Work Plan: Petaluