lower Salinas Valley include leaf and head lettuces, strawberries and various other vegetable row crops (MCAC 2016) while the majority of crops in the Pajaro Valley include caneberries (raspberries and blackberries), strawberries, apples and vegetable row crops (SCCAC 2016). The fertile soil in the flat valley bottoms of these two regions makes growing conditions ideal and agricultural operations highly profitable. The agriculture industry contributes billions of dollars to the local economies on an annual basis.

Due to decades of intensive agricultural production, groundwater and surface waters in the lower regions of the Salinas and Pajaro Watersheds are compromised by a variety of nonpoint source pollutants, especially nitrate. Harter and Lund with UC Davis (2012) found that nitrogen fertilizer and animal wastes are the primary sources of nitrate pollution in the Salinas Valley and that over one third of domestic and irrigation wells in the Salinas Valley exceed the drinking water standard for nitrate (10 mg/L as N) deemed acceptable by the California Department of Public Health. Comparable agricultural practices and histories of the Pajaro Valley make similar findings likely.

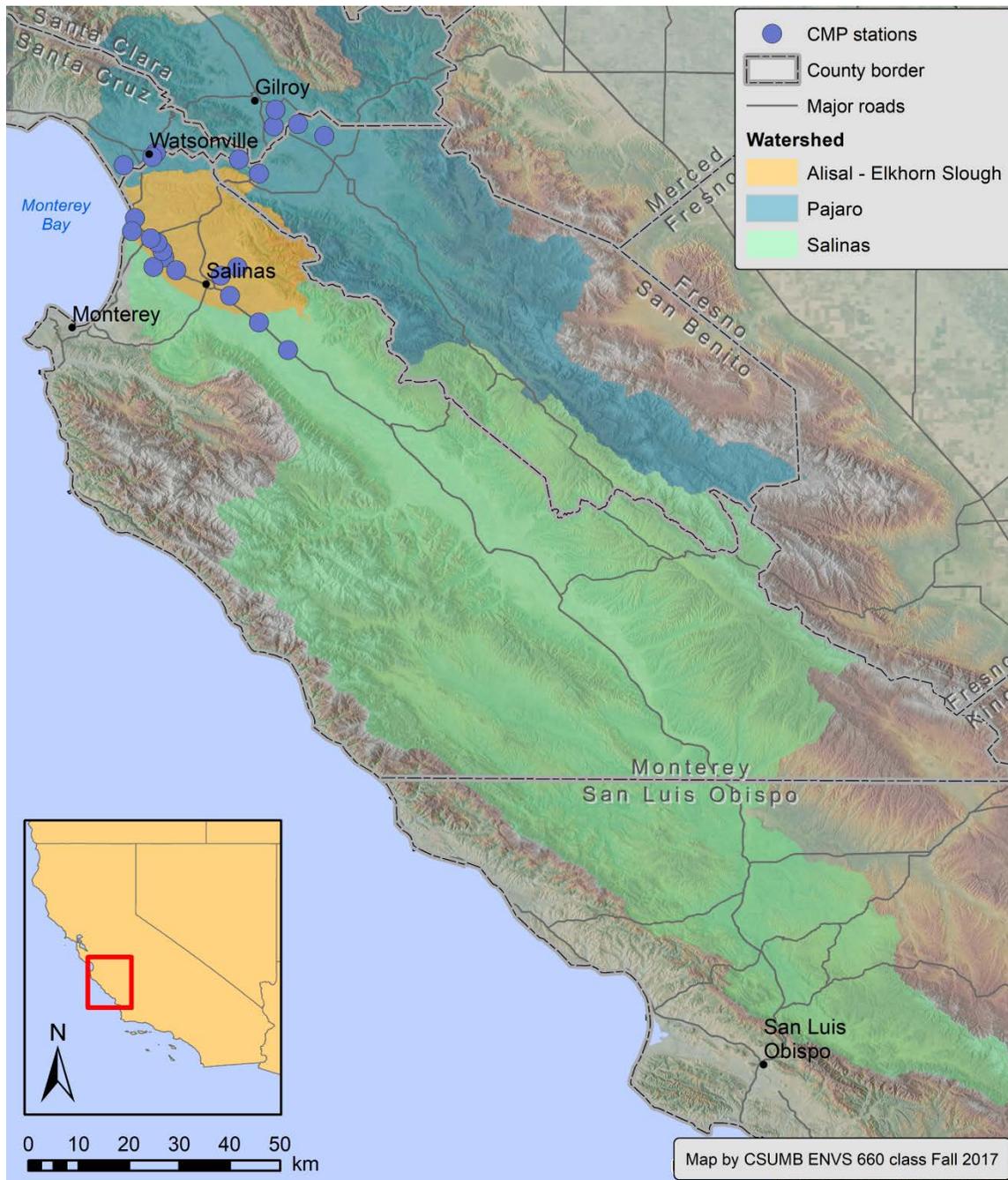


Figure 1. Watersheds and Cooperative Monitoring Program water quality monitoring sites assessed in this study. Located on the Central Coast of California, the study area falls under the jurisdiction of the Central Coast Regional Water Quality Control Board.

The 2010 List of Impaired Waterbodies, pursuant to Section 303(d) of the federal Clean Water Act, designated fifteen waterbodies in the Lower Salinas watershed and five water bodies in the Pajaro River watershed as impaired by nitrate pollution (CCRWQCB

2011) while others were listed for a variety of other pollutants including, but not limited to: un-ionized ammonia, low dissolved oxygen, orthophosphate and chlorophyll *a* (CCRWQCB 2013). The effects of high nutrient concentrations are far reaching and can include drinking water contamination and negative impacts on natural systems. In fact, high nutrient concentrations in coastal freshwater systems have been linked to marine mammal and shorebird mortality in the Monterey Bay National Marine Sanctuary (Scholin et al. 2000, Work et al. 1993).

Few deny that water pollution on the Central Coast is a negative externality of local agricultural production, but developing a region-wide solution to lessening the problem is, and will continue to be, a difficult task. However, little progress can be made in the region if the efficacy of best management practices is not assessed. Time and money, on the part of growers and regulators, will certainly be wasted on practices that do not contribute to water quality improvement if assessments are not made. Therefore, the first step in achieving improved region-wide water quality is determining which practices are most effective.

1.3 Project Goals

The primary goal of this project was to assess whether correlations exist between nutrient BMP implementation by farms and water quality results in the lower Salinas and Pajaro Valleys for the 2016 calendar year.

Specifically, we aimed to address the following questions:

1. Are there particular subwatersheds that are “hot spots” for high nutrient concentrations?
2. Are the ILRP nutrient BMPs the best practices for improving water quality?
3. Is there a relationship between particular management practices and CMP water quality results?
4. Are there particular regions where practice adoption is low?

2 Methods

2.1 Overview

We performed our analysis using CMP 2016 chemical analyte data and 2016 ILRP data for all farms that reported to the Regional Board for that year. We delineated subwatersheds (catchments) using ArcGIS, determined the practices that occurred upstream of each CMP station, cleaned and combined CMP (response) and ILRP BMP (predictor) data for model input and built a random forest model using our compiled dataset to identify any correlations between BMP implementation and water quality. We assessed all correlations between predictors and response variables on watershed and catchment scales. Catchment scale refers to either response or predictor data associated with a portion of landscape that drains directly to a CMP station and excludes upstream contributions. Watershed scale refers to data from both the local catchment and all upstream catchments.

2.2 Catchment delineation

2.2.1 Data sources

The GIS analysis included the following geospatial data:

- Digital elevation model (DEM), 1/3 arc-second (10m) resolution, The National Map, U.S. Geological Survey (USGS), 2013. Projection: GCS NAD 1983
- National Hydrography Dataset (Flowline, subregion 1806). U.S. Geological Survey, 2013, National Hydrography Geodatabase: The National Map. Available from: [<https://viewer.nationalmap.gov/viewer/nhd.html?p=nhd>], accessed [2017 November 03].
- CMP water quality monitoring station latitude and longitude data from the California Environmental Data Exchange Network's (CEDEN) Station Lookup List.
- 2014 ranch atlas (polygon), Monterey County Agricultural Commissioner's Office, April, 2014. Extracted from California Ag Permits (CAP) system. Projection: GCS NAD 1983. Available from: [<http://montereycountyopendata-12017-01-13t232948815z-montereyco.opendata.arcgis.com/datasets/2011-ranchmap-final>], accessed [2017 November 03].

2.2.2 Catchment analysis

We delineated catchments using CMP monitoring sites as pour points using ArcGIS Desktop version 10.5 to link farms and their management practices to the CMP site where their surface water runoff drains (Figs 2 and 3; Table 1). Using ModelBuilder, we:

- Burned NHD stream layer 10 m into the DEM (Raster Calculator tool).
- Filled spurious sinks to account for small errors and imperfections in the DEM (Fill tool).
- Created a raster that determined flow direction based on each cell’s steepest downslope neighboring cell (Flow direction tool).
- Created a raster of accumulated flow in each cell based off the flow direction raster (Flow accumulation tool).
- Snapped CMP sites to the implied streams (done manually using Edit Features).
- Determined watershed boundaries by calculating the contributing cells above each flow gauge using the flow direction raster (Watershed tool).
- Corrected small errors in the watershed delineation (Dissolve tool).
- Projected watershed boundaries to a Projected Coordinate System to calculate areas in acres (Raster to polygon and Project tools, calculated geometry in attribute table).
- Determined which Tier 3 farms were split between more than one watershed; manually cut farm polygons to calculate acreage in each watershed.
- Joined farm points to respective watersheds, denoted by station codes, and exported resulting attribute table to Excel for data analysis (Spatial Join).

