ABSTRACT

HYDROLOGIC CLASSIFICATION OF ALTERED RIVER IN CALIFORNIA

by

Daisy Guitron

University of California, Davis, 2020
Under the Supervision of Professor Samuel Sandoval Solis

Required environmental flows can mimic components of the natural flow regime, to support the lifecycle needs of fish, amphibians, riparian vegetation, birds, and wildlife. Human development has caused changes in flow regimes, disconnected habitats (upstream, floodplain), increased water temperature, brought in invasive species, increased consumptive use (salinity changes), changing magnitude, timing, frequency, and rate of change in flows. Natural streamflow patterns vary worldwide. In California, riverine ecosystems adapt to a Mediterranean climate: floods in wet winters, snowmelt flows in spring and low flows in summer. Humans have modified the natural river flow patterns in California by storing water during winter and releasing during summer and diverting water from streams. Resulting alterations to the natural flow regimes have degraded riverine ecosystems. Both intense climatic variability and profoundly altered rivers increase the importance of understanding the diversity of streamflow patterns. The present study quantifies the human alteration on flow regimes in California by categorizing impaired flow regime classes from human alteration. A systematic framework based on statistical approaches characterizes and predicts hydrologic class for impaired flows in California is presented. A total of 813 gauges out of 1,810 were deemed non-reference gauges after filtering by location based on ArcGIS, visual inspection of every stream gage, drainage basin from recent high-resolution imagery, and screen out of gages with insufficient daily streamflow data. The statistical methods used to determine the impaired streamflow classification are Pearson Correlation, Non-metric Multidimensional Scaling (NMDS), Hierarchical Clustering using Ward's Algorithm, Tukey's box and whisker plots, and Classification and Regression Tree Analysis (CART). After this initial classification, a heuristic analysis was performed to verify that gages downstream of reservoirs were adequately classified and reduced the mixed alteration class. This methodology resulted in nine altered flow classes representing distinct flow sources and hydrologic characteristics. The nine impaired streamflow classes reflect the land use impairment in California: *urban* (high, low, and medium density), *agriculture* (high, medium, and low crop density), *dams and reservoirs*, *forestland and land use change* (deforestation, logging, fires, cattle stocking and grazing, cannabis production), and *mixed alteration*. The impaired streamflow classification assesses, locates and evaluates the extent of each type of impairment and can be used to estimate how the current impaired flow regime diverges from the natural flow regime.

Dedication

This thesis is dedicated to my parents.

For their endless love, support, and their constant "Echale Ganas"

Table of Figures

Table of Tables

Acknowledgments

Table of Contents

1. Introduction

Rivers are under immense pressure as growing populations demand more from rivers. Alterations of flow are a primary contributor for the degradation of river ecosystems and native species reduction. Flow alterations are influenced by water management and land activities, including diversions, reservoir construction, groundwater withdrawals and depletion, agricultural runoff, impervious surfaces, such as roads. The modification of reservoir operations to control the timing and magnitude of flow releases for environmental benefits is an emerging approach for mitigating the negative ecological impacts of dams while preserving essential water management functions (Richter and Thomas 2007; Arthington 2012; Richter et al. 2003; Ai et al. 2013; Lane et al. 2014). A native riverine ecosystem composition is tightly linked to natural hydrologic variability. Freshwater is an escalating conflict, where a growing realization that human society must modify its behavior to ensure long term ecological vitality of riverine ecosystems. A goal of ecosystem management is to sustain ecosystem integrity by protecting native biodiversity. Faced with the complexity inherent in natural systems, achieving that goal will require that resource managers explicitly describe desired ecosystem structure, function, and variability; characterize differences between current and desired conditions; define ecologically meaningful and measurable indicators that can mark progress toward ecosystem management and restoration goals and incorporate adaptive strategies into resource management plans (Swanson, 2005). By managing river flows for water supplies and power generation, water management agencies have inadvertently caused considerable degradation of riverine ecosystems and associated biodiversity (Richter & Richter 2000).

In much of the world, growing populations are reducing available freshwater supplies. This quest for ecological sustainability typically centers on managing human uses of water such that enough water is available for use by future generations. There is a considerable need for new approaches in meeting human needs for water while conserving the riverine ecosystems ecological integrity. Ecological degradation has been an unintended consequence of water management as a result of a lack of understanding of water flows necessary to sustain freshwater ecosystems.

Variability counteracts the main goals of water resource management. Conventional water management has sought to dampen the natural variability of river flows to attain controlled and dependable water supplies for various sources such as domestic and industrial uses, irrigation, navigation, and hydropower and to moderate extreme water conditions such as floods and droughts. When alteration to the natural flow becomes excessive, variability in river flows changes is expected in the physical, chemical, and biological conditions and functions of natural freshwater ecosystems. This extreme change has a cost that is both high to biodiversity and society. The ultimate challenge is designing and implementing a water management program that stores and diverts water for human purposes in a manner that does not cause affected ecosystems to degrade.

This implies that a limit exists for the amount of withdrawn water from a river, and a limit in the degree to which the shape of a river, and a restriction in the degree to which the shape of the river's natural flow patterns are altered. The ecosystem's requirements for water define limits like these. Human involvement that exceeds these limits will compromise the ecological integrity of the affected ecosystems, resulting in loss of native species and ecosystem products and services for society (Ritcher et al., 2003). But the river degradation has not been quantitatively estimated; this is important for understanding where and how flow regimes have been altered. Stressing the critical need to classify and characterize the type of alteration that is occurring. River classification serves various essential purposes; one is by assigning rivers or river segments to a particular kind, relationships between ecological metric and flow alterations can be developed for an entire river type based on data obtained from a limited set of rivers of that type within the region of California

The central scientific question of this study is to identify patterns of streamflow alteration using publicly available data. Thus, the overall goal of this research is to develop a method for determining a streamflow altered classification. The state of California is used as a case of study. The central hypothesis is that it is possible to develop a method for classifying streamflow gages using publicly available alteration indicators (land use and infrastructure disturbance indicators). The specific objectives of this research are: (1) identify the non-reference streamflow gages, (2) classify gages by their impairment for the entire state of California, and (3) spatially predict the altered streamflow classes throughout the river network. Figure 1 illustrates a breakdown of the research goals with corresponding steps. This is further explained in great detail in section 4 of the methodology.

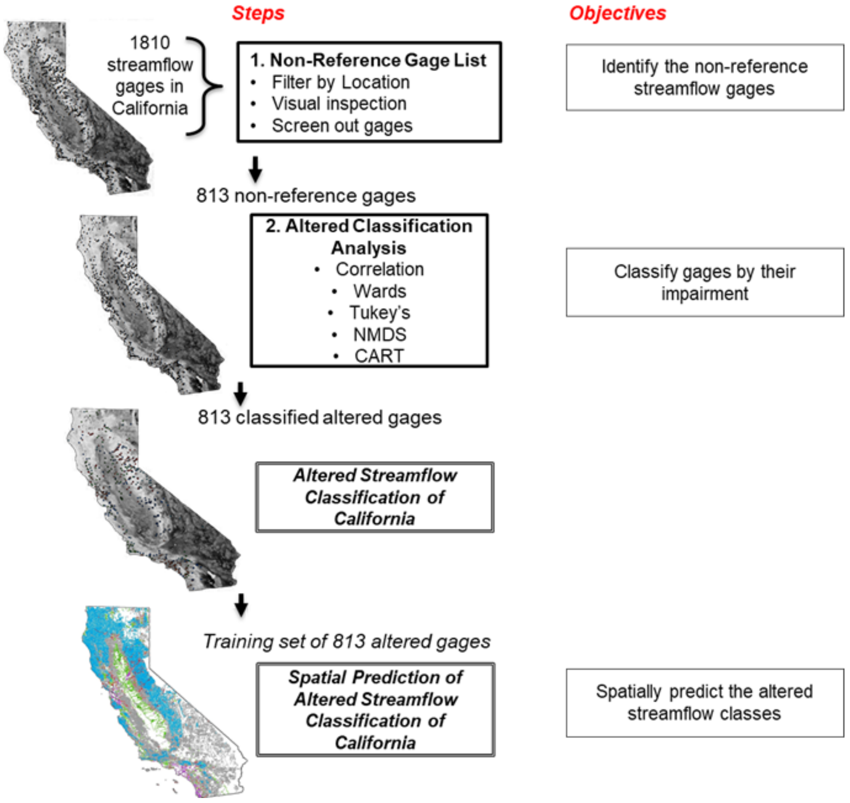


Figure 1.Process in achieving the streamflow gauge impaired classification of California

This work is innovative in terms of quantitatively quantifying and spatially predicting patterns of streamflow alteration at the gage and 200-m reach scale, specifically for the state of California. This research study defines a method (using available data throughout the U.S) that can be applied in other regions. Given the long history of water use in California, this study allows us to identify the types of streamflow alteration, and in conjunction with other variables of low, medium and high alteration, it can provide a road map for prioritization. This study can be used in combination with the natural streamflow classification to assess the patterns of alteration throughout the state of California. Information from this study can be used for restoration purposes, by identifying and locating the extent and type of alteration, or where there may or may not be good ecosystem health. This research study builds upon site-specific and spatial prediction analysis of databases that will support the development of policies aimed to improve the health of riverine ecosystems throughout the state. Ultimately, this will serve as a new approach for finding a solution to conserving riverine ecosystems and their ecological integrity. The expected outcome will be a foundational step to apply water management plans as a restoration process.

2. Literature Review

A primary issue in water management is the availability of freshwater to concurrently meet water demands for the growing human populations while ensuring the integrity of freshwater ecosystems. This poses a challenge, which will become more difficult with human population growth, intensified land use, and climate change. A promising approach to this issue is integrating the concept of environmental flows to achieve reliable water supply while protecting the health of the ecosystem

Environmental flows describe the quantity, timing, and quality of water flows required to sustain freshwater and estuarine ecosystems and the human livelihoods and well-being that depend upon these ecosystems (Acreman et al., 2004). Environmental flows are an integral part of the continuity of the hydrologic cycle to produce outcomes beneficial to species, ecosystems, and people. The hydrologic cycle flows from place to place and from time to time, supplying water to aquatic ecosystems (Arthington et al., 2017). The goal is to develop a hydrologic classification for the impaired rivers in California by applying established hydrologic and ecological techniques with minimal resources and data requirements.

2.1 Environmental Flows

Aquatic ecosystems need water and other inputs like sediment and debris to stay healthy and provide environmental benefits to people. Environmental flows are critical contributors to the health of rivers and aquatic ecosystems (Scanlon et at., 2003). Depriving a river of these flows damages the entire aquatic ecosystem and threatens the communities who depend on it. The long-term absence of environmental flows puts the existence of dependent ecosystems at risk, and the lives, the livelihood and the security of downstream communities and industries.

Environmental flows provide a flow regime that would be adequate in quantity, quality, and timing for sustaining the health of rivers. Managing environmental flows is a complex process, as it is difficult to transform environmental studies and policies into action. Important factors for effective environmental flow management include a commitment from governments and stakeholders, sufficient resources, training and institutional capacity to manage water resources, enforcement, and adaptation. The five common major components of flow are extreme low flows, low flows, high flow pulses, small floods, and large floods. These components describe the variable river environment that an organism experiences. Table 1 lists some key ecological roles of each environmental flow component (combining large and small floods):

Table 1.Related the environmental flow component to the ecological roles (Sklar et at., 1998)

Environmental Flow Component	**Ecological Roles**
Low (Base) Flows	Provide adequate habitat space for aquatic organisms Maintain suitable water temperatures, dissolved oxygen, and water chemistry Provide drinking water for terrestrial animals Keep fish and amphibian eggs suspended Enable fish to move to feeding and spawning areas
Extreme Low Flows	Enable recruitment of certain floodplain plants Purge invasive, introduced species from aquatic and riparian communities Concentrate prey into limited areas to benefit predators

High Flow Pulses	Shape physical character of river channel including pools, riffles Prevent riparian vegetation from encroaching into the channel Restore normal water quality conditions after prolonged low flows, flushing away waste products and pollutants[18] Aerate eggs in spawning gravels, prevents siltation
Floods	Provide migration and spawning cues for fish Trigger new phase in the life cycle (e.g., insects) Enable fish to spawn on a floodplain, provide a nursery area for juvenile fish Provide new feeding opportunities for fish, waterfowl Recharge floodplain water table Maintain a balance of species in aquatic and riparian communities Create sites for recruitment of colonizing plants Deposit gravel and cobbles in spawning areas

Two types of flows are analyzed within environmental flows: reference and non-referenced. These types of flows help indicate the type of natural flow regime. Referenced flows are river flows without storage or diversions. Non-referenced flows are often a time series that is influenced by upstream disturbances of infrastructure, land-use change, or water diversions. These occur in rivers manipulated for human demands indirectly or directly. In systems where water is already over-allocated, a challenge of environmental flows may be reallocating or conserving water and returning it to the river.

2.2 Natural Streamflow Classification

Stream flow is a key element in the ecology of both rivers and streams. Knowledge of the natural flow regime facilitates the assessment of whether specific hydrologic attributes have been altered by humans in a particular stream and the establishment of specific goals for stream-flow restoration (Falcone et al.,2010). Riverine biota has evolved in the context of a “natural flow regime” – the quantity, timing, and variability of flow unaffected by human influences over many years – and quantifying that flow regime is important for maintaining ecosystem function and natural biodiversity (Richter et al., 1996; Poff et al., 1997; Poff et al., 2009). The natural flow regime is often only possible to be characterized by estimating flow characteristics based on nearby stream gauges of reference quality since human influences have altered most streams. This addresses the need for a development of a spatially explicit reach-scale hydrologic classification for California for both unimpaired and impaired streamflow data.

A study conducted by Lane et al. (2018), provides a broad-scale hydrologic framework upon which flow-ecology relationships could subsequently be established towards reach-scale environmental flow applications in a complex, highly altered Mediterranean region. This methodology identifies eight natural flow classes representing distinct flow sources, hydrologic characteristics, and catchment controls over rainfall-runoff response (Lane et al., 2018). Nine classes were identified for the State of California in Figure 2. This serves as a building block to improve our understanding of the diverse natural streamflow patterns needed to support future development in management applications.