Table 1. Cooperative Monitoring Program water quality monitoring sites used in analysis. All sites are located in the lower reaches of the Salinas, Pajaro and Alisal area watersheds.

<u>Station Code</u>	
305CAN	309ALG
305CHI	309ASB
305COR	309BLA
305FRA	309CCD
305LCS	309ESP
305PJP	309GAB
305SJA	309JON
305TSR	309MER
305WSA	309MOR
	309NAD
	309OLD
	309QUI
	309TEH

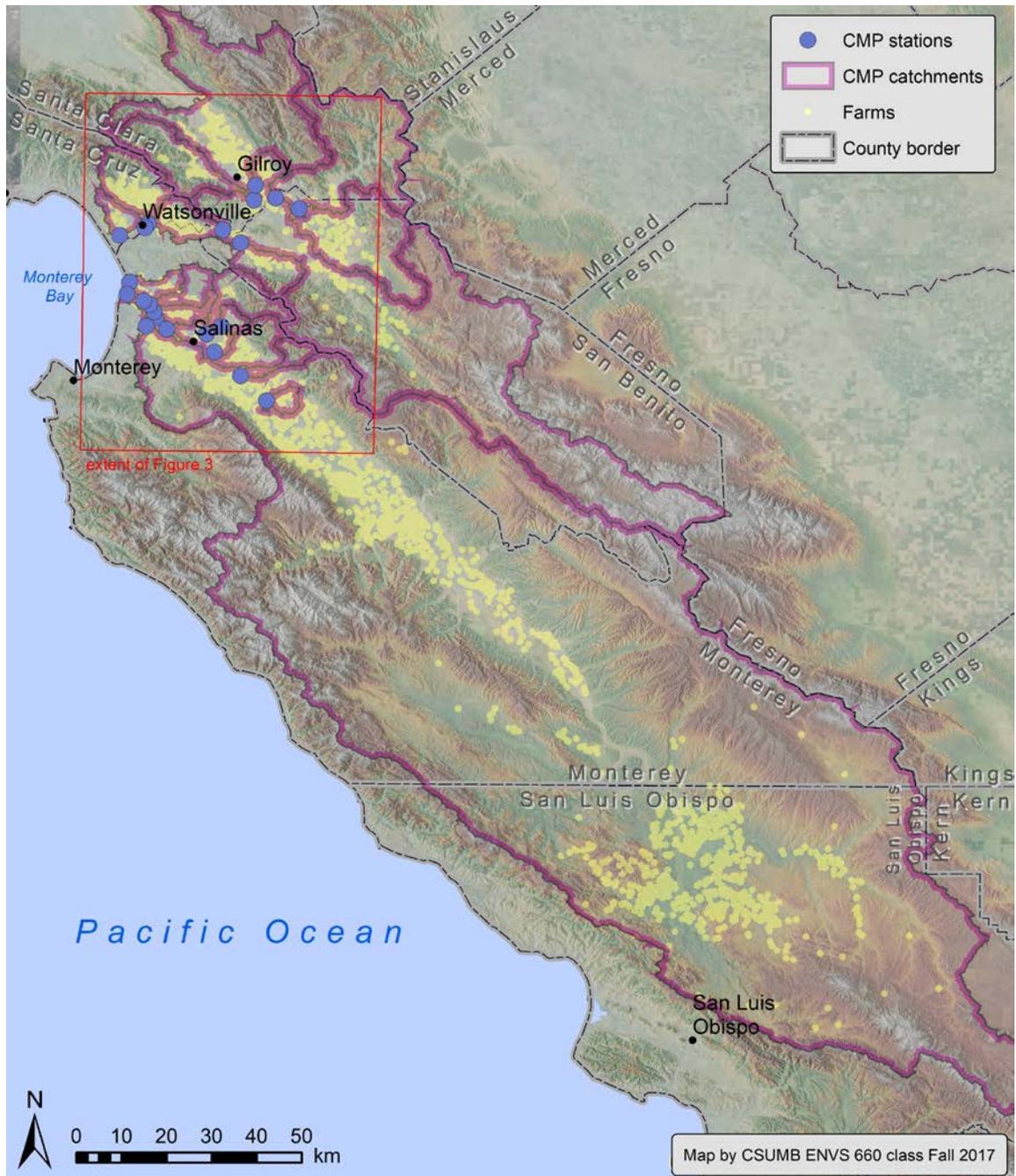


Figure 2. Watersheds and Cooperative Monitoring Program water quality monitoring sites assessed during this study. Located on the Central Coast of California, the study area falls under the jurisdiction of the Central Coast Regional Water Quality Control Board.



Figure 3. Catchment delineations in the lower Salinas watershed, denoted by CMP station code.

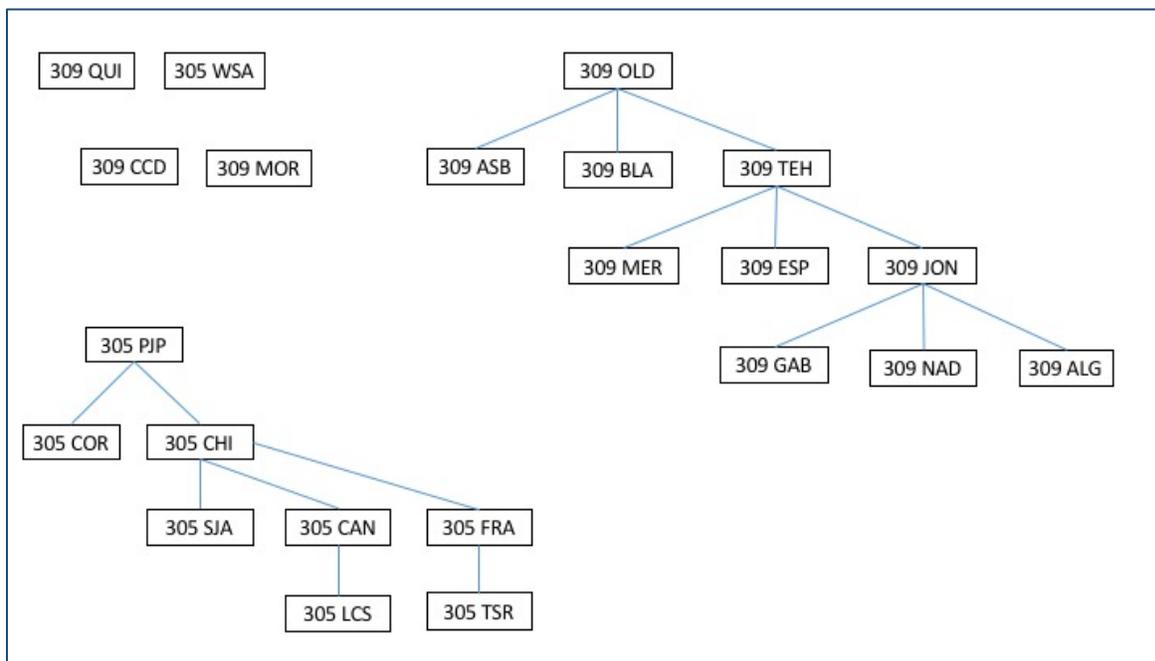


Figure 4. Catchment flow chart used to inform our data preparation process. Note: these catchment relationships were based on both elevation connectivity, and informed by local experts (S. Lopez, pers. comm.), but were not ground-truthed. Watershed data for each individual CMP station was calculated by combining data from the local catchment with upstream catchments (see Table 2).

2.3 Data management

2.3.1 Data sources

We used the following data as model inputs:

- CMP water quality data, 2016, CEDEN Station Lookup List.
- CMP discharge data, 2016, CEDEN.
- ILRP Farm management practice data, 2016, CCRWQB, shared with us by the Monterey National Marine Sanctuary.

2.3.2 CMP Data Cleaning

We extracted chemical concentration data for four analytes at 22 CMP monitoring sites (Table 1) and calculated average concentrations for four different temporal scales at each of the sites. Temporal scales were chosen to investigate whether certain times of the year showed specific patterns of predicted water quality in relation to BMP implementation. We calculated the average concentrations at each CMP site on the following time scales: October to March (winter), April to September (summer), January to December (annual) and dates closest to the recorded first flush event of October 15

(first flush) (Bilginsoy and Ruby 2017). Initial cleaning steps were performed using RStudio (RStudio Team 2015), and temporal calculations were performed manually in Microsoft Excel. The following list outlines the CMP data cleaning process:

- Removed lab blank and lab quality assurance samples.
- Selected Salinas and Pajaro watershed CMP sites (Table 1).
- Extracted relevant columns: StationCode, SampleDate, MatrixName, Analyte, Unit, Result, ResultQualCode, MDL, RL and QACode.
- Extracted relevant analytes: “Ammonia as N, Total”, “Nitrate + Nitrite as N, Total”, “Nitrogen, Total, Total” and “Phosphorus as P, Total”.
- Calculated analyte values listed as “non-detect” to be one-half of the minimum detectable limit (MDL).
- Separated each temporal range into data sheets according to date ranges for annual, winter, summer and first flush. Note: the closest dates to the first flush event were October 25 and 26, 2016.
- Calculated watershed-scale concentration response variables by averaging analyte concentrations for each monitoring site per temporal range in Microsoft Excel using the following equation:
 - $=\text{AVERAGE}(\text{IF}(\text{Analyte Code}=\text{range of analyte cells}*(\text{CMP site code}=\text{range of CMP site codes}),\text{range of concentration cells},""))$
- Calculated catchment-scale concentration response variables by subtracting upstream concentrations from downstream concentrations according to the catchment flow chart (Fig. 4), summarized in Table 2.
- Calculated watershed-scale load response variables using the CMP discharge data by averaging discharge per site and multiplying by the corresponding average analyte concentration.
- Calculated catchment-scale load response variables by subtracting upstream load from downstream load (Table 2).