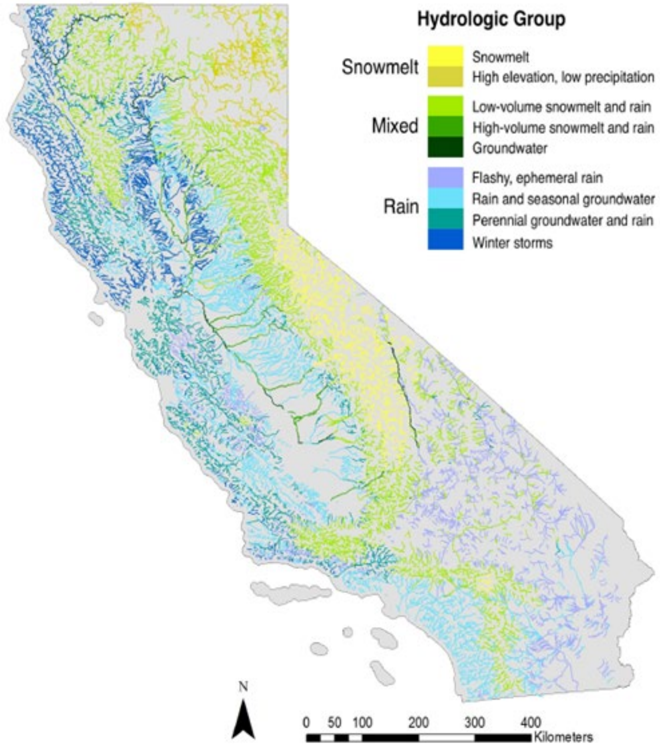


Figure 2. Natural Streamflow Classification of California from Noelle et al. (2020)

Alterations to flow regimes for water management objectives have degraded river ecosystems worldwide (Falcone et al., 2010). Changes are mostly seen in Mediterranean climate regions like California, where there are strong climatic variability and highly adapted riverine species to flooding and drought disturbances. Defining environmental flow targets for the Mediterranean river is quite complex. Thus, stressing the need to improve our understanding of the diversity of streamflow patterns to support future development of effective flow targets that can potentially aid in management applications. This study uses available impaired streamflow time series and generally publicly available geospatial datasets.

2.3 Assessing Hydrological Alterations within Ecosystems

Human land and water use are substantially altering hydrologic regimes all around the world, specifically in California. Evaluations of impaired hydrologic changes are needed to advance research to support ecosystem management and restoration plans. The biotic composition, structure, and function of aquatic and riparian ecosystems depend largely on the hydrologic regime. Interannual variation in hydrological conditions is essential to successful life-cycle completion as these conditions play a huge role in population dynamics of species through influences on reproductive success, natural disturbance, and biotic completion (Swanson, 2005). The interannual variation of hydrologic regimes is necessary to sustain the native biodiversity and evolutionary potential of aquatic ecosystems.

Modifications of hydrologic regimes can indirectly alter the composition, structure, or function of aquatic, riparian, and wetland ecosystems through their effects on physical habitat characteristics, including water temperature, oxygen content, water chemistry, and substrate particle sizes (Swanson, 2005). Effective ecosystem management of aquatic systems requires that existing hydrologic regimes be characterized using biologically relevant hydrologic parameters. Also, the degree to which human-altered regimes differ from natural or preferred conditions as it needs to be related to the status and trends of the biota. Sustaining ecosystem integrity requires ecosystem management that tests the need to maintain or restore the natural characteristics of the hydrologic regime.

Only a few studies have carefully examined hydrologic influences on ecosystem integrity, in part because statistical tools used are poorly suited for characterizing hydrologic data into biologically relevant attributes. The lack of statistical tools restricts further knowledge on the effects of hydrologic alterations on ecosystem integrity. This highlights the necessity of protecting or restoring natural hydrological regimes. The challenge lies in identifying and recording anthropogenic effects in the natural flow regimes of streams and rivers on public lands. On the contrary, there is a way to identify a river that has been highly modified.

Highly modified rivers are considered to be those that (1) have a high proportion of their total length converted to reservoirs, (2) have a high percentage of their total annual stream flow diverted and managed for societal uses, (3) have a high proportion of their total annual streamflow stored in reservoirs, and (4) have a large proportion of their total length channelized or lined by levees (Yarnell et al., 2010). The four characteristics rarely occur in the same river, but every one of them can significantly affect the riverscape, in terms of constraining e-flow implementation and ecosystem restoration potential.

A way to achieve highly modified rivers is by shifting the focus on ecological and geomorphological functionality of particular aspects of the flow regime, like considering geomorphic context and emphasizing spatiotemporal diversity at critical locations in the riverscape. A functional flow is a component of the hydrograph that provides a distinct geomorphic or ecological function (Pasternack et al. 2010 and 2011; Yarnell et al., 2015 and 2020). These functions may include geomorphic processes, ecological processes, or biogeochemical processes. These processes operate in three dimensions, such as longitudinally, laterally, and vertically. Also, they are connected to the timing, duration, and frequency of natural flows. Therefore, functional flows reflect the natural patterns of flow variability. Operational flows maximize the benefits from limited environmental flow allocations in highly modified rivers with complex water demands.

Some flow perturbations result from intentional releases and diversions outside the natural flow timing to generate electricity and deliver water elsewhere or to divert one stream segment into another. During high floods that exceed dam gates and valves capacities, there can be uncontrolled flows for reservoirs. Geomorphic processes, riverine species, and food webs are most likely to be sustained if dam operations mimic natural patterns of daily, seasonal, and annual variation of river flow. However, these operations often conflict with the demands of energy production and water, which is exemplified by California's Mediterranean climate. During wet winters, the runoff level is at its highest, but electrical and agricultural demands for water peak during dry summers.

Huge alterations to the natural hydrologic patterns in rivers affect a diverse array of species. Species that evolve with predictable annual flood-drought cycles are vulnerable to disruption of the synchrony between stable low-flow conditions and reproduction when river regulation alters the timing of historical free-flowing conditions. For example, the timing of spawning, hatching of eggs, and rearing of juveniles in anadromous Pacific salmon is shaped by the seasonal cycles of runoff, and ill-timed flow fluctuation can scour or dewater eggs and kill fry (Faustini, 2012). In parallel, when natural disturbance regimes are suppressed, river hydrology mimics regions with disparate seasonality, facilitating the recruitment of nonnative species. California's rivers exemplify the intensity of water resource development in which relations between flow management and persistence of native riverine biota are required to inform decision making.

Rivers change constantly through erosion and deposition. Normally, this can lead to displacements of the stream bed and channel line. On the other hand, this can be artificially magnified by human activity for either flood prevention or by canalization. The involvement of human activity, in erosion, leads to loss of habits and reduction in the biological communities in the river. There is a difference between sediment having occasional movement and permanent displacement. The natural order has some displacement of deposited sediments and rocks as it is a normal process in flowing water resulting in minimal impact on the biological community. Displacement on a large scale tends to prevent colonization by organisms. Each river type has a range of natural hydrologic variations that regulates characteristic ecological process and habitat characteristics, and that represents the reference condition against which ecological responses to alteration are measured across multiple river segments falling along a gradient of hydrological alterations (Chapman, 1996).

The long-term environmental degradation often outweighs the benefits. Understanding the degree to which anthropogenic activities have altered flows is critical for developing effective conservation strategies. Many studies have demonstrated that alterations of the natural flow regime are associated with changes in biological assemblages and altered hydrology is one of the dominant factors reported to affect the composition and health of aquatic species assemblages (Yarnell et al., 2015). Assessing flow alterations requires estimates of flows expected in the absence of human influence. Although more constant streamflow is desirable to support human use, such changes to natural variability across seasons, including a reduction to high-magnitude flows during rainy winters and warm spring snowmelt periods and augmentation to low-season flows during dry summers, have been shown to have ecological consequences (Yarnell et al., 2015). Collectively, alteration to natural streamflow patterns has been documented to have negative effects on California's aquatic biota, and there is evidence that restoring components of natural hydrology can provide substantial ecological benefits (Yarnell et al., 2015).

2.4 Ecosystem and Ecology Reconciliation

Humans have changed the land surface in some way that will affect the flora and fauna of the area (Franklin, 2007). In other words, the amount of land left over for the millions of other species that live on our planet continues to diminish, leaving a small amount left over that is not nearly enough to sustain. Rosenzweig (2014) developed the idea of reconciliation ecology. In his own words, "reconciliation ecology is the science of inventing, establishing and maintaining new habitats to conserve species diversity in places where people live, work and play.

In a seminar at UC Davis, the ecologist Micheal Rosensweig talks about using the reconciliation approach to improving California's aquatic habitat. There are two distinct threats of a mass extinction when habitat is removed: qualitative and quantitative remedy. One is answered through the reservation ecology approach which makes sure no habitat disappears together. This action has saved species throughout the world by slowing extinction rates. The issue arises when quantity is stressed. On the other hand, a quantitative remedy is answered by reconciliation ecology to make sure enough habitat exists. Rosenzweig proposes that the solution to this challenge is reconciliation ecology. Rosenzweig stresses that people should not think that the only way to make a profit is by destroying things, but rather to start thinking about coexisting with nature in a proactive way where we redesign our future where both succeed.

A serious effort from scientists and the general public will be required to establish and maintain new habitats in cities and towns. This means that scientists must find a way where species can potentially coexist with humans. There are species that will never be able to live within human populated areas, like large mammal species. But many species may be able to thrive within cities if their habitat requirements are taken into consideration (Franklin, 2007). This idea is achievable with a group effort, it must take the cooperation of several individuals to successfully reach this idea of living in harmony with species.

Although Rosenzweig first mentioned the term in the early 2000s (Rosenzweig 2014), the concept of reconciliation ecology is not new. There have been efforts made to provide some guidelines for creating policies and regulations to reconcile human and environmental uses of water. In the recent Managing California's Water From Conflict to Reconciliation, one of the sections of the chapter discusses three general conservation strategies—reservation, restoration, and reconciliation—to create long-term solutions, focuses on reconciliation as a way to deal with major environmental problems, and discusses legal means to achieve reconciliation (Hanak et al.,2011). Reconciliation requires actions that create better conditions for species and humans. Reconciliation recognizes that humans so completely dominate the planet that conservation of species and their habitats depends on integrating native ecosystem functions into ecosystems shaped by human activity (Hanak et al.,2011), structures whose functions are controlled by continual human management. For example, maintaining native fish in rivers below dams in California requires not only adequate flow releases from the dam but releases with appropriate temperatures and volumes on a schedule that follows the natural flow regime (Hanak et al.,2011). Reconciliation actions will not bring back historical conditions, but they create a new era. Examples such as stabilizing the list of endangered species and allowing for recovery in numbers and environments that allow native species to exist for the long-term. Currently, there is a proposed reconciliation strategy alternative for the Delta that blends both news of ecosystems and humans and its called "eco-friendly Delta" as shown in Figure 3.

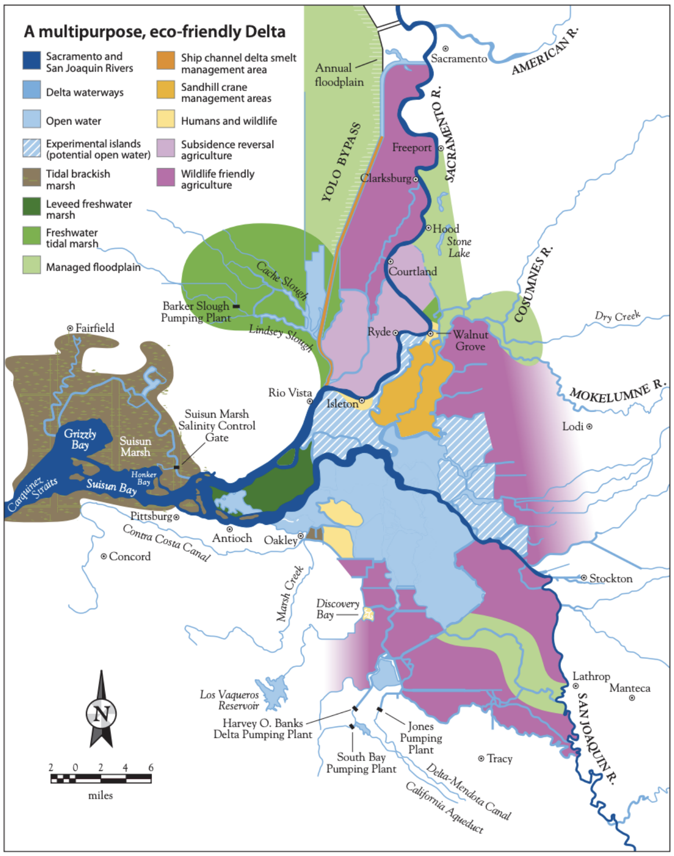


Figure 3. A reconciled, "eco-friendly" Sacramento–San Joaquin Delta would have multipurpose land and water uses. Source: Lund et al. (2010)

This new Delta, described in more detail in Lund et al. 2010, seeks to accommodate inevitable future changes (higher sea level, earthquakes, additional permanently flooded islands, and changing inflows as a result of climate shifts), seeks to maintain substantial and profitable agricultural use of Delta lands in ways that support native wildlife, and creates or improves aquatic habitats and functions needed to support desirable fish species (tidal marsh and open water habitat, along with variable hydrology and salinity) (Lund et al., 2010). This proves that reconciliation ecology can be done, but it will take the effort of both scientists and the general public to successfully carry out the approach.

3. Case Study: The State of California

California is a geographically diverse area that has a Mediterranean climate with warm, dry summers and mild, wet winters. There are two main things that affect California in terms of the water cycle, climate, and the extreme gradient in elevation. Atmospheric rivers produce up to fifty percent of California's precipitation annually and sixty-five percent seasonally (Arcuni, 2019). Atmospheric rivers start at the coastal range moving upwards cold temperature resulting in rain. Once in the Central Valley, the atmospheric rivers get warm. In these areas, there is no precipitation and it is often referred to as the rain shadow effect. The bank of moisture will then travel to the Sierra Nevada mountains coming in as two different types. In the higher altitude, it will be snow and in the lower altitude less than 5 thousand feet above sea level it will be in the form of rain, this elevation at which precipitation falls as snow or rain is called the snow line.

California will always be inseparably linked to water resources as it continues to shape the state’s development. It is the single most vital resource to California’s farms, urban centers, industry, recreation, and environmental preservation. Distributing and sharing this resource is the most basic issue affecting water supply in California as it requires getting the water to the right place at the right time while not harming the aquatic species and the environment. Nearly 75 percent of the available surface water originates in the northern third of the state (north of Sacramento), while 80 percent of the demand occurs in the southern two-thirds of the state (Hanak et al. 2011). The distribution often conflicts between competing interests over the use of available supplies.

California’s networks of dams, canals, levees, and water treatment plants, along with the laws, regulations, and institutions that govern them, have evolved over the course of more than 160 years. The historical foundations of today’s water systems date back to the 1800s. The laws, policies, and infrastructure of today derive from the laissez-faire approaches to water during and immediately after the Gold Rush in the mid-1800s, the drive to develop local water supplies in the late 1800s, and the local, state, and federal efforts in the 20th century to redistribute water throughout California, creating one of the most complex and ambitious water supply and flood control systems in the world (Hanak et al.,2011). The 1970s brought a new concern, as society acted to protect the ecological health of the state’s waters. Ever since, California has struggled with the apparent conflict between ecosystem and water management (Hanak et al.,2011). The Central Valley Project was the world's largest water and power project devised in 1933.

These historic events of water development projects have caused the alteration we see in the State of California. The evolution has been a response to rapid population growth, demographics, demands, droughts, floods, and lawsuits. Figure 4 demonstrates the natural streamflow classification and the network conveyance and storage infrastructure operated by different agencies.

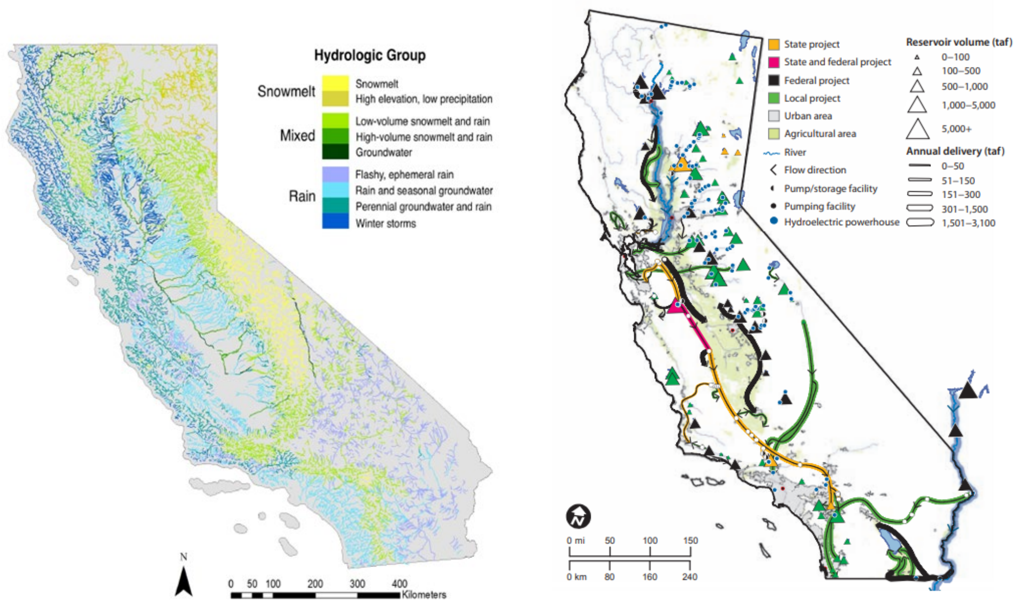


Figure 4. California's network or conveyance and storage infrastructure (Hanak et al. 2011)

Streamflow alterations have been monitored and recognized for a long time, the main contribution to the body of knowledge of this study is to propose a method to classify alteration and identify the locations where it is happening. In the 1979 California Water Atlas (Figure 5), the observed (in yellow) and estimated natural flow (in blue) were estimated for the water year 1975 (October 1974 through September 1975). The nearer of the pair is representing the actual flow measured at the gaging station (in yellow), while the farthest diagram represents the hypothetical natural condition flow (in blue) as it would have been if there were no artificial diversions or storage facilities. In Figure 5, each diagram represents streamflow past a gaging station in water years (from October to September).

Figure 5. Measured and Unimpaired Streamflows (Rumsey, 1979)

4. Method

The overall goal of this research is to develop a method for establishing an altered streamflow altered classification. Thus, the hydrologic classification development consisted of three steps (Fig. 2). Step 1 addresses the identification of the non-reference streamflow gage list (further explained in Section 4.1), step 2 addresses classifying the gages by their impairment (further explained in Section 4.2), and step 3 addresses the spatially prediction of the altered streamflow classes (further explained in Section 4.3). Figure 6 shows the workflow for accomplishing the three main goals of this. The following sections will provide a thorough explanation of each step of the process.

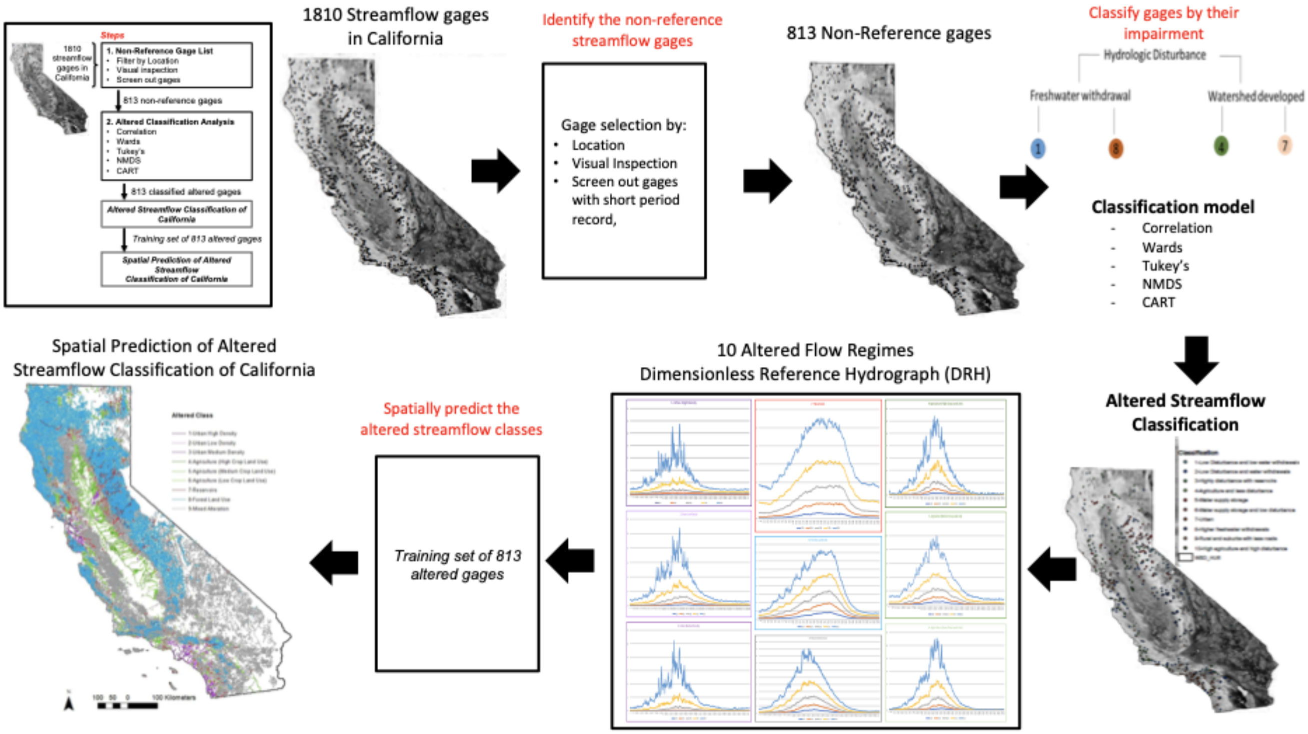


Figure 6. Overview of the methods used to determine the Impaired Streamflow Classification

4.1 Non-reference gauge list

The U.S. Geological Survey (USGS) streamflow gage data, Gages II, were analyzed to determine the non-reference streamflow gages (Falcone et al., 2010). This dataset, referred to as GAGES II (Geospatial Attributes of Gages for Evaluating Streamflow, version II), provides geospatial data and classifications for stream gages maintained by the USGS. The sites comprise all USGS stream gages in the conterminous United States with at least 20 years of complete-year flow record from 1950–2019. From the USGS list of streamflow gages in California, there were a total of 1810 gage stations (Falcone et al., 2010). Three primary sources of information were used in identifying reference-quality streamgages: (1) filtering by location based on ArcGIS location, (2) visual inspection of every stream gage and drainage basin from recent high-resolution imagery, and (3) screen out of gages.

Filtering by location based on ArcGIS location

The USGS GAGES II database contained 1810 gage stations located in the states of California, Nevada, Oregon, Arizona, and the border of Mexico. For the purpose of this research only the ones located in California were selected. The sites not selected may be of use for this research, but are still as important and can be used when a classification is done to other states.

Visual inspection of every stream gage and drainage basin from recent high-resolution imagery,

A mapping and analytics platform (ArcGIS) was used to visually inspect the sites. The information used was the site name, location (latitude and longitude), California's main rivers and streams, and a basemap. The basemap served as a reference to identify if the sites were located along any water infrastructure, e.g. intakes, aqueducts, ditches, drains, canals, diversions, hydropower plants, etc. This research main objective is to develop a classification for the rivers that have been altered, not of the streamflow in water infrastructure. Thus, gage sites that were located on water infrastructure were not considered.

Screen out of gages

Gage screen out required a careful analysis of site inspection of location and information presented from USGS. Each site had information regarding common identifiers, gage type (Reference, Non-reference, and NA), and period of record from 0 to 39 years. The common identifier (ComID) of the NHDFlowline feature of the gage location is used. The sites that did not have a ComID were not selected. Part of this research study is to spatially predict the altered streamflow classes and without a ComId the spatial prediction would not be possible. Each gage had additional breakout point notes like redundant gauge II reference site, insufficient record, failed preliminary OE screen, NA, probable hydro alteration, no evidence that was carefully investigated prior to deciding on whether it should be selected or not. Sites whose period of record was less than 5 years were not selected.

4.2 Altered Streamflow Classification

In designing a non-reference streamgage dataset for predicting altered flow regimes, we need a population of sites having adequate long-term streamflow data from watersheds primarily influenced anthropogenically. The creation of the Non-reference gage list was explained in the previous section and will be used for the classification on altered streamflow. Our intent was to apply a consistent set of criteria to identify watersheds with adequate streamflow records, with anthropogenic influences based on the most recent data available in California. To accomplish this, we assembled a large dataset incorporating features of these watersheds using StreamCat and GAGEs II. The StreamCat Dataset provides summaries of natural and anthropogenic landscape features for ~2.65 million streams, and their associated catchments, within the contiguous USA (Hill, Weber, Leibowitz, Olsen, Thornbrugh, 2015). The GAGES II provides disturbance variables that have been used to estimate the degree of alteration in other studies (Falcone 2010). For this study, we selected 33 predictor variables that represent physical attributes (5 variables) water storage (3 variables), agriculture (3 variables), urban land use (12 variables), forest land use (8 variables), water quality (1 variable) and water withdrawals (1 variable).

These predictor variables values were taken from Streamcat and GAGES II, they were used as input parameters for the methods used to determine the non-reference classification. As illustrated in figure 4, the overall statistical methods are Pearson Correlation, Non-metric Multidimensional Scaling (NMDS), Hierarchical Clustering using Ward's Algorithm, Tukey's box and whisker plot, and Classification and Regression Tree Analysis (CART). These will be explained in detail in sections 4.2.1. and 4.2.2.

4.2.1 Eliminating redundant variables

Pearson Correlation was the analysis used to evaluate the correlation among all predictor variables (33 indicators). This method identifies which variables are highly, linearly correlated. If two attributes were highly correlated, one of the two would be removed. The removal process between any two, if this was to occur, would be based on whoever had the lowest cross-validation when combined with the other indicators. The correlation analysis will be done using the 33 predictor variables.

4.2.2 Statistical Approach

The statistical approaches used are Non-metric Multidimensional Scaling (NMDS), Hierarchical Clustering using Ward's Algorithm Classification and Regression Tree analysis, Tukey's honestly significant differences and box and whisker plots, and cross-validation.

NMDS is a statistical approach used to better understand how the sites cluster in multivariate space and which predictor variables are driving the clustering. The NDMS is not used as a final statistical classification, it is only used identifying key variables.

Box and whisker plots are used to observe differences in individual influences of the predictor variables and to interpret the classification based on differences between the nine anthropogenic influences. Analysis of variance between all of the groups provides Tukey's honestly significant differences. This provided a heuristic approach for defining the class names. To determine the number of final classes in the hierarchical clustering, we determined which predictor variables had a higher cross-validation.

4.2.3 Altered Classification Analysis

Hierarchical Clustering using Ward's Algorithm utilizes variance to determine which sites have the most similarities and which sites are most dissimilar. The amount of dissimilarity between sites is represented through the vertical axis of the associated dendrogram. The connections at higher levels represent combinations of more dissimilar sites. The classification out of this method was verified by visually inspecting the box and whisker plots.

The Classification and Regression Tree (CART) analysis is used to achieve a multivariate classification that makes physical sense with respect to the reduced disturbance indicators. CART is a classification tree that splits all sites into smaller groups based on values at each site. On the other hand, Ward’s hierarchical clustering starts with individual sites and combines sites into larger groups. Our goal is to have a high classification tree prediction rate defined by hierarchical clustering. The prediction rate will be used when performing cross-validation to better understand the classification.