Table 2. Equations for calculating catchment–scale response variables and watershed–scale predictor metrics based on catchment relationships and movement of water between them.

Watershed to Catchment Equations (Responses)	Catchment to Watershed Equations (Predictors)
<i>Pajaro Watershed</i>	
305CAN = 305CAN - 305LCS	305CAN = 305CAN + 305LCS
305CHI = 305CHI - (305SJA + 305CAN + 305FRA)	305CHI = 205CHI + 305SJA + 305CAN + 305FRA
305COR = 305COR	305COR = 305COR
305FRA = 305FRA - 305TSR	305FRA = 305FRA + 305TSR
305LCS = 305LCS - (305COR + 305CHI)	305LCS = 305LCS
305PJP = 305PJP - (305COR + 305CHI)	305PJP = 305PJP + 305COR + 305CHI
305SJA = 305SJA	305SJA = 305SJA
305TSR = 305TSR	305TSR = 305TSR
305WSA = 305WSA	305WSA = 305WSA
<i>Salinas Watershed</i>	
309ALG = 309ALG	309ALG = 309ALG
309ASB = 309ASB	309ASB = 309ASB
309BLA = 309BLA	309BLA = 309BLA
309CCD = 309CCD	309CCD = 309CCD
309ESP = 309ESP	309ESP = 309ESP
309GAB = 309GAB	309GAB = 309GAB
309JON = 309JON - (309GAB + 309NAD + 309ALG)	309JON = 309JON + (309GAB + 309NAD + 309ALG)
309MER = 309MER	309MER = 309MER
309MOR = 309MOR	309MOR = 309MOR
309NAD = 309NAD	309NAD = 309NAD
309OLD = 309OLD - (309ASB + 309BLA + 309TEH)	309OLD = 309OLD + 309ASB + 309BLA + 309TEH
309QUI = 309QUI	309QUI = 309QUI
309TEH = 309TEH - (309MER + 309ESP + 309JON)	309TEH = 309TEH + 309MER + 309ESP + 309JON

2.3.3 ILRP Data Cleaning

We extracted farm management data from the ILRP dataset. We calculated four predictor metrics for each practice: the number of farms that implemented each practice, the proportion of farms implementing each practice, the acreage used for each practice, and the proportional acreage used for each practice. These metrics were calculated at catchment and watershed scales to investigate whether upstream BMP implementation affected downstream response variables. All steps were performed manually in Microsoft Excel. The following list outlines the ILRP data cleaning process:

- Extracted relevant columns from the ILRP Annual Compliance Form (ACF) tab: AgRanchID, Trade_Secret, Nutrient_Implementation.
- Parsed the nutrient implementation column and reorganized so that the presence/absence of each BMP was recorded in its own unique column (practices P1 through P15).

- =IF(COUNTIF(Range of farm’s reported BMPs,"BMP_1")=present, Insert Present, leave blank)
- P1 through P12 represent the 12 recommended nutrient implementation BMPs, P13 represents a response of “Other”, P14 represents “No Practices Implemented”, and P15 represents “No Data Reported” which includes farms claiming trade secrets (Table 3).
- Calculated catchment–scale predictor metrics for the number of farms, proportion of farms, acreage of farms, and proportional acreage of farms that used each implementation practice.
- Calculated watershed–scale predictor metrics by summing the related catchment–scale predictor metrics for each implementation practice (Table 2).

Table 3. ILRP nutrient management practice implementation codes.

Nutrient Implementation	
Code	Description of Management Practice
P1	Evaluated how much fertilizer crop needs and timing of application
P2	Scheduled fertilizer applications to match crop requirements
P3	Measured nitrogen concentration in irrigation water and adjusted fertilizer nitrogen applications accordingly
P4	Measured soil nitrate or soil solution nitrate and adjusted fertilizer nitrogen applications accordingly
P5	Used precision techniques to place fertilizer in the root zone, to ensure crop uptake, with minimal runoff and deep percolation (e.g. fertigation)
P6	Measured nitrogen in plant tissue and adjusted fertilizer nitrogen applications
P7	Measured phosphorus in soil and adjusted fertilizer phosphorus applications
P8	Measured nitrogen and phosphorous content of applied manures and other organic amendments
P9	Mixed and loaded fertilizers on low runoff hazard sites (e.g. away from creeks and wells)
P10	Used urease inhibitors and/or nitrification inhibitors
P11	Modified crop rotation to use beneficial cover crops, deep rooted species, or perennials to utilize nitrogen.
P12	Used treatment systems to remove nitrogen from irrigation runoff or drainage water (e.g. wood chip bioreactor)
P13	Other, describe in Farm Plan and submit upon request
P14	None
P15	Answer left blank

2.3.4 Data matrix creation

Based on the NHD stream layer, catchment spatial relationships and input from technical advisors, we created a flowchart to inform our data preparation process. We joined all response variables and predictor metrics by CMP water quality monitoring site code (n = 22) (Fig. 5).

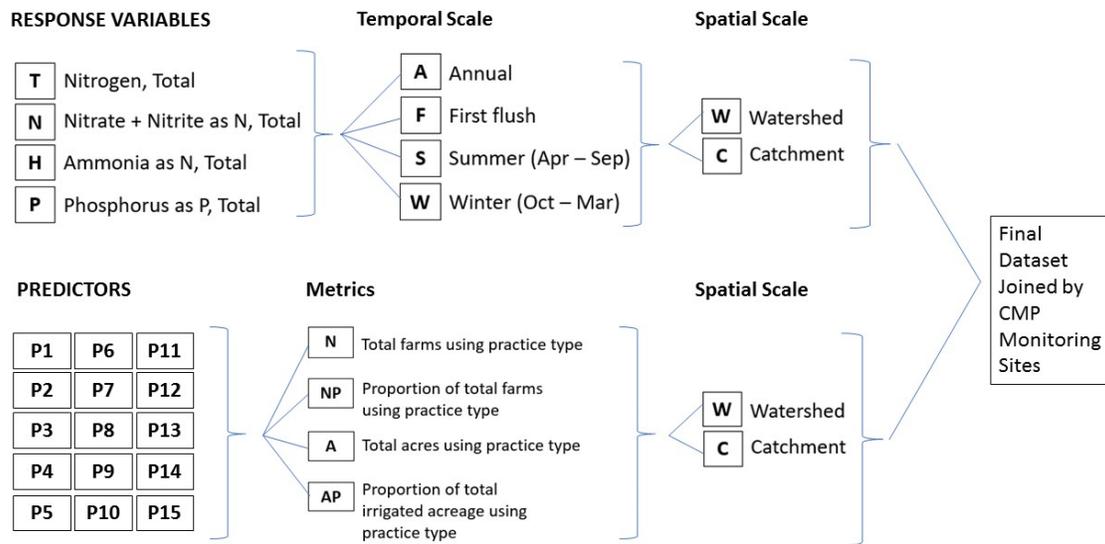


Figure 5. Data processing flow chart. This chart displays all data combinations used in the model. P1–P15 are codes that refer to on–farm management practices explained in Table 6.

2.4 Random forest statistical analysis

Random forest (RF) is a nonparametric empirical modeling technique that is based on an ensemble of decision trees fitted to a dataset (Breiman 2001). Fitting a single classification tree to one dataset can lead to models that are overfit (Cutler et al. 2007). To avoid overfitting, random forest models fit hundreds of classification trees using different subsets of data from the original dataset. The data not included in each classification tree is then used to estimate the accuracy of that tree. Random forest averages predictions across all trees within the forest to produce an ensemble prediction that are robust when applied to many ecological datasets (Fox et al. 2017).

We developed 64 separate random forest models, one for each of our eight analyte responses (four concentrations and four loads) at four temporal and two spatial scales using Rstudio (Rstudio Team 2014). Each initial model used 120 best management

practices (predictors) which were used to predict CMP water quality data (responses).

We used a coarse variable selection method to refine the top performing model for each analyte and to reduce the number of predictors in each model. We extracted the importance of each predictor from the variable selected model and ranked them based on their importance in predicting the response data. RF estimates importance by calculating the percent increase of mean squared error (%incMSE) of prediction when the model is built with the predictor of interest randomized. A positive %incMSE value indicates the model performance decreased when the predictor in question is randomized, indicating it is important in the model. Instead of building a model by examining each individual predictor's performance, because of the time limitations we used an abbreviated model selection technique. To select predictors, we took the top 30 predictors based on %incMSE from each of the best performing models and developed a new model. We then re-ranked the top 30 predictors based on their performance in the new model and analyzed the five most important in each model.

Finally, we summarized the average importance of individual predictors across all four models by adding the %incMSE of each predictor and dividing by the number of models.