4.2.4 Flow regime - Dimensionless Non-reference Hydrographs (DNH)

The eFlows Functional Flow Calculator (FFC) quantifies key hydrologic aspects of the annual flow regime from any daily streamflow time series which was used to the 813 streamgages. The FFC produces dimensionless reference hydrographs (DRHs) that serve as a descriptive visual tool of continuous daily and inter-annual streamflow patterns considering the reference gages used for the natural streamflow classification (Lane et al. 2018). The Dimensionless Non-reference Hydrographs (DNH) are calculated for each non-reference gage by dividing daily streamflow data by the water year's average annual flow across all water years of flow data. The 10th, 25th, 50th, 75th, and 90th percentile flows over the entire reference period of record are determined for each date of the water year and plotted. This will serve as an illustration of the range of non-dimensionalized flow that occurs across the water year at a daily time-step

4.3 Spatial Prediction

Key variables were used as input data into machine learning algorithms to build a model that provided the best spatial prediction throughout the river network. Altered classes were predicted in the river network using three machine learning algorithms: random forest, support vector machine, artificial neural networks. Each model was trained using a ten-fold training data set, meaning the model was trained with 90% of the data end evaluated with the remaining 10% of the data left out. The cross validation technique was used to assess the performance of the best machine-learning model.

5. Results

5.1 Non-reference gage list

This section describes the methods, steps, and assumptions made for determining the list of *Non-reference streamflow gauges* for the state of California. Streamflow gauges that are deemed as *Reference*, are those that do not exhibit human-induced alteration in their streamflow time series data (Lane et al. 2018). *Non-reference* gauges are those gauges that exhibit anthropogenic alteration in the streamflow time series data and are located along the river network (i.e. not including intakes, ditches, aqueducts, canals, etc.). Of the 1810 gage stations USGS located in California, these two conditions must be met for a gage to be deemed as Non-Reference. This condition is important because the objective of the project is to develop a classification for rivers that have been altered, not to classify the alteration in the water infrastructure, which can be monitored by streamflow data in intakes, aqueducts, ditches, drains, etc. Figure 7 shows the process and analysis for selecting Non-reference gages.

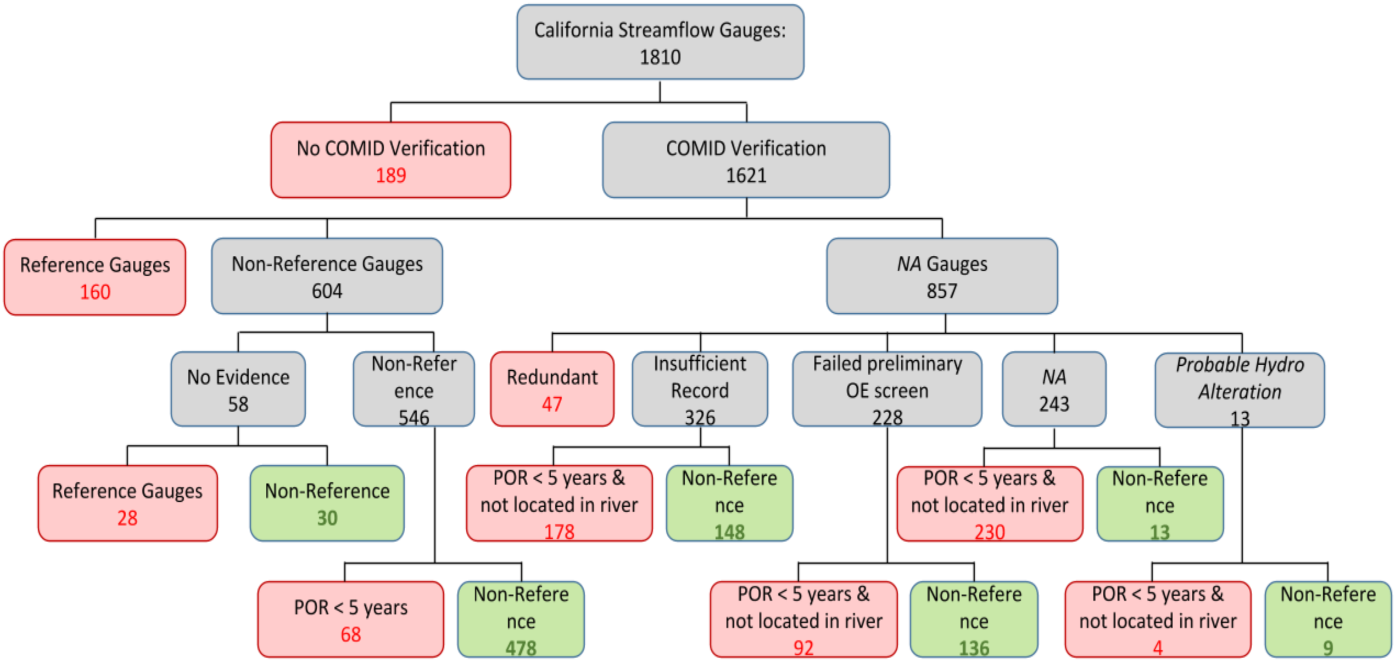


Figure 7. Streamflow gauge analysis to determine the definitive list of Non-Reference streamflow gauges

First, a verification on the ComID location was done for each gage, resulting in only 1621 of the 1810 gauges located in California. The remaining 189 gage stations were considered: not available, not found, or suspect and were not considered in this study. The filtering of the following analysis can be seen in Figure 7.

Second, we analyzed the list of 1621 streamflow gauges, of which 160 gage sites were classified as *Reference*, 604 as *Non-Reference*, and 857 as *NA*. For the 160 *Reference* streamflow gauges, we inspected their periods of record and verified that no records were available before or after the period for which they were deemed Referenced. The authors compared the 160 *reference* streamflow gauge of GAGES II with the 223 *reference* gauges identified by Lane et al. (2018). Out of the 160 reference gauges identified in this study, only 131 coincide with those identified by Lane et al. (2018). The remaining 29 *Reference* gauge stations were not included in Lane et al. (2018) because of the short period of record, incomplete data, or suspicious information.

Third, we further analyzed the 604 gage stations considered *Non-Reference*. Out of these 604 gages, 546 were already classified as *Non-Reference* by Falcone et al. (2010) and 58 classified as *No-Evidence*. The list of 546 streamflow gauges was visually inspected and only 493 are included in the definitive *Non-Reference* list. For the remaining 58 *No-Evidence* gages (referring to no evidence of hydro-alteration during the specified periods), we evaluated if there was streamflow data before or after the period of *Reference* record. 30 streamflow gauges found to have data after outside the period that was considered *Reference* were also included in the definitive *Non-Reference* list.

Fourth, we split the list of 857 gauge stations classified as “NA” into five groups: (1) 47 gauge stations classified as *Redundant with gauge II reference site*, (2) 326 gauge stations classified as *Insufficient Record*, (3) 228 gauge stations classified as *Failed preliminary OE screen*, (4) 243 gauge stations classified as *NA*, and (5) 13 gauge stations classified as *Probable Hydro Alteration*. The 47 gage stations were disregarded from the definitive *Non-Reference* list because through a visual inspection there was no evidence of hydrologic alteration and they were noted as *Redundant* in the USGS gauge list.

For the remaining three groups (326, 228 and 243 streamflow gauge lists) the following analysis was conducted. First, we analyzed the name of each gauge station and disregarded any gauge that included one of the following infrastructure names: diversion, canal, ditch, aqueduct, drain, release, weir, conduit, combined, powerhouse, bypass, tunnel, powerplant. The goal here was to eliminate any gauge located along anthropogenic infrastructure rather than along the river corridor. Second, we spatially located each gauge and added a buffer of 5 meters of diameter. Then we selected those that were located within the river corridor and disregarded the 500 gages that were not (178 from the insufficient record, 92 from Failed preliminary OE screen and 230 from NA). From this analysis, 297 gage stations were included in the definitive Non-Reference list. Finally, the 13 gages classified as Probable Hydro Alteration were visually inspected and only 9 were included in the definitive Non-Reference list.

In the end, there were 829 gage stations selected as part of the definitive list of which only 813 had actual data (flow and dates). Thus, the final count for the streamflow gages used to identify impaired classification in California is 813.

5.2 Altered Classification Analysis

From the set of 1810 sites from USGS, we identified 813 as non-reference quality stream gages. We collected data from StreamCAT and Gages II for 33 variables for the 813 non-reference streamflow gages, these variable are considered a proxy for documented streamflow alteration in California: physical attributes (5 variables) water storage (3 variables), agriculture (3 variables), urban land use (12 variables), forest land use (8 variables), water quality (1 variable) and water withdrawals (1 variable). (Table 2).

Table 2. Analysis contained 33 variables

Indicator name	**Description**	**Predictor type**
Stor_Nor_2009	Normalized upstream (from gage) reservoir storage from 1950-2006.	Water Storage

DamDensCat	Density of georeferenced dams within catchment (dams/ square km) based on the National Inventory of Dams	
Canals_Pct	Canals/ditches/pipelines/artificial_path in the watershed	Agriculture
CropsNLCD06	Percent of watershed in cultivated crops (NLCD class 82)	
Pct_Irrig_Ag	Percent of watershed in irrigated agriculture, from published USGS sources	
Npdes_Maj_Dens	Density of NPDES (National Pollutant Discharge Elimination System) "major" point locations in watershed	Water Quality
DevNLCD06	Percentage of developed land from the 2006 National Land Cover Dataset	Urban Land Use
PctUrbOp2006Ws	Percent of watershed area classified as developed, open space land use (NLCD 2006 class 21)	
PctUrbLo2006Ws	Percent of watershed area classified as developed, low-intensity land use (NLCD 2006 class 22)	
PctUrbMd2006Ws	Percent of watershed area classified as developed, medium-intensity land use (NLCD 2006 class 23)	
PctUrbHi2006Ws	Percent of watershed area classified as developed, high-intensity land use (NLCD 2006 class 24)	
PopDen2010Ws	Mean population density (people/square km) within watershed	
Roads_Km_Sq_Km	Distance of stream gage to nearest major pollutant discharge site	
PctUrbOp2006Cat	Percent of catchment area classified as developed, open space land use (NLCD 2006 class 21)	
PctUrbLo2006Cat	Percent of catchment area classified as developed, low-intensity land use (NLCD 2006 class 22)	
PctUrbMd2006Cat	Percent of catchment area classified as developed, medium-intensity land use (NLCD 2006 class 23)	
PctUrbHi2006Cat	Percent of catchment area classified as developed, high-intensity land use (NLCD 2006 class 24)	
PopDen2010Cat	Mean population density (people/square km) within catchment	

Freshw_Withdrawal	Freshwater withdrawals in the United States	Water Withdrawals
PctConif2006Ws	Percent of watershed area classified as evergreen forest land cover (NLCD 2006 class 42)	Forest Land Use
PctConif2006Cat	Percent of catchment area classified as evergreen forest land cover (NLCD 2006 class 42)	
PctIce2006Ws	Percent of watershed area classified as ice/snow land cover (NLCD 2006 class 12)	
PctIce2006Cat	Percent% of catchment area classified as ice/snow land cover (NLCD 2006 class 12)	
PctFire2010Cat	Percent Forest loss to fire (fire perimeter) for 2010 within catchment	
PctFire2010Ws	Percent Forest loss to fire (fire perimeter) for 2010 within watershed	
PctFrstLoss2013Cat	Percent Forest cover loss (Tree canopy cover change) for 2013 within catchment	
PctFrstLoss2013Ws	Percent Forest cover loss (Tree canopy cover change) for 2013 within watershed	
ElevWs	Mean watershed elevation (m)	Physical attributes
Precip8110Ws	PRISM climate data - 30-year normal mean precipitation (mm): Annual period: 1981-2010 within the watershed	
WsAreaSqKm	Watershed area (square km) at NHDPlus stream segment outlet, i.e., at the most downstream location of the vector line segment	
RunoffCat	Mean runoff (mm) within catchment	
Runoffs	Mean runoff (mm) within watershed	

5.2.1 Eliminating redundant variables

A correlation analysis was done to see if any of the 33 parameters were highly correlated. Having parameters that are correlated with others makes the analysis redundant. To minimize this and in order to achieve the highest cross validation performance, two runs were performed.

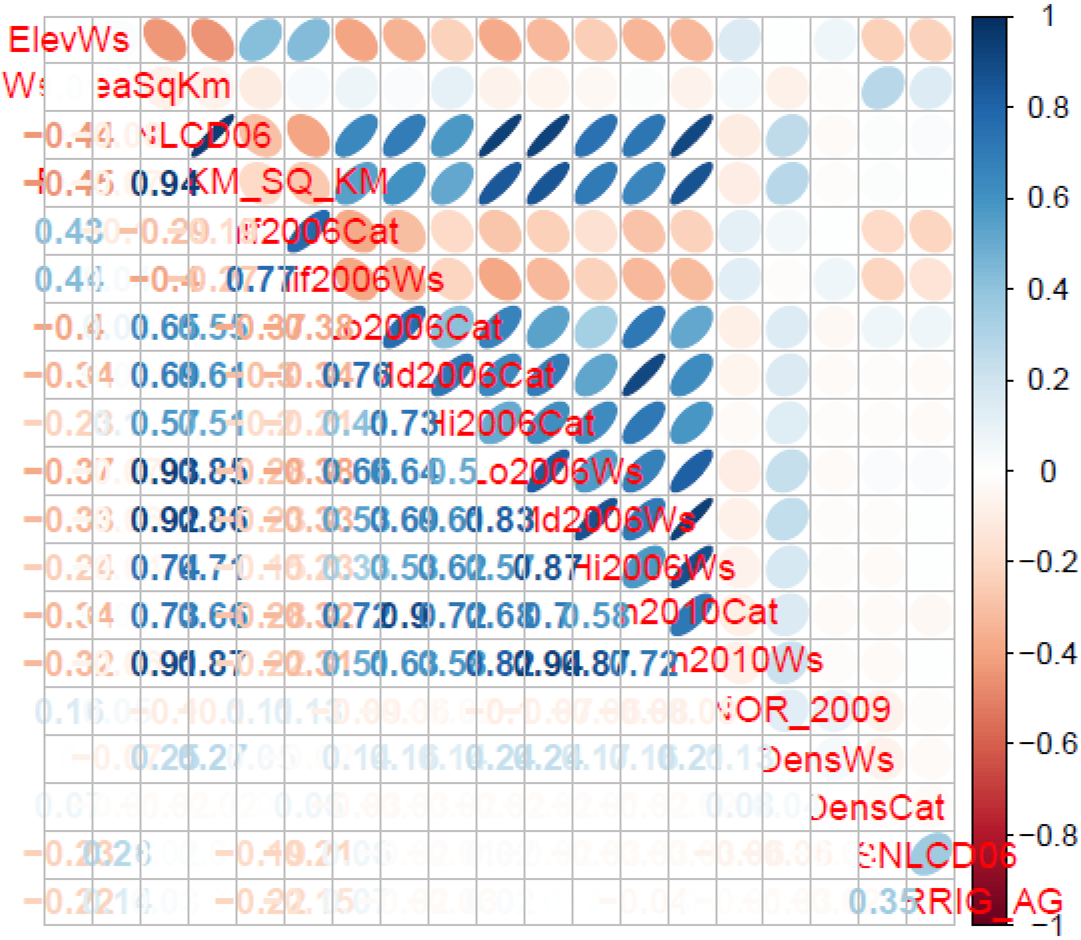


Figure 8. Pearson Correlation for all 33 variables

The first correlation analysis was performed to see the similarities between all the parameters combined as illustrated in Figure 8. This was the initial step in minimizing the variables in the study. After the initial correlation analysis, eight variables were selected because they were variables representative of the well-documented impairments and they did not exhibit high correlation among themselves. The eight variables are listed in Table 3 and the corresponding correlation is illustrated in Figure 9.

Table 3. Eight indicators of disturbance estimated

Indicator name	**Description**
PctUrbLo2006Cat	Percent of catchment area classified as developed, low-intensity land use (NLCD 2006 class 22)

PctUrbMd2006Cat	Percent of catchment area classified as developed, medium-intensity land use (NLCD 2006 class 23)
PctUrbHi2006Cat	Percent of catchment area classified as developed, high-intensity land use (NLCD 2006 class 24)
PopDen2010Cat	Mean population density (people/square km) within catchment
Stor_Nor_2009	Normalized upstream (from gage) reservoir storage from 1950-2006.
DamDensWs	Density of georeferenced dams within watershed (dams/ square km) based on the National Inventory of Dams
CropsNLCD06	Percent of watershed in cultivated crops (NLCD class 82)
Pct_Irrig_Ag	Percent of watershed in irrigated agriculture, from published USGS sources

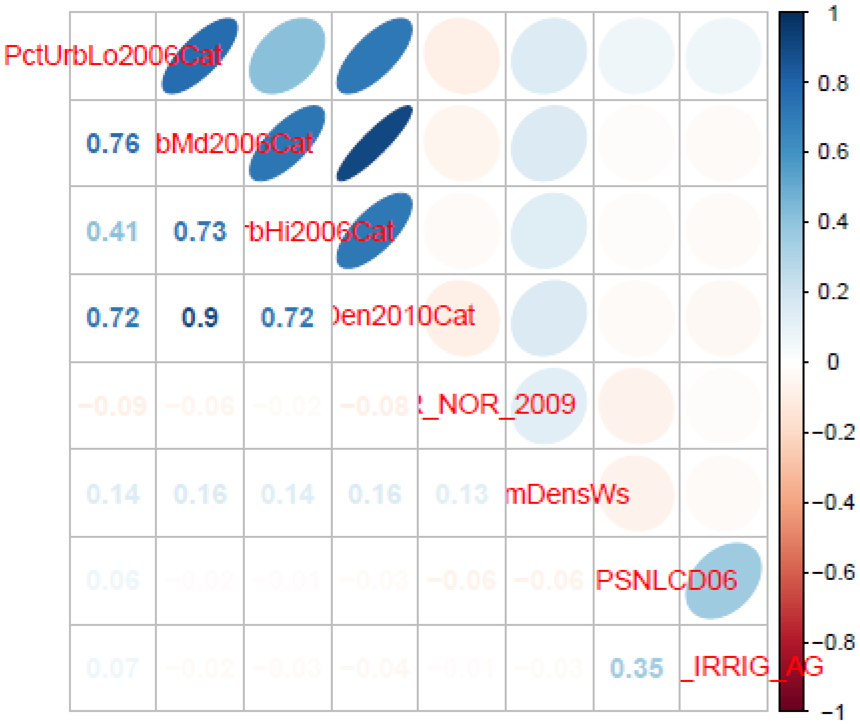


Figure 9. Results of the Pearson Correlation analysis for the nine indicators of disturbance

5.2.2 Statistical Approach

The statistical approaches used are Non-metric Multidimensional Scaling (NMDS), Hierarchical Clustering using Ward's Algorithm Classification and Regression Tree analysis, Tukey's honestly significant differences and box and whisker plots, and cross-validation.

NMDS is a statistical approach used to better understand how the sites cluster in multivariate space and which of the eight indicators are driving the clustering. As demonstrated in Figure 10, the indicators are fairly clustered.

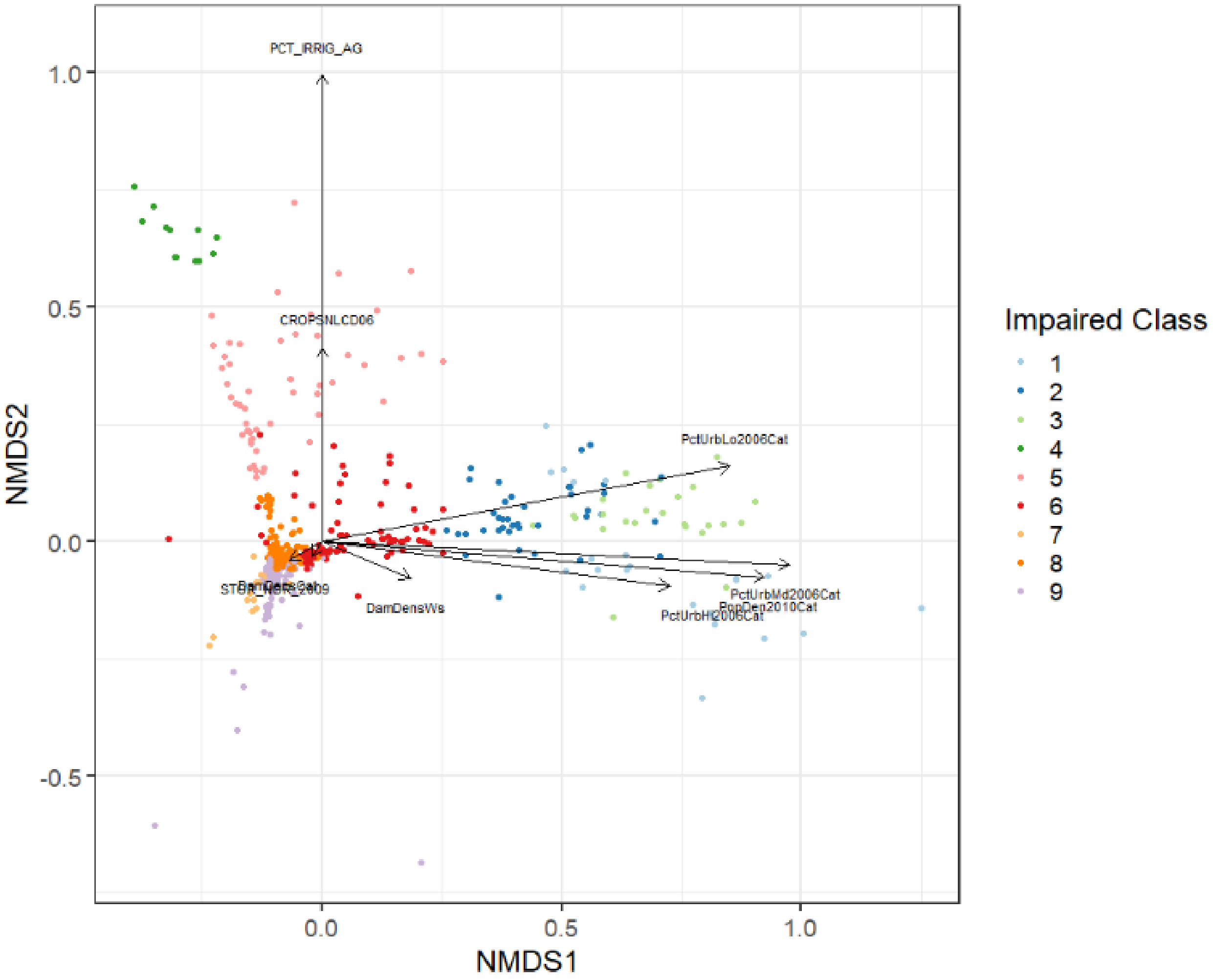


Figure 10. NMDS Plot (Note: Colors defined by final anthropogenic influences)

Box and whisker plots are used to observe differences in individual influences and to interpret the classification based on differences between the eight anthropogenic influences. Figure 11 shows how each class in the classification is distinguished by its own characteristics. Class 1, 2, and 3 demonstrate strong influence with respect to urbanization, as opposed to class 7 which is clearly distinguished by dams.

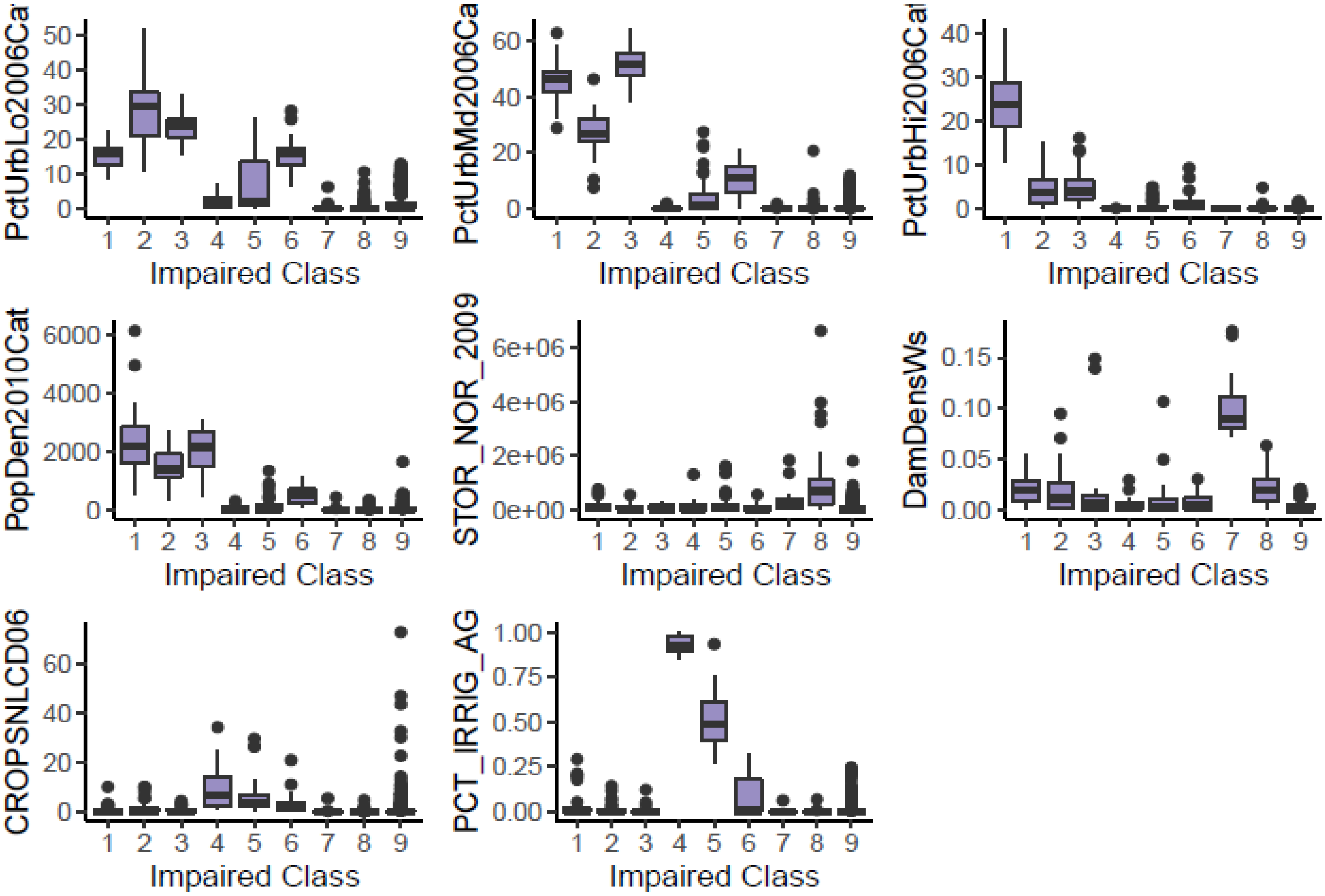


Figure 11. Box and whisker plots representative of statistical differences between anthropogenic types for each anthropogenic influence.