3 Results

3.1 Model results

Nutrient management practices reported via ILRP explained up to 46% of the variation in nutrient concentrations, indicating that practices had measurable effects on water quality. Management practices related to measurement and evaluation of fertilizer needs were associated with lower nutrient concentrations in summer. However, practices that lowered nutrient concentrations in summer were associated with increased loads of ammonia during first flush.

The random forest analysis yielded six models with an R^2 greater than 0.2 (i.e., 20% of the variation within the response data explained by the predictor data) (Table 4).

Table 4. R^2 values of initial models using all 120 predictors. Red values indicate $R^2 > 0.2$ (20% of the response variation explained by predictors), and bolded values indicate the best performing models for each analyte.

Analyte	Spatial Scale	Temporal Scale			
		Annual	First Flush	Summer	Winter
Ammonia as Nitrogen	Catchment (mg/l)	0.06	-0.42	0.07	0.18
	Catchment Load	-0.07	-0.45	-0.18	0.17
	Watershed (mg/l)	-0.04	-0.21	-0.35	0.09
	Watershed Load	-0.57	0.21	-0.22	-0.19
Nitrate + Nitrite	Catchment (mg/l)	-0.52	-0.68	-0.32	-0.61
	Catchment Load	-0.33	-0.69	-0.21	-0.48
	Watershed (mg/l)	0.17	0.17	0.34	0.06
	Watershed Load	0.21	-0.16	0.19	0.07
Phosphorus	Catchment (mg/l)	0.20	-0.14	0.21	0.16
	Catchment Load	-0.39	-0.39	-0.25	-0.48
	Watershed (mg/l)	0.03	-0.12	-0.07	0.01
	Watershed Load	-0.34	-0.03	-0.22	-0.42
Total Nitrogen	Catchment (mg/l)	-0.37	-0.50	-0.33	-0.52
	Catchment Load	-0.36	-0.57	-0.03	-0.54
	Watershed (mg/l)	0.15	0.16	0.38	0.10
	Watershed Load	0.12	-0.20	0.21	0.12

We examined the top models for each analyte: total nitrate/summer/watershed, nitrate+nitrite/summer/watershed, ammonia/first flush/watershed load, and phosphorus/summer/catchment. We used the coarse variable selection method and re-ran these models using the top 30 predictors from the initial models. After the coarse variable selection, the R² improved for the top models (Table 5).

Table 5. Comparison of R² values of the top performing models for each analyte, before and after variable selection and model refinement.

Model name	Initial model	Refined model
HFWL	0.21	0.45
NSW	0.34	0.46
PSC	0.21	0.35
TSW	0.38	0.45

Results for individual models (Table 6):

Table 6. Summary of the top-performing refined analyte models and the top five predictors based on % increase mean squared error (%incMSE). Correlations were inferred from partial dependence plots (Figures 6 through 9).

Model	R2	Metric	Spatial scale	Correlation	%incMSE
TSW (total nitrate, summer, watershed)	0.45	Proportion of farms not reporting practice	Watershed	positive	6.67
		Proportion of total irrigated acres not reporting practice type	Watershed	positive	6.07
		Proportion of farms not reporting practice	Catchment	positive	5.85
		Proportion of farms evaluating fertilizer needs and application	Catchment	negative	5.15
		Proportion of farms evaluating fertilizer needs and application	Watershed	negative	4.93
PSC (phosphorus, summer, catchment)	0.35	Total acres measuring N and P content of manures and organic amendments	Watershed	negative	5.49
		Total acres scheduling fertilizer applications to match crop requirements	Watershed	negative	4.96
		Total farms not reporting practice	Watershed	negative	4.26
		Total farms scheduling fertilizer applications to match crop requirements	Watershed	negative	4.15
		Total acres modifying crop rotation to utilize nitrogen	Watershed	negative	3.91
NSW (nitrate/nitrite, summer, watershed)	0.46	Proportion of farms not reporting practice	Watershed	positive	6.64
		Proportion of total irrigated acres not reporting practice type	Watershed	positive	6.00
		Proportion of farms not reporting practice	Catchment	positive	5.94
		Proportion of total irrigated acres not reporting practice type	Catchment	positive	4.63
		Proportion of farms measuring N in plant tissue and adjusting N applied	Catchment	negative	2.91
HFWL (Ammonia as N, First Flush, Watershed Load)	0.45	Total farms measuring N in plant tissue and adjusting N applied	Catchment	positive	4.29
		Total farms measuring N in soils and adjusting N applied	Catchment	positive	4.11
		Total farms measuring P in soils and adjusting P applied	Catchment	positive	4.05
		Proportion of total irrigated acres measuring N and P content of manures and organic amendments	Catchment	positive	3.70
		Total farms scheduling fertilizer applications to match crop requirements	Catchment	positive	3.66

Total Nitrogen–Summer–Watershed scale model

Three of the top five predictor metrics in this model were farms that did not disclose their management practices and were also positively correlated to increased total nitrogen concentrations. The other top two predictor metrics were farms that evaluated their fertilizer quantity and timing needs, and were negatively correlated meaning decreased total nitrogen concentrations. (Fig. 6).

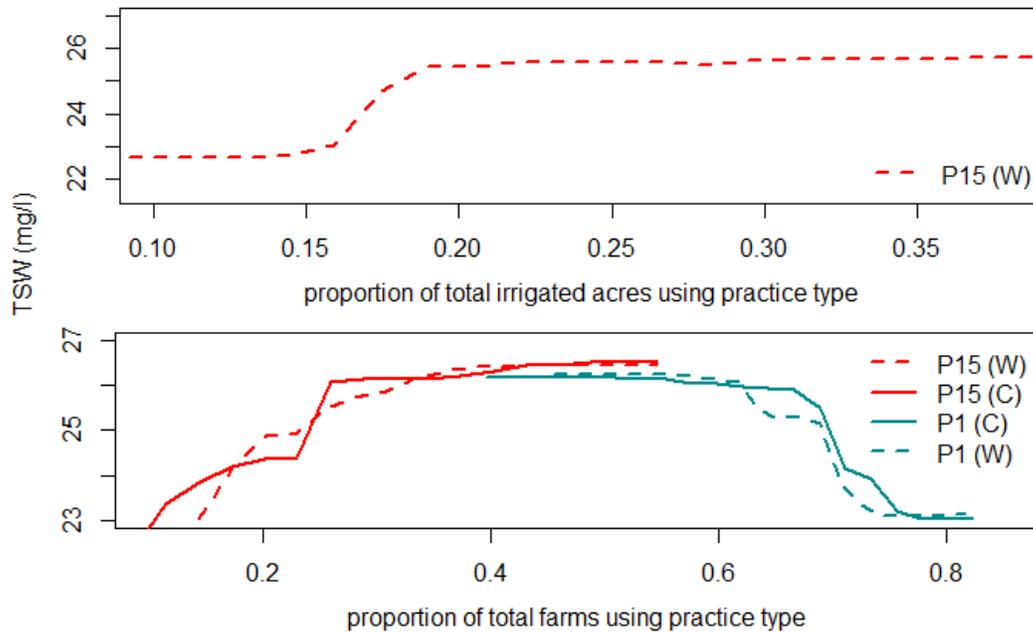


Figure 6. Partial dependence plots of the top five predictors in the TSW model (Total Nitrogen, summer, watershed). W or C refers to watershed or catchment spatial scale of the predictor, respectively.

Nitrate+Nitrite-Summer-Watershed scale model

Similarly, the top four predictor metrics in this model were farms that did not report what management practices they used. These metrics were also positively correlated to higher concentrations of nitrate+nitrite. The final metric was farms that measured plant tissues for nitrogen and adjusted fertilizer application, and was the only negatively correlated metric of the top five (Fig. 7).

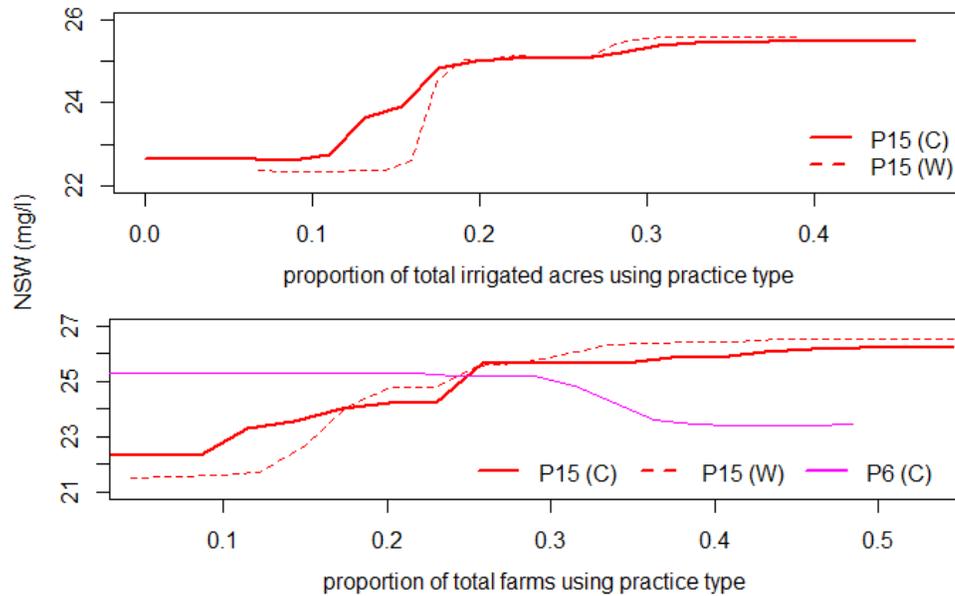


Figure 7. Partial dependence plots of the top five predictors in the NSW model (nitrate/nitrite, summer, watershed). W or C refers to watershed or catchment spatial scale of the predictor, respectively.