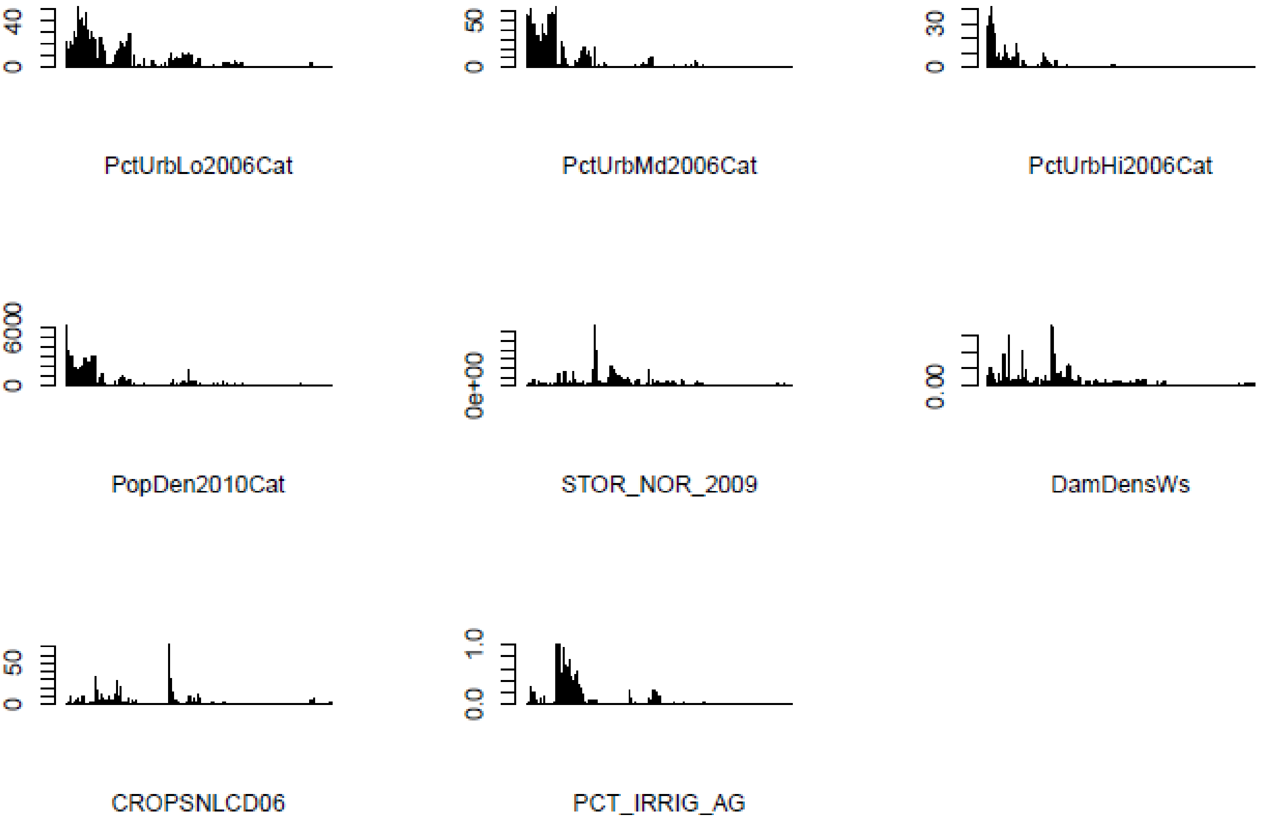


Figure 12. Channel attribute values in order of hierarchical dendrogram.

The impaired classification analysis approaches used are Hierarchical Clustering using Ward's Algorithm and CART. Figure 13 shows the layout of Ward's algorithm with just breaking out the classes without any level defined (height). Connections at higher levels represent combinations of more dissimilar sites.

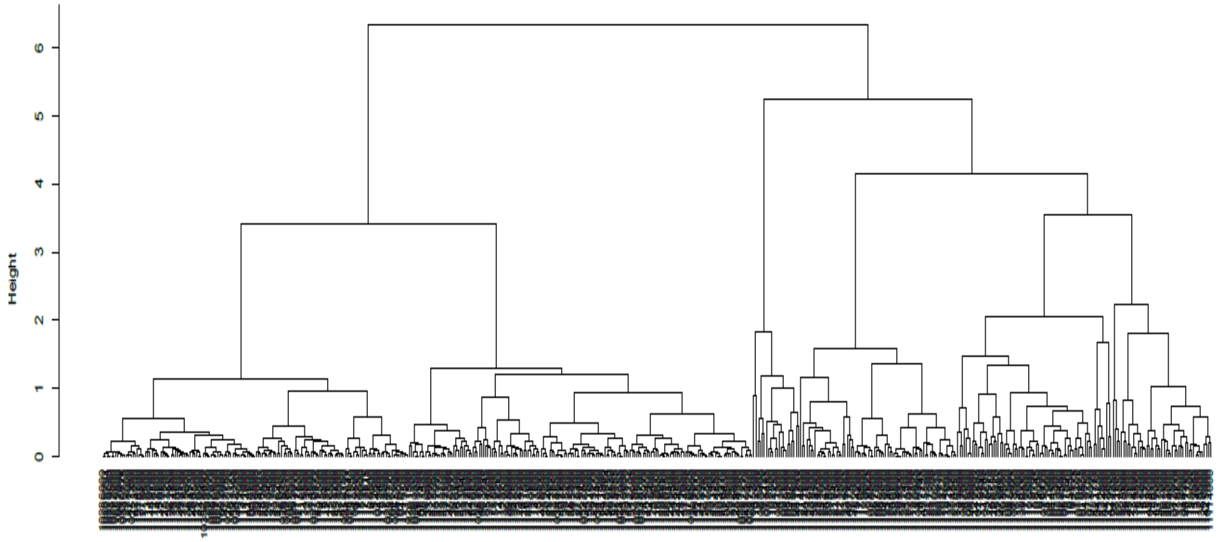


Figure 13. Dendrogram of hierarchical clustering with Ward's algorithm

After the Hierarchical-clustering algorithm is complete it breaks down the classification to the designated 9 classes as demonstrated in Figure 14.

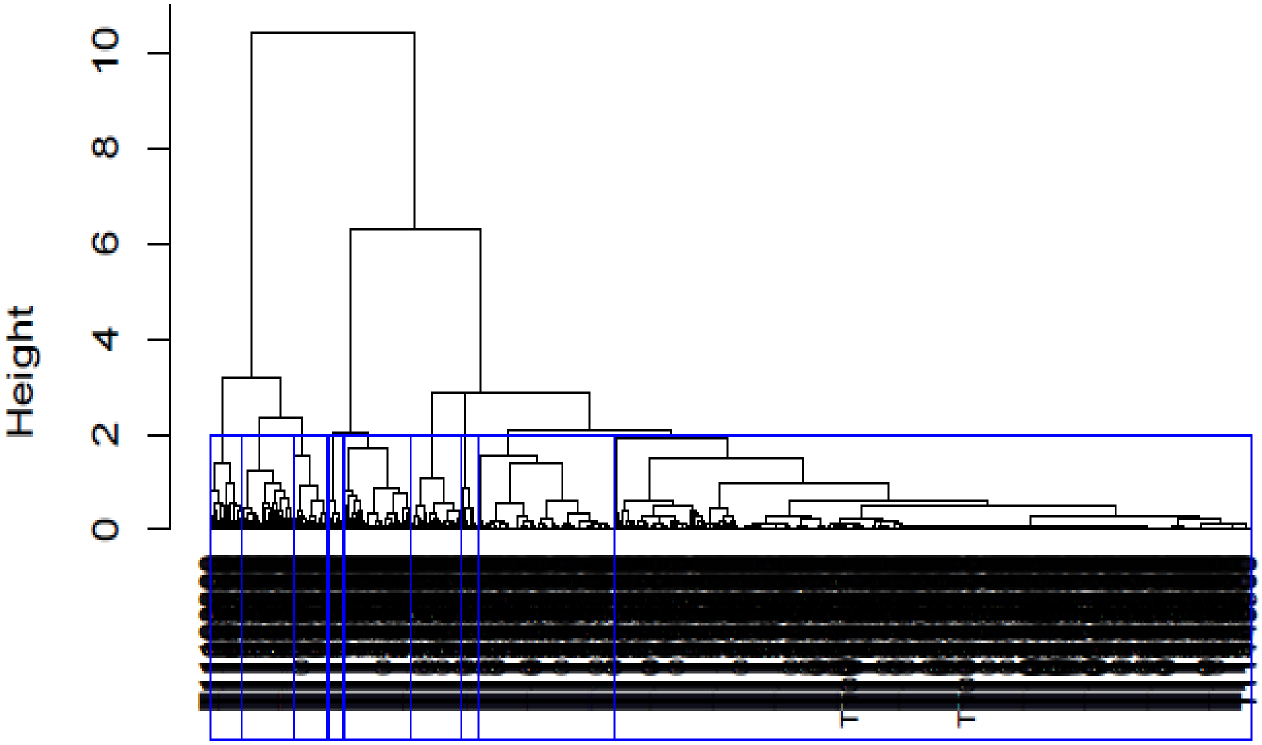


Figure 14. Dendrogram of hierarchical clustering with Ward's algorithm

Classification tree analysis uses individual channel attributes to achieve the same classification as the hierarchical clustering. Ultimately showing how the eight parameters are broken down into the 9 classes as seen in Figure 15.

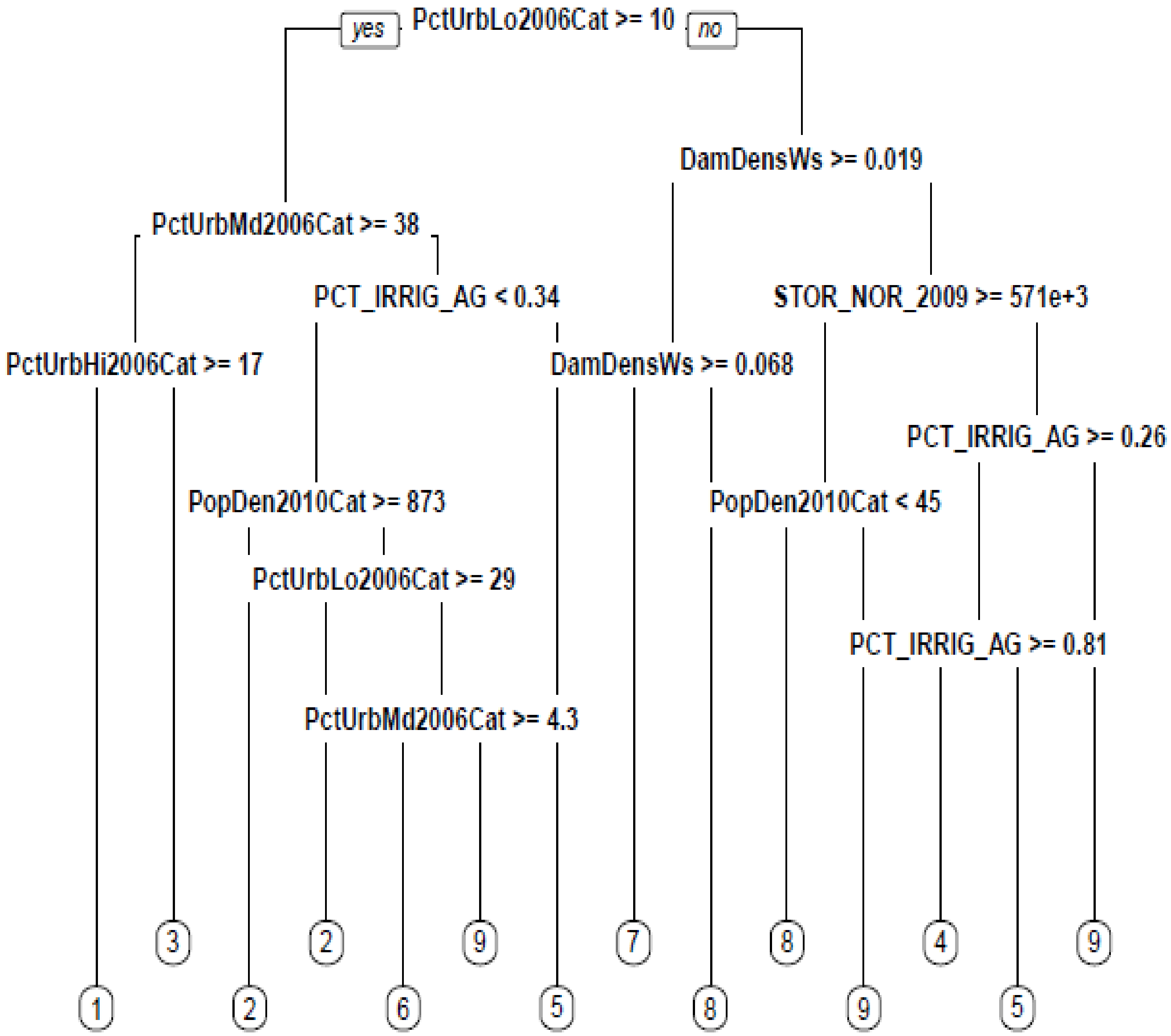


Figure 15. Classification tree for the impaired classes.

5.2.3 Altered Classification Analysis

The class names were determined according to variables identified from the classification tree as listed in Table 4.

Table 4. Classification and Frequency for each of the classes

Classes	Quantity of Gages	Frequency
1. Urban High Density	25	3%
2. Urban Low Density	40	5%
3. Urban Medium Density	27	3%
4.Agriculture (High Crop Land Use)	12	1%
5. Agriculture (Medium Crop Land Use)	53	7%
6. Agriculture (Low Crop Land Use)	39	5%
7. Reservoirs (High Density Low Volume)	13	2%
8. Reservoirs (Low Density and High Volume)	106	13%
9. Mixed Alteration	498	61%

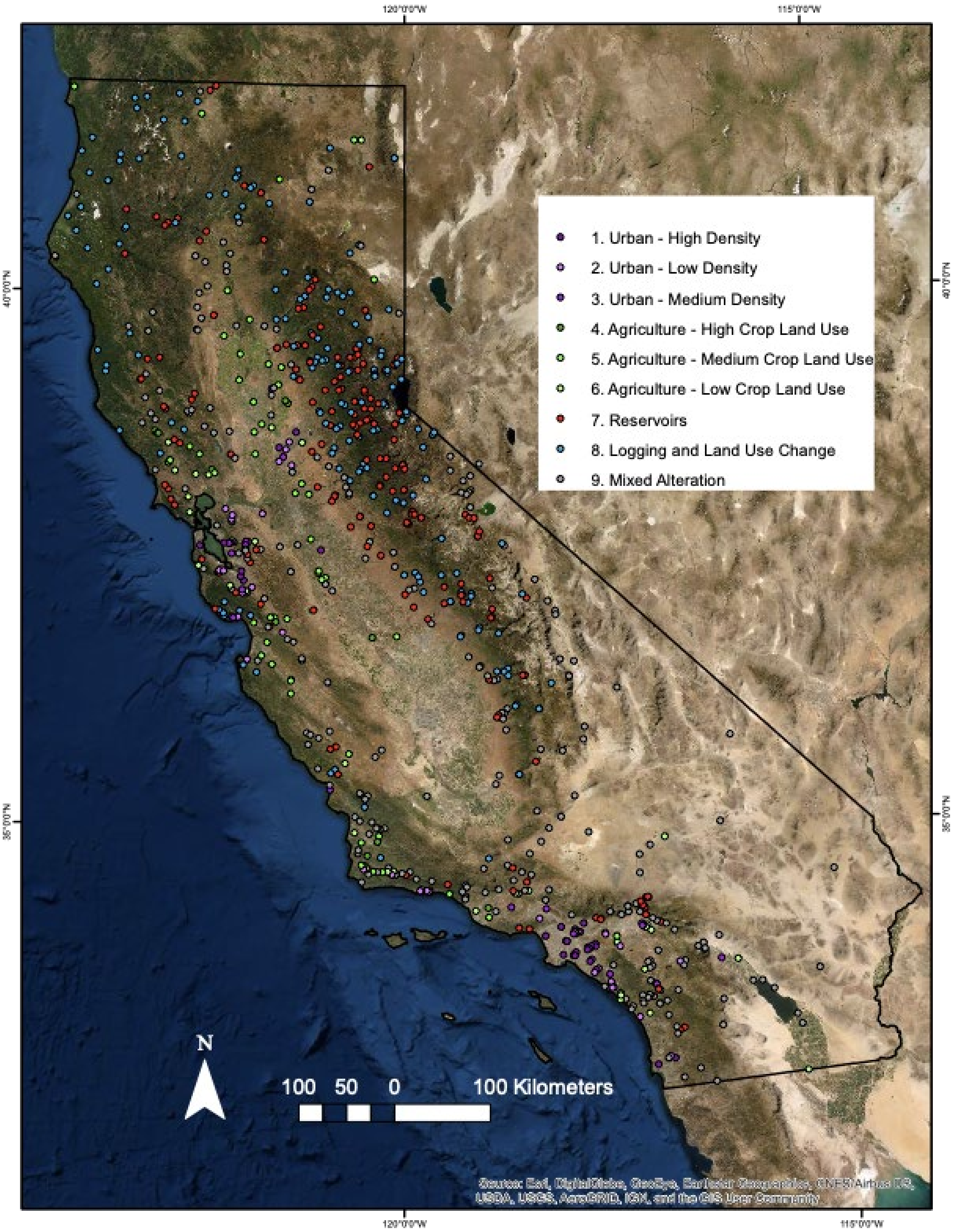
Reducing the number of gages deemed as a mixed alteration

From the previous results, most of the gauges were deemed as mixed alteration, a total of 498 (61 %). Because of this, the following heuristic analysis was performed to further classify the mixed alteration ages. First, in order to reduce the number of gages in the Mixed Alteration class, we performed a separate classification analysis that included forest variables as predictors (PctConif2006Ws – percent of conifers land use in the watershed and PctConif2006Cat – percent of conifers land use in the catchment). The classes identified from this second analysis were dominated by the forest classes and no reservoir class was identified in the analysis. This second classification confirmed the urban and agriculture classes were classified adequately, all the gages classified as agriculture and urban were consistent in both classifications. We identified the gages classified as Mixed Alteration that were classified as forest and assign this class to those gages.

Second, we combined the water storage classes (classes 7 and 8 of Table 4) into a single Reservoir class. Then, we visually inspected the 813 non-reference gauges and those that were immediately downstream of a reservoir and were not classified as water storage classes, we manually classified them as reservoirs. Table 5 shows the final impaired streamflow classification for the 813 non-reference streamflow gages. Figure 16 shows the spatial distribution of the non-reference gages classified into nine impaired flow classes distinguished across California. Table 6 explains each class.

Table 5. Classification and Frequency for each of the classes

Classes	Quantity of Gages	Frequency
1. Urban High Density (UH)	25	3%
2. Urban Low Density (UL)	40	5%
3. Urban Medium Density (UM)	27	3%
4.Agriculture (High Crop Land Use) (AgH)	12	1%
5. Agriculture (Medium Crop Land Use) (AgM)	53	7%
6. Agriculture (Low Crop Land Use) (AgL)	39	5%
7. Reservoirs (Dam)	185	23%
8. Forestland and Land Use Change (FLU)	185	23%
9. Mixed Low Alteration (Mix)	247	30%



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

Figure 16. Map of the reach-scale impaired hydrologic classification of California based on the classification tree model.

Table 6. Description of the nine altered classes

Classification	Description
1. Urban High Density (UrH)	Gages and river reaches in this class exhibit a high percent of catchment area classified as high-intensity urban land use with high population density. These sites are located in the areas of San Francisco, South Bay, Modesto, Sacramento, Los Angeles, Santa Monica, and San Diego.
2. Urban Low Density (UrL)	Gages and river reaches in this class exhibit a high percent of catchment area classified as low-intensity urban land use with low population density. These sites are located in the areas of Sacramento, Elk Grove, Walnut Creek, Palo Alto, Hollister, Santa Cruz, Monterey, Pomona, Anaheim, Oceanside, and Santa Barbara. This is the most common class related to urban land use.
3. Urban Medium Density (UrM)	Gages and river reaches in this class exhibit a high percent of catchment area classified as medium-intensity urban land use with medium population density. These sites are located in the areas of San Leandro, Santa Cruz, Los Gatos, Hayward, San Bernardino, Palm Desert, Santa Ana, and Morro Bay.
4.Agriculture - High Crop Land Use (AgH)	Gages and river reaches in this class exhibit the highest percent of watershed in cultivated crops and a high percent of watershed in irrigated agriculture. These sites are located in the areas of Santa Maria, Santa Ynez, Santa Clara, Gilroy, San Joaquin Valley, Napa Valley, Santa Rosa, Dry Creek, and Sacramento Valley.
5. Agriculture - Medium Crop Land Use (AgM)	Gages and river reaches in this class exhibit a medium percent of watershed in cultivated crops and medium percent of watershed in irrigated agriculture. These sites are located in the areas of Central Valley, Napa Valley, Dry Creek, Pajaro Valley, Salinas, Santa Maria, Santa Ynes, Smith River Valley, and Santa Clara the Modoc Plateau. This is the most common class related to agriculture.
6. Agriculture - Low Crop Land Use (AgL)	Gages and river reaches in this class exhibit a low percent of watershed in cultivated crops and low percent of watershed in irrigated agriculture. These sites are located in the areas of Russian River, Santa Rosa, Sonoma Valley, Novato, Half Moon Bay, San Clemente, Indio, Oxnard, Santa Clarita, and Salinas.

7. Reservoirs (Dam)	Gages and river reaches in this class exhibit upstream reservoir storage and high density of georeferenced dams within watershed (dams/ square km) based on the National Inventory of Dams. These sites are located in areas of the Sierra Nevada foothills called rim dams (Shasta, Oroville, Folsom, New Bullard's Bar, New Melones, Friant dam, Lake McClure, Pine Flat), high elevation hydropower reservoirs in the Sierra Nevada, water supply reservoirs in the Trinity, Klamath, Russian, Eel, Salinas, and Santa Ynez, and small storage reservoirs in Southern California and along the coast of California.
8. Forestland and Land Use Change (FLU)	Gages and river reaches in this class are located in the forestland area of California dominated by trees generally greater than 5 meters tall, and typically with the forest cover greater than 20% of total vegetation. They exhibit land use change and the associated streamflow alteration by : • deforestation, logging and clear cutting due to timber extraction that produces reduction in time of concentration, increase in peak flows, erosion of soil, sediment transport and degradation of water quality, • cattle grazing and stocking that produce beneficial services by reducing understory biomass that prevent devastating fires, however if this activity is not well managed in the riparian corridor and meadows it can produce change in the hydrology, straighten and deepen of channels with the subsequent desiccation of meadows, erosion of channel banks, and degradation of water quality; and • land use change from forest to agriculture for cannabis production that produces similar streamflow alteration to clear cutting with the potential of runoff of pesticides and fertilizers used in high slope agriculture production areas. These sites are located in all national forests in California (Klamath, six rivers, Trinity, Shasta – Trinity, Lassen, Plumas, Tahoe, El Dorado, Stanislaus, Sierra Sequoia, Mendocino, Los Padres, Angeles, San Bernardino, and Cleveland) and the areas close to national forest with high forest density, such as land holdings of timber industry.
9. Mixed Low Alteration (Mix)	Gages and river reaches in this class are located in the Central Valley, California's deserts and dry land areas (Paso Robles and East San Joaquin).

	In the Central Valley, they are located in the foothills along the perimeter of the Central Valley, in the areas of transition between altered classes, for example from dam alteration in the Sierra Nevada to agriculture in Central Valley, where the main impairment is land-use modification, such as native vegetation to orchards or urban development, or they are located in the catchments of small tributaries, e.g. Antelope, Creek, Battle Creek, Mariposa Creek where similar land-use change occurs. This impaired class occurs in the dry land areas of California, namely Paso Robles and the east side San Joaquin basin, where the main impairment is land-use modification from native lands use to agriculture. In addition, this impaired class occurs in California's deserts (Mojave Desert, Death Valley, Owens) where there might be few impairments, but high rates of natural erosion and expansion of urban areas are the main impairment. This is the largest altered class in California.

5.2.4 Flow Regime- Dimensionless Non-reference Hydrographs (DNH)

A dimensionless reference hydrograph, DNH, is a scalable representation of reference hydrology based on streamflow data from unimpaired streamflow gauges in a hydrologic stream class. The y-axis is expressed in dimensionless units by dividing daily streamflows by average daily streamflow for that water year. Figure 17 shows the (DNHs) of the nine impaired streamflow classes.

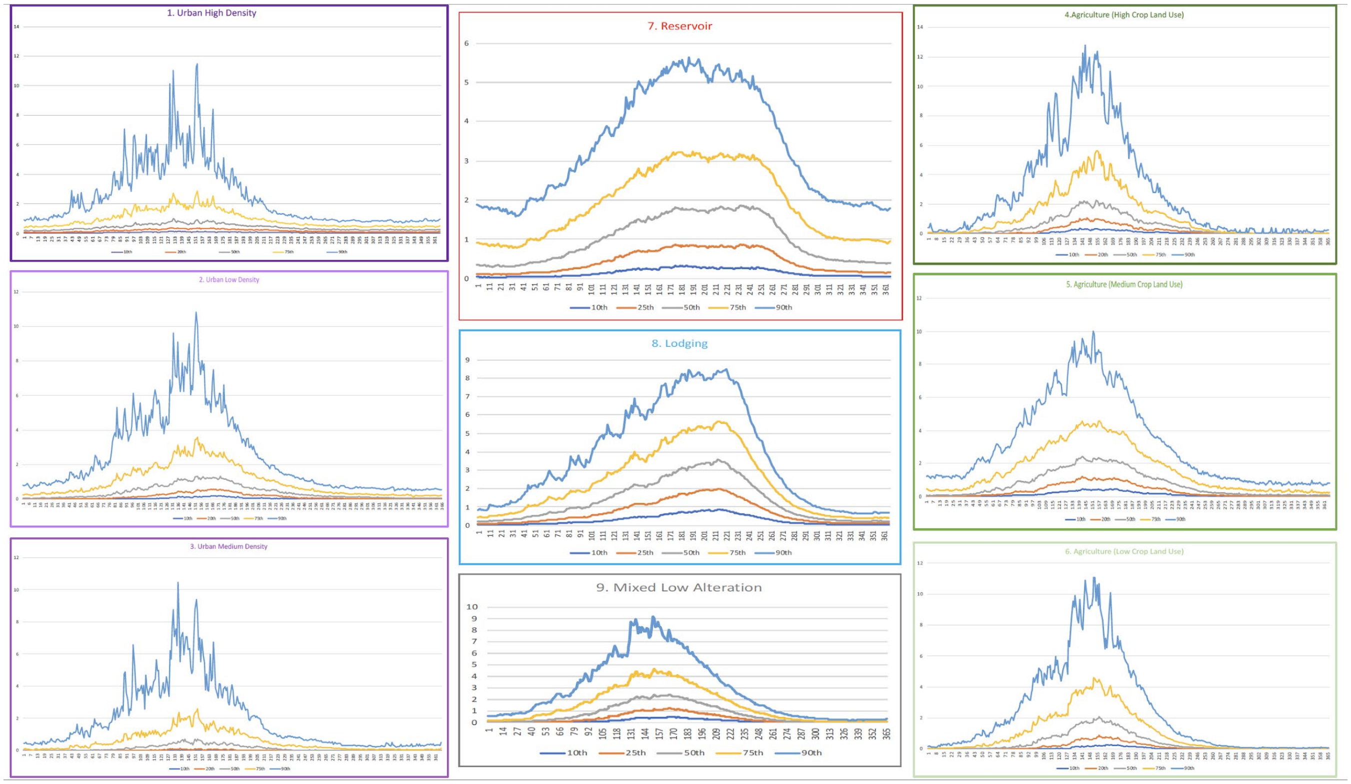


Figure 17. Dimensionless reference hydrographs for the nine altered classes

5.3 Spatial Prediction

Predictor variables are the parameters used to classify river alteration. The statistical methods used in the previous sections resulted in nine impaired flow classes across California. The eight predictor variables were estimated at every 200-m river reach throughout the entire river network. The nine impaired streamflow classes were predicted in the river network using three machine-learning algorithms: random forest, support vector machine, artificial neural networks. Each model was trained using a ten-fold training data set, meaning the model was trained with 90% of the data end evaluated with the remaining 10% of the data left out. A cross validation technique was used to determine the algorithm that had the best performance, in essence, what was the percentage of the sites left out that were classified correctly.