Ammonia-First Flush-Watershed scale load model

The top five predictor metrics in this model were positively correlated with increased ammonia load. All metrics were associated with farm management practices that measured and/or adjusted fertilized application: measured nitrogen in plant tissue and adjusted fertilizer application, measured soil nitrate and adjusted fertilizer application, measured soil phosphorus and adjusted fertilizer application, measured nitrogen and phosphorus in applied manure/organic amendments, and schedule fertilized applications to match crop requirements (Fig. 8).

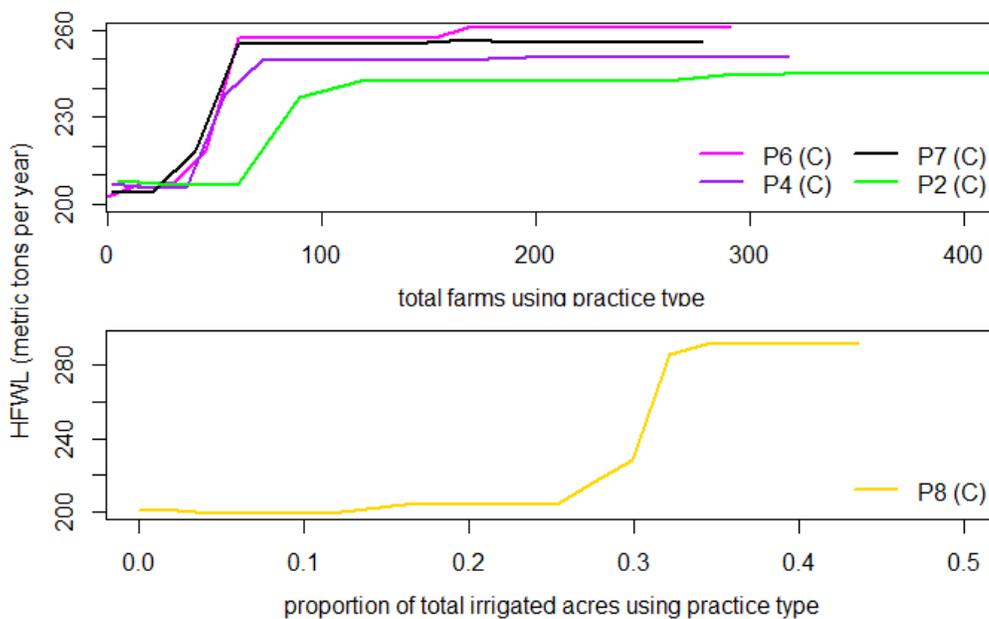


Figure 8. Partial dependence plots of the top five predictors in the HFWL model (NH₄, first flush, watershed load). W or C refers to watershed or catchment spatial scale of the predictor, respectively.

Phosphorus–Summer–Catchment scale model

The top five predictor metrics in this model were negatively correlated, meaning lower phosphorus concentrations: measured nitrogen and phosphorus in applied manure/organic amendments, scheduled fertilizer application to match crop requirements, did not disclose practices used, modified crop rotation to use beneficial cover crops and nitrogen utilizing plants. Of the practices disclosed, most of them appear to be lower-cost forms of management, and are otherwise not closely related (Fig. 9).

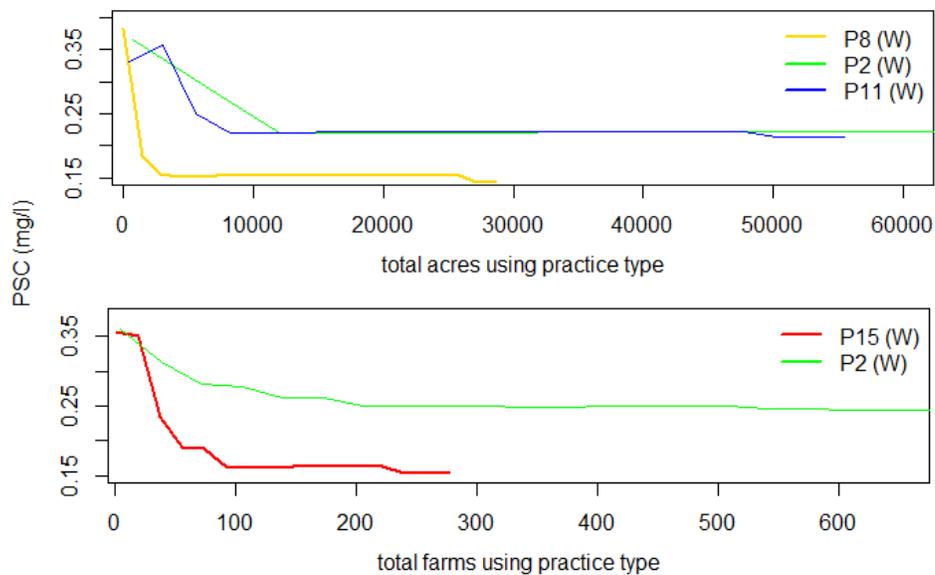


Figure 9. Partial dependence plots of the top 5 predictors in the PSC model (phosphorus, summer, catchment). W or C refers to watershed or catchment spatial scale of the predictor, respectively.

Across the top four refined models, the top four most important predictors were: classifying practices as trade secrets (P15), evaluating fertilizer needs and timing (P1), scheduling fertilizer to match crop requirements (P2) and measuring nitrogen and phosphorus content of applied organics (P8) across the top models for each analyte (Table 7)).

Table 7. Summary of best management practices (predictors) from the top performing models ranked by the weighted sum % increase mean squared error (%incMSE).

	Best Management Practices							
	P15	P1	P2	P8	P6	P4	P7	P11
Total %incMSE across top four refined models	46.06	10.08	12.77	9.19	7.19	4.11	4.05	3.91
Weighted Sum of %incMSE	11.52	2.52	3.19	2.30	1.80	1.03	1.01	0.98
Ranked predictor importance	1	2	3	4	5	6	7	8

We used these models, predictors and responses to address the questions listed in section 1.3:

3.2 Are there particular subwatersheds that are “hot spots” for high nutrient concentrations?

Water quality monitoring sites showed highest surface water nutrient concentrations in the lower Salinas watershed where most of the irrigation runoff is collected by tile drains and piped to irrigation ditches (sites 309BLA, 309TEH, 309MER, 309ASB, 309JON).

Total nitrogen as N and nitrate + nitrite as N analyte results exhibited similar patterns both spatially and temporally with highest concentrations recorded during the summer months (Figures 10 and 11). Nitrate + nitrite concentrations were especially high in the summer of 2016 in the lower Salinas watershed (23–78 mg/L). Ammonia concentrations were generally low across the Pajaro watershed and higher in the Salinas watershed across most temporal scales (Fig. 12). Total phosphorus as P concentrations were highest in the winter months across both the Salinas and Pajaro watersheds (0.59–2.98 mg/L and 0.59–1.06 mg/L, respectively; Fig. 13).

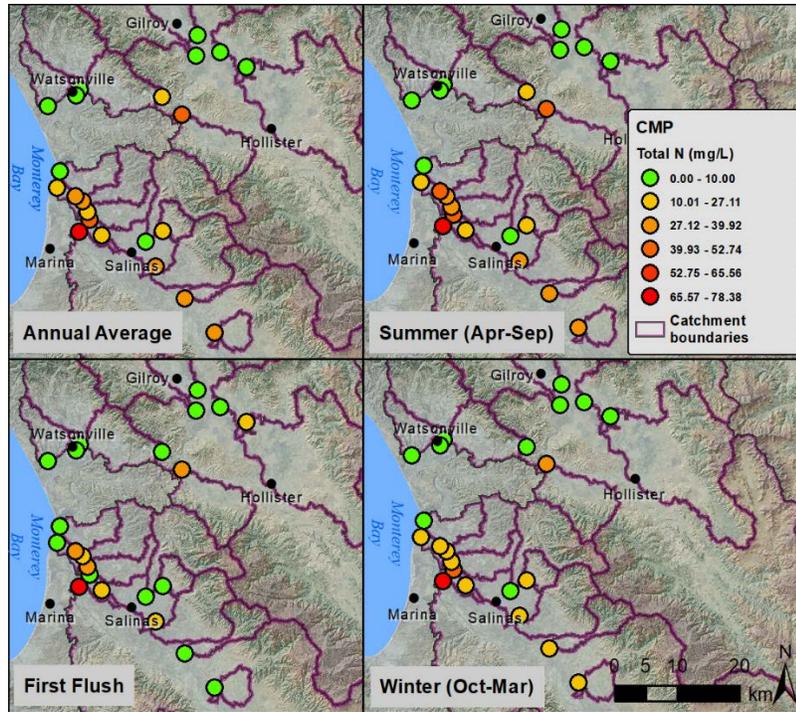


Figure 10. 2016 average total nitrogen concentrations as N at 22 CMP water quality monitoring sites over four temporal scales.

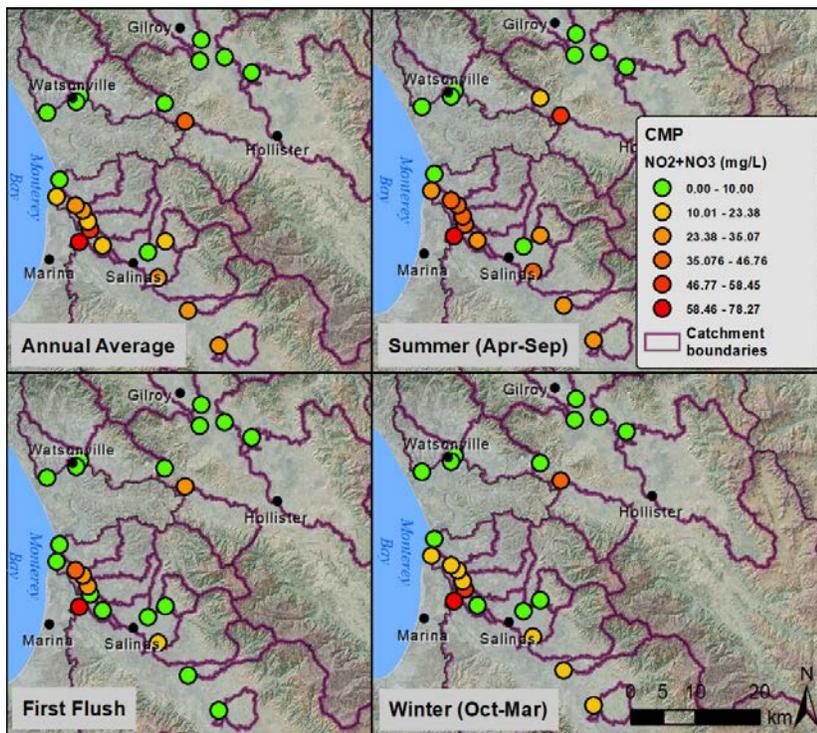


Figure 11. 2016 average nitrate + nitrite concentrations as N at 22 CMP water quality monitoring sites over four temporal scales.

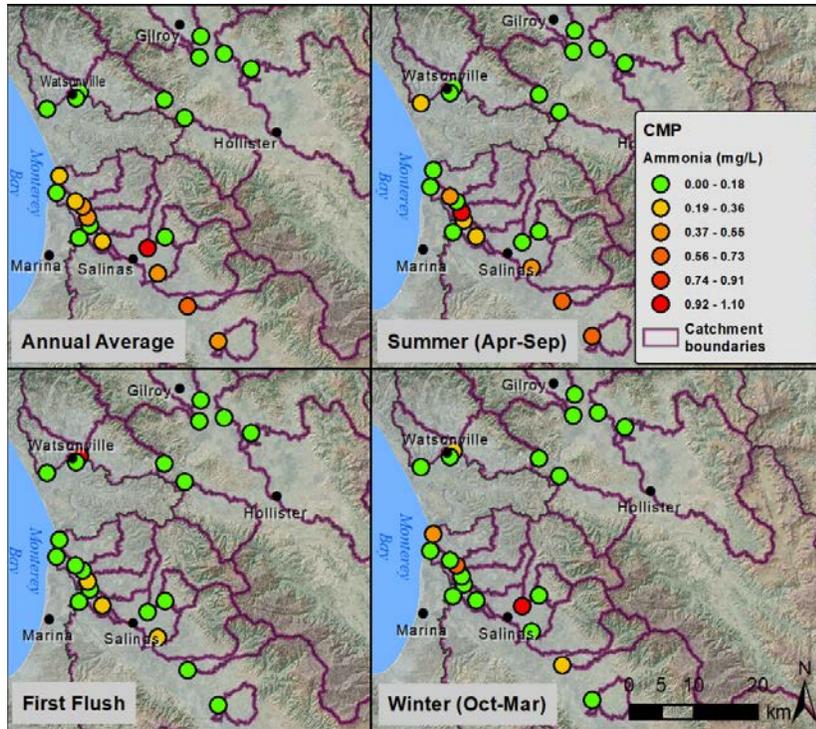


Figure 12. 2016 average ammonia concentrations at 22 CMP water quality monitoring sites over four temporal scales.

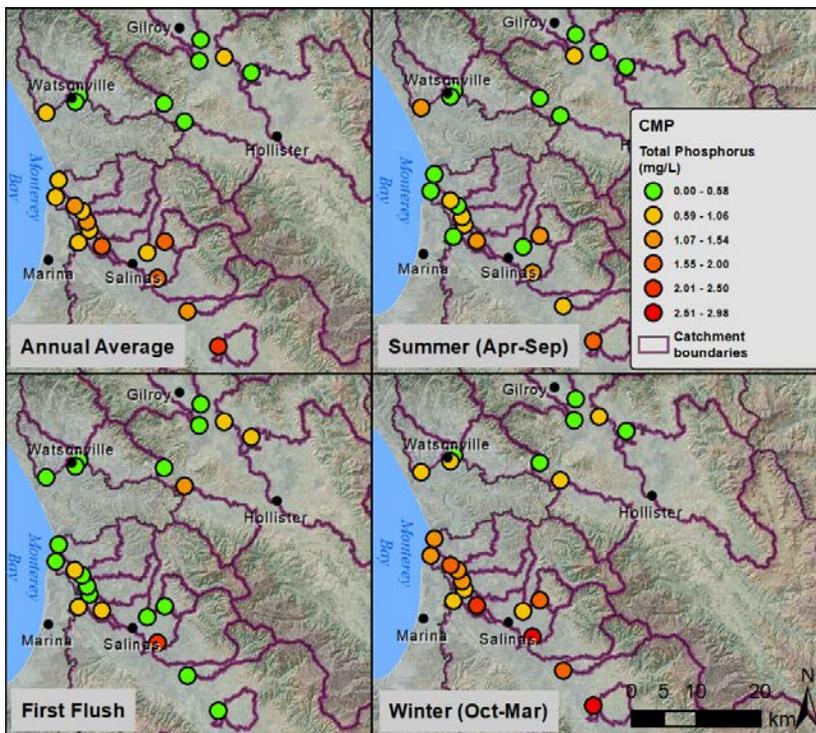


Figure 13. 2016 average phosphorus concentrations as P at 22 CMP water quality monitoring sites over four temporal scales.

3.3 Is there a relationship between particular management practices and CMP water quality results?

We examined the top four predictors (classifying practices a trade secret (P15), evaluating fertilizer needs and timing (P1), scheduling fertilizer to match crop requirements (P2) and measuring nitrogen and phosphorus content of applied organics (P8)) across the top models for each analyte (Table 7)).

Farms classifying their practices as a trade secret were predictors of all four analytes, and had both positive and negative correlations with nutrient concentrations depending on the model. Increases in farms not reporting their management practices were strongly correlated with higher levels of total nitrogen and nitrate+nitrate in summer. Non-reporting was also negatively correlated with phosphorus levels in summer

Evaluating how much fertilizer crop needs and timing of application (P1) was correlated with reduced total nitrate concentration in summer.

The scheduling of fertilizer applications to match crop requirements (P2) resulted in reduced phosphorus concentrations in summer, however was positively correlated with ammonia concentration during the first flush period.

Measuring nitrogen and phosphorus content of applied manures and other organic amendments (P8) was negatively correlated with phosphorus concentrations in summer, but positively correlated with ammonia during first flush.

3.4 Are there particular geographies where practice adoption is low?

Management practice adoption rates and distribution around the region varies among the top four predictor practices (Fig. 14). Generally, high proportions of acreage are evaluating fertilizer needs and timing (P1). Most catchments have more than 50% of acreage using P1, save for one catchment in the area immediately surrounding the city of Salinas (309TEH). 309TEH also had the lowest proportion of acreage scheduling fertilizer to match crop requirements (P2) at 45%. Acreage proportions measuring nitrogen and phosphorus content of applied organics (P8) were generally low, and 309GAB and 305FRA had virtually no acreage where this practice was utilized. Farms that did not report their management practices (P15) were more concentrated in the southwest of the study area and around the City of Salinas. P15 acreage proportion ranged from 0 (305CAN) to 46% (309TEH).