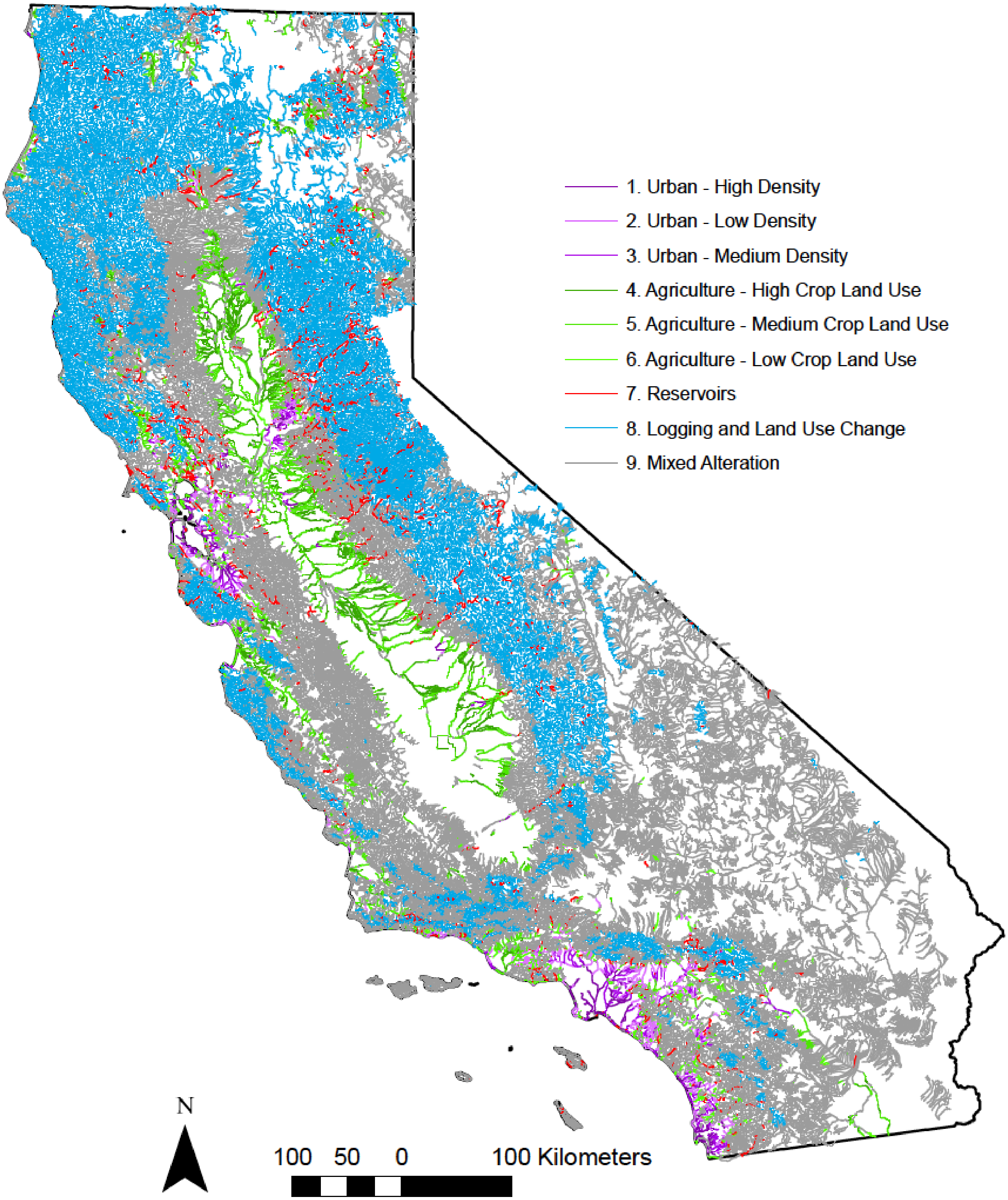
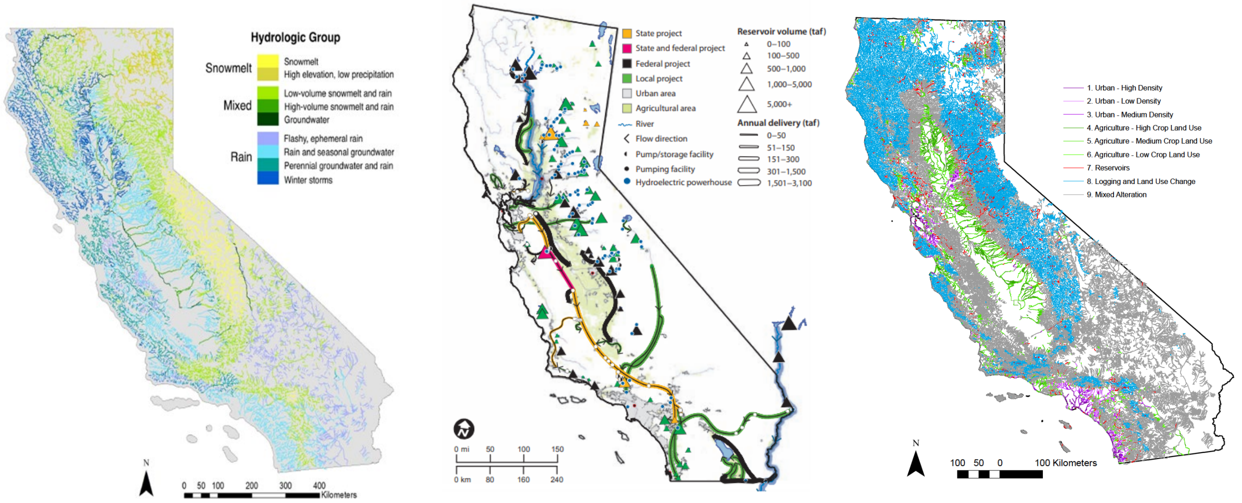

1. Urban - High Density
2. Urban - Low Density
3. Urban - Medium Density
4. Agriculture - High Crop Land Use
5. Agriculture - Medium Crop Land Use
6. Agriculture - Low Crop Land Use
7. Reservoirs
8. Logging and Land Use Change
9. Mixed Alteration
N
100 50 0 100 Kilometers

Figure 18. Spatial prediction of the Streamflow for the nine altered classes

6. Conclusions

Human impact on ecosystems can be viewed in various ways, but refusing to acknowledge the magnitude of streamflow alteration due to human activity will pose greater threats in the future. What is the future river ecosystem in human-dominated landscapes? In California, even worldwide, flow regimes are being modified by various anthropogenic impacts inducing higher risks to the health of river ecosystems and the loss of ecosystem services that are valuable by society. The survival, viability and evolution of river ecosystem is threatened by human activity and dependent of actions of society as a whole to restore and protect this fragile resources. The side effects of the maltreatment and lack of management now threatens our future and that of biological diversity. The ability to evaluate natural stream flow, which is not altered by human activities, would be enhanced by the existence of a nationally consistent and up-to-date database of gages in relatively undisturbed watersheds. Conservation philosophy, science, and practice must be framed against the reality of human-dominated ecosystems, rather than the separation of humanity and nature underlying the modern conservation movement (Western 2001). The consequence of one harms the other, both are interlinked and affect the future of ecosystems and everyone's well-being. The future contains further loss of species diversity and wild habitat, accelerated erosion, sediments, along with the loss of ecosystem services.



(a) (b) (c)

Figure 19. Map of the natural streamflow class (a), main water infrastructure (b) and impaired streamflow classification (c)

California is an altered state where the spatial extent and type of alteration were not known before and now it is. There are 4 types of alterations such as Agriculture, Dams, Urban, and Forestland and Land Use Change, that make up 69% of the Non-Reference gages. The largest alteration classification is Mixed Alteration that is located in places where few land-use changes or in California's deserts. These classification data sets (Natural and Altered Streamflow Classification) can help bridge the gap between the natural flow regime, locations of current impairment, and what can be done to reduce the alteration between the two ideas. Ultimately, this is one of the pieces of a large puzzle for conserving and restoring river ecosystems.

7. References

Acreman, M. C., and M. J. Dunbar. "Defining Environmental River Flow Requirements – a Review." Hydrology and Earth System Sciences, vol. 8, no. 5, 2004, pp. 861–876., doi:10.5194/hess-8-861-2004.

Agglomerative Algorithm. Journal of Classification 31 : 274–295. DOI: 10.1007/s00357-014-9161-z [online] Available from: http://arxiv.org/abs/1111.6285 (Accessed 21 March 2018)

Arcuni, P. (2019, January 16). Rivers in the Sky: What You Need to Know About Atmospheric River Storms. Retrieved from https://www.kqed.org/science/1935067/rivers-in-the-sky-what-you-need-to-know-about-atmospheric-river-storms

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26 : 32–46. DOI: 10.1111/j.1442-9993.2001.01070.pp.x [online] Available from: http:

//onlinelibrary.wiley.com/doi/10.1111/j.1442-9993.2001.01070.pp.x/abstract (Accessed 10 October 2017)

Arthington, Angela H. Environmental Flows: Saving Rivers in the Third Millennium. University of California Press, 2012.

Chapman, Deborah. “Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring - Second Edition.” *Water Quality Assessments*, 1996, pp. 1–609., doi:10.4324/noe0419216001.ch7.

Clarke KR. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18 : 117–143. DOI: 10.1111/j.1442-9993.1993.tb00438.x [online] Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1442-9993.1993.tb00438.x/abstract (Accessed 10 October 2017)

De’ath G, Fabricius KE. 2000. Classification and Regression Trees: A Powerful yet Simple Technique for Ecological Data Analysis. Ecology 81 : 3178–3192. DOI: 10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

Falcone, J.A., D.M. Carlisle, D.M. Wolock, and M.R. Meador, 2010. GAGES: A Stream Gage 706 Database for Evaluating Natural and Altered Flow Conditions in the Conterminous United 707 States. Ecology 91(2):621-621. DOI: 10.1890/09-0889.1

Faustini, John. “River Classification: An Overview.” *A White Paper Submitted to the Southeast Aquatics Resource Partnership*, US Fish and Wildlife Service, Region 4, 15 June 2012, southeastaquatics.net/resources/sifn-resources/documents/general-sarp-instream-flow-resources/river-classification-overview-by-john-faustini.

Franklin, Kimberly. “Investigation.” *Reconciliation Ecology*, Tree of Life Project, 2007, tolweb.org/treehouses/?treehouse_id=4558.

Hanak, E., J. Lund, A. Dinar, B. Gray, R. Howitt, J. Mount, P. Moyle, and B. Thompson. (2011) Managing California’s Water: From Conflict to Reconciliation. San Francisco: Public Policy Institute of California.

Hill, R.A., M. Weber , S. Leibowitz , T. Olsen , and D.J. Thornbrugh. The Stream-Catchment (StreamCat) Dataset: A database of watershed metrics for the conterminous USA. JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION. American Water Resources Association, Middleburg, VA, USA, 9, (2015).

James, A. Falcone, Daren M. Carlisle, David. M. Wolock and Micheal R. Meador. (2010). GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. Ecology Society of America. Ecology Archives E091-045.

Kruskal JB. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29 : 1–27. DOI: 10.1007/BF02289565 [online] Available from: https://link.springer.com/article/10.1007/BF02289565 (Accessed 7 February 2018)

Lane, B.A., Sandoval-Solis, S., Stein, E.D., Yarnell, S.H., Pasternack, G.B. and Dahlke, H.E. (2018). Beyond metrics? The role of hydrologic baseline archetypes in environmental water management. Journal of Environmental Management. DOI: 10.1007/s00267-018-1077-7

Lane, B.A., Sandoval-Solis, S., Stein, E.D., Yarnell, S.H., Pasternack, G.B. and Dahlke, H.E. (2018). Beyond metrics? The role of hydrologic baseline archetypes in environmental water management. Journal of Environmental Management. DOI: 10.1007/s00267-018-1077-7

Lund, J., E. Hanak, W. Fleenor, W. Bennett, R. Howitt, J. Mount, and P. Moyle, Comparing Futures for the Sacramento-San Joaquin Delta, University of California Press, Berkeley, CA, February 2010.

Murtagh F, Legendre P. 2014a. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification 31 : 274–295. DOI: 10.1007/s00357-014-9161-z [online] Available from: https://link.springer.com/article/10.1007/s00357-014-9161-z (Accessed 9 October 2017)

Murtagh F, Legendre P. 2014b. Ward's Hierarchical Clustering Method: Clustering Criterion and

Patterson, N.K., Lane, B.A., Sandoval-Solis, S., Pasternack, G.B., Yarnell, S.M., Qiu, Y. (2020). A hydrologic feature detection algorithm to quantify seasonal components of flow regime. J. of Hydrology. Accepted Feb/2020

Richter, B. D., & Richter, H. E. (2000). Prescribing flood regimes to sustain riparian ecosystems along meandering rivers. Conservation Biology, 14(5), 1467–1478.

Richter, Brian D., et al. "Ecologically Sustainable Water Management: Managing River Flows For Ecological Integrity." *Ecological Applications*, vol. 13, no. 1, 2003, pp. 206–224., doi:10.1890/1051-0761(2003)013[0206:eswmmr]2.0.co;2.

Rosenzweig, M. (2014, June). Tactics for Conserving Diversity: Global Vertebrate Patterns Point the Way.Public Talk presented at UC Davis Center for Watershed Sciences and explored a "reconciliation" approach to improving California's aquatic habitat. https://californiawaterblog.com/2014/07/23/reconciling-ecosystem-and-economy/

Rumsey, D. (1979). David Rumsey Map Collection. Retrieved January 24, 2010, from https://www.davidrumsey.com/blog/2010/1/21/landmark-1979-california-water-atlas-debuts-online

Scanlon, John. "Flow : the Essentials of Environmental Flows." 2003, doi:10.2305/iucn.ch.2003.wani.2.en.

Sklar, Fred H., and Joan A. Browder. "Coastal Environmental Impacts Brought About by Alterations to Freshwater Flow in the Gulf of Mexico." Environmental Management, vol. 22, no. 4, Jan. 1998, pp. 547–562., doi:10.1007/s002679900127.

Swanson, Steve. "Indicators of Hydrologic Alteration (IHA)." *Encyclopedia of Hydrological Sciences*, 2005, doi:10.1002/0470848944.hsa249.

Tukey JW. 1991. The Philosophy of Multiple Comparisons. Statistical Science 6 : 100–116. [online] Available from: http://www.jstor.org/stable/2245714 (Accessed 21 March 2018)

Ward JHJ. 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58 : 236–244. DOI: 10.1080/01621459.1963.10500845 [online] Available from: https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10500845 (Accessed 21 March 2018)

Western (2001) Western D. Human-modified ecosystems and future evolution. Proceedings of the National Academy of Sciences USA. 2001;98(10):5458–5465. doi: 10.1073/pnas.101093598.

Yarnell, Sarah M., et al. "Ecology and Management of the Spring Snowmelt Recession." *BioScience*, vol. 60, no. 2, 2010, pp. 114–127., doi:10.1525/bio.2010.60.2.6.

Yarnell, Sarah M., et al. "Functional Flows in Modified Riverscapes: Hydrographs, Habitats and Opportunities." *BioScience*, vol. 65, no. 10, May 2015, pp. 963–972., doi:10.1093/biosci/biv102.