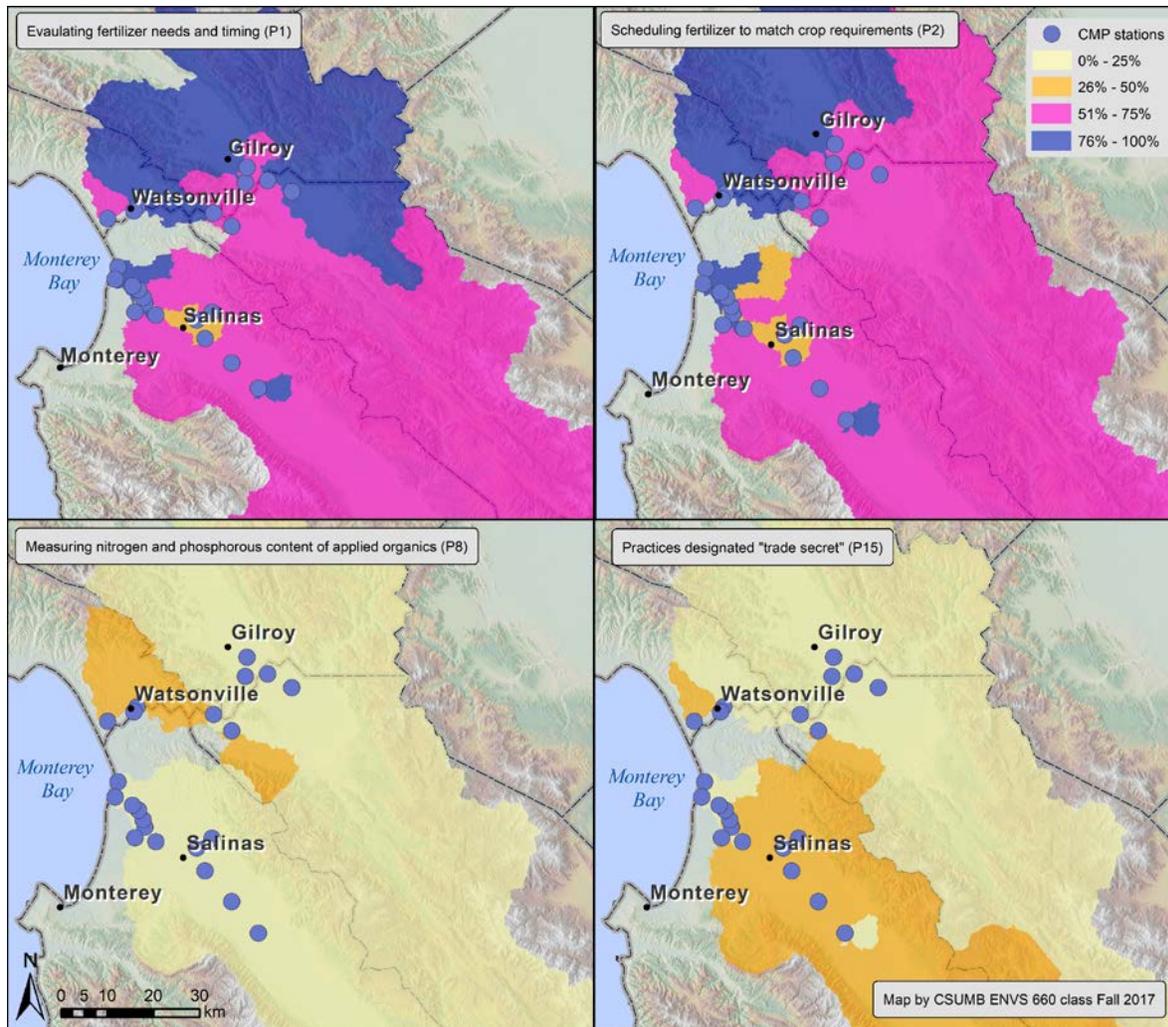


Figure 14. Management practice adoption rate and distribution for top overall predictors of nutrient concentration by proportion of irrigated acreage per catchment.

3.5 Are the ILRP BMPs the best management practices for improving water quality?

Considering the magnitude of the dataset, the limited scope of this study and the complexities of the model, it would be premature to draw such conclusions at this point in the process. Expanded analysis and in-depth consideration of management practices highlighted here may offer further insight into this important question. The results presented here only offer preliminary results.

4 Discussion

Correlations between management practices and nutrient concentrations in this analysis do not provide explicit explanations of causation. It is possible that the management practices found to be highly correlated with improved water quality are also correlated with other practices not accounted for in these models.

While the ammonia model (first flush, watershed load) explained the highest percentage of variation in water quality, the predictors in this model were all positively correlated with nutrient concentrations, opposite of what we observed in the other models. There are several possibilities for these results. Practices that reduce total nitrogen and nitrate can cause higher levels of ammonia. Organic fertilizer, applied to crops to reduce nitrogen and nitrate, is related to the practice of measuring nitrogen and phosphorus content of applied manure and organic amendments (P8). Similarly, adjusting fertilizer nitrogen and phosphorus (P6 and P7) in the late summer to reduce total nitrogen and nitrate may cause higher nutrient runoff during the first flush. Increased ammonia can also be correlated with other, unaccounted for factors like tillage practices or soil pH. A larger, more comprehensive data set is required to validate our results.

4.1 Study limitations

The issue of agriculture's influence on surface water quality in California is highly complex and contentious, as evidenced by multiple agricultural orders culminating in limited success and differing perspectives among stakeholders. Constrained by large data sets and a restrictive timeframe, we focused on the correlation between nutrient management practices and surface water quality. ILRP and CMP datasets include several other parameters and variables that should be considered for a more complete understanding of which practices farmers can adopt to improve overall water quality on the Central Coast. As we analyzed our results, we determined model and scope limitations, unaccounted for hydrologic intricacies requiring ground-truthing, and the need to incorporate additional practices and years of data.

Scope limitations

We narrowed the scope of our research based on the five-week time limit. We confined our predictor variables to farm nutrient implementation practices, assuming that the application and management of fertilizers has a significant effect on water quality.

The irrigation practices and the specific hydrologic conditions of different farms and geographic areas also have a profound effect on surface water quality, especially with regard to nitrate. Sprinkler-fertigated, tile-drained areas with perched groundwater tables and channels with baseflow and/or tidal influence can all influence surface water quality. An understanding of on-farm practices and hydrologic connections requires a knowledge of catchment-specific geology and hydrology as well as intensive ground-truthing, which is unrealistic for a study of this scale. These hydrologic intricacies do make a difference on the individual farm scale, but modeling overall relationships between farm management practices and water quality over several watersheds is not necessarily dependent on them.

Model assumptions and limitations

We used random forest to model which nutrient implementation BMPs best predict water quality analyte concentrations on a temporal and spatial scale, making several decisions that limited the interpretation of our results. We used farms from all tiers in this analysis. Tier 3 farms – those that may have the most impact on water quality – were treated the same as Tier 1 and Tier 2 farms. We did not use crop type as a predictor of water quality. Certain crops, like strawberries or lettuce, are associated with increased fertilizer and water use, and therefore may have a greater effect on water quality. Finally, the distance between individual farms and CMP site is not accounted for. BMPs of farms several kilometers away from their corresponding CMP sites, for example, were weighted the same as BMPs implemented on farms just meters away from the site. We isolated the effect of nutrient implementation BMPs by eliminating these other factors from our model, though an exhaustive analysis should take these variables into account.

The random forest model was limited by sample size and controlled by the predictors and variable selection process. The sample size was $n = 22$, based on the 22 CMP sites in our study area. Incorporating additional water quality information, both from other CMP locations and any other available water quality data, raising n to at least 40, would create a more robust model. Including data from watersheds in other areas would help show widespread trends between management practices and water quality. We sorted the predictors by percent increase MSE, arbitrarily choosing the top 30 for additional refinement. Following this refinement, we chose the top 5 to 7 predictors to plot partial dependence for further analysis. Predictor selection was therefore somewhat arbitrary and can be refined in the future with a formal variable selection process.

CMP sites

Mapping of 2016 average nutrient concentrations at CMP sites showed that the poorest surface water quality across the study area exists in the lower Salinas watershed. These waterways generally transport agricultural runoff from fields with tile drain systems, gathering water with high nutrient concentrations. Agricultural land that is not equipped with tile drains could be the source of high nutrient contamination too, as surface agricultural water infiltrates the soil and moves laterally into nearby creeks. Therefore, although surface water nutrient concentrations register as highest in the lower Salinas watershed, other locations that sporadically show up as “hotspots” should be looked at closely when developing management plans for different subwatersheds.

The random forest analysis encompassed water quality data from 22 CMP sites located in the Salinas and Pajaro watersheds. We refined our analysis by using only those CMP sites in the lower Salinas watershed due to more immediate regional significance. This refinement led to the exclusion of five CMP sites in the upper Salinas valley, forming a long, large catchment associated with 309BLA that contained the majority of, mainly Tier 1, farms in the valley. Concerned that the 309BLA catchment skewed our results, we re-ran our model without it, and found R^2 values increased for the top three models and that several top predictors changed. However, we decided to continue the analysis using this large catchment to include more farms in the model. Future studies should incorporate additional CMP sites to create smaller catchments in the area, or weight management practices by distance from CMP site.

We initially based CMP site catchment extents on a GIS analysis, determining flow direction and accumulation in an unaltered system based on a 2013 USGS DEM. The system (Salinas and Pajaro watersheds) is, however, significantly altered by extensive agriculture and urbanization. We adjusted CMP catchment flow progression based on field reports of hydrologic connectivity (S. Lopez, pers. comm., Dec. 7, 2017). This contribution rendered our analysis more accurate, and indicates that more extensive analysis, field study and consultation with regional experts is necessary to account for the dynamic relationships between catchments.

ILRP

This dataset did not include each farm’s total acreage for most of the data points. Instead, we used the reported irrigated acreage to assess the impact range of nutrient implementation BMPs at the catchment and watershed levels.

4.2 Areas for Future Study

This is an initial study assessing correlations between agricultural fertilizer BMPs and surface water quality. Model results inform stakeholders and policy makers on possible correlations between farm management practices and surface water nutrient concentrations on temporal and spatial scales, while alerting them to possible limitations of these datasets. We recommend further study to include:

- Further analysis on a refined list of best management practices and suites of associated practices
- Addition of remaining 305 and 309 CMP sites and/or additional water quality data to increase model strength
- Addition of crop type, farm tier and irrigation BMPs into models
- Ground truthing of farm drainage, including irrigation canals, runoff ditches, pumping and tailwater to determine hydrologically correct catchment boundaries and interactions
- Include irrigation well nitrate concentration monitoring data as a predictor of surface water pollution trends
- Relate discharge data and irrigation practices in a similar analysis to better understand how practices are affecting water quality
- Examine results over a longer temporal period to determine if BMP benefits occur immediately or over a longer time span
- Calculation and modeling of nutrient loads to determine the total amount of nutrients entering the marine sanctuary

5 Conclusion

We found that management practices explained up to 46% of the variation in nutrient concentration, with practices P15, P1, P2, and P8 having the largest contributions. Variations in water quality are related to farm practices, although further identification of important and unimportant farm practices is required to improve accuracy of predictive models. Results of our model do not imply causation, but rather show a correlation between management practices P15, P1, P2 and P8 and significant differences in analyte concentrations. P3, P5, P9, P10, P11 and P12 management practices showed little to no correlation with water quality. However, further investigation is required to determine if inclusion of these practices could potentially improve our model accuracy and be effective at reducing nutrient load.

Because our analysis shows that farms that indicated their practices were a “trade secret” were a top predictor for water quality, both negatively and positively correlated, the actual effect of any practices they may be implementing is inconclusive. Future research should address this data gap, as well as methods to improve the strength of the model.

This analysis aimed to illuminate general trends, not specific results from individual farms. The methodology used in this study can serve as a trial that informs future research needs, and should not be taken out of context or over-interpreted.

7 References

Bilginsoy A, Ruby A. 2017. Annual water quality report: 2016 city of Capitola first flush. Coastal Watershed Council. [Internet]. [cited 2017 November 23]. Available from: <https://coastal-watershed.org/wp-content/uploads/2017/02/2016-City-of-Capitola-First-Flush-Report-Final.pdf>.

Breiman L. 2001. Random Forest. Machine Learning 45(1): 5–32.

[CCRWQCB] Central Coast Regional Water Quality Control Board. 2011. Order No. R3-2011-0006, Appendix G: Water quality conditions in the Central Coast Region related to agricultural discharges. [Internet]. [cited 2017 November 22]. Available from: https://www.waterboards.ca.gov/centralcoast/board_info/agendas/2011/march/Item_14/14_att7.pdf

[CCRWQCB] Central Coast Regional Water Quality Control Board. 2012. Order No. R3-2012-0002: Conditional Waiver of Waste Discharge Requirements for Discharges from Irrigated Lands. [Internet]. [cited 2017 November 22]. Available from: https://www.waterboards.ca.gov/centralcoast/water_issues/programs/ag_waivers/docs/ag_order/final_agorder_atta_032612.pdf

[CCRWQCB] Central Coast Regional Water Quality Control Board. 2013. Total Maximum daily loads for nitrogen compounds and orthophosphate for the Lower Salinas River and Reclamation Canal, and the Moro Cojo Slough subwatershed, Monterey County, CA. Final Project Report. Item No. 9 Attachment 2 for the January 31–February 1, 2013 Water Board Meeting. [Internet]. [cited 2017 November 11]. Available from: https://www.waterboards.ca.gov/rwqcb3/board_info/agendas/2013/jan/Item_9/item_9_att2.pdf

[CCRWQCB] Central Coast Regional Water Quality Control Board. 2017. Order No. R3-2012-0011: Conditional Waiver of Waste Discharge Requirements for Discharges from Irrigated Lands. [Internet]. [cited 2017 November 22]. Available from: https://www.waterboards.ca.gov/centralcoast/water_issues/programs/ag_waivers/docs/ag_order3/ag_order3.0_approved.pdf

Cutler DR, Edwards TC Jr., Beard KH, Cutler A, Hess KT, Gibson J, and Lawler JJ. 2007. Random Forests for Classification in Ecology. Ecology 88(11): 2783–2792.

Fox EW, Hill RA, Leibowitz SG, Olsen AR, Thornbrugh DJ, Weber MH. 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environmental Monitoring and Assessment* 189(316):1–20.

Harter R, Lund JR. 2012. Addressing nitrate in California's drinking water, with a focus on Tulare Lake Basin and Salinas Valley groundwater. Report for the State Water Resources Control Board. [Internet]. [cited 2017 November 11]. Available from: <http://groundwaternitrate.ucdavis.edu/files/138956.pdf>

[MCAC] Monterey County Agricultural Commissioner. 2016. Monterey County 2016 crop report. [Internet]. [cited 2017 November 11]. Available from: <http://www.co.monterey.ca.us/home/showdocument?id=27601>

RStudio Team. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Available from: <http://www.rstudio.com/>

[SCCAC] Santa Cruz County Agricultural Commissioner. 2016. County of Santa Cruz 2016 crop report. [Internet]. [cited 2017 November 11]. Available from: <http://www.agdept.com/Portals/10/pdf/2016%20Crop%20Report%20for%20Webpage.pdf>

Scholin CA, Gulland F, Doucette GJ, Benson S, Busman M, Chavez FP, Cordaro J, DeLong R, De Vogelaere A, Harvey J, Haulena M. 2000. Mortality of sea lions along the central California coast linked to a toxic diatom bloom. *Nature* 403(6765):80–4.

Starner K, White J, Spurlock F, Kelley K. 2006. Pyrethroid Insecticides in California Surface Waters and Bed Sediments: Concentrations and Estimated Toxicities. California Department of Pesticide Regulation. [Internet]. [cited 2017 November 22]. Available from: http://www.ccamp.org/ccamp/documents/starner_pyreth06.pdf

Work TM, Barr B, Beale AM, Fritz L, Quilliam MA, Wright JL. 1993. Epidemiology of domoic acid poisoning in brown pelicans (*Pelecanus occidentalis*) and Brandt's cormorants (*Phalacrocorax penicillatus*) in California. *Journal of Zoo and Wildlife Medicine* 1:54–62.

8 Appendix

```
660 Random Forest Modeling Code.
#clean space
rm(list=ls())
graphics.off()
# set working directory
setwd("E:\\ENVS660\\John\\RFmodel\\171209")
library(randomForest)
##Let's bring in the Predictor data and have a look at that:
#preds <- read.csv ("ILRP_Predictor_METRICS_WATERSHEDS_FINAL_171120.csv", header=TRUE, #sep=",")
#upload predictor and response data
dat <- read.csv("RespPred_All_171209.csv", header=TRUE, sep=",")
#subset the predictors
preds <- subset(dat[,3:122])
#extract model names for R2 output table
model_nams <- colnames(dat[123:186])
##### Create Random Forest Function #####
RF_model <- function(df,model_nams,na.rm = FALSE, ...){
  preds      <- subset(df[,3:122])
  R2         <- c(NA)
  results    <- data.frame(model_nams,R2)
#Create the loop that uses the Random Forest function created above, and inputs the results into
#a table with model names called "PsuedoR2"

for (i in seq_along(model_nams)) {
  require(randomForest)
  n <- model_nams[i]
  k <- df[n]
  m <- na.omit(cbind(k,preds))
  rf <- randomForest(m[,1] ~ m[,2]+ m[,3]+ m[,4]+ m[,5]+ m[,6]+ m[,7]+ m[,8]+ m[,9]+ m[,10]+
m[,11]+ m[,12]+ m[,13]+ m[,14]+ m[,15]+ m[,16]+ m[,17]+ m[,18]+ m[,19]+ m[,20]+ m[,21]+ m[,22]+ m[,23]+
m[,24]+ m[,25]+ m[,26]+ m[,27]+ m[,28]+ m[,29]+ m[,30]+ m[,31]+ m[,32]+ m[,33]+ m[,34]+ m[,35]+
m[,36]+ m[,37]+ m[,38]+ m[,39]+ m[,40]+ m[,41]+ m[,42]+ m[,43]+ m[,44]+ m[,45]+ m[,46]+ m[,47]+
m[,48]+ m[,49]+ m[,50]+ m[,51]+ m[,52]+ m[,53]+ m[,54]+ m[,55]+ m[,56]+ m[,57]+ m[,58]+ m[,59]+
m[,60]+ m[,61]+ m[,62]+ m[,63]+ m[,64]+ m[,65]+ m[,66]+ m[,67]+ m[,68]+ m[,69]+ m[,70]+ m[,71]+
m[,72]+ m[,73]+ m[,74]+ m[,75]+ m[,76]+ m[,77]+ m[,78]+ m[,79]+ m[,80]+ m[,81]+ m[,82]+ m[,83]+
m[,84]+ m[,85]+ m[,86]+ m[,87]+ m[,88]+ m[,89]+ m[,90]+ m[,91]+ m[,92]+ m[,93]+ m[,94]+ m[,95]+
m[,96]+ m[,97]+ m[,98]+ m[,99]+ m[,100]+ m[,101]+ m[,102]+ m[,103]+ m[,104]+ m[,105]+
m[,106]+ m[,107]+ m[,108]+ m[,109]+ m[,110]+ m[,111]+ m[,112]+ m[,113]+ m[,114]+ m[,115]+
m[,116]+ m[,117]+ m[,118]+ m[,119]+ m[,120]+ m[,121], data = m, importance=T)
  RSQ <- rf$rsq
  q <- RSQ[length(RSQ)]
  results[i,2] <- q
}
print(results)
}
PsuedoR2 <-RF_model(dat,model_nams)
#write.csv(PsuedoR2,"RF_r2_171209.csv")
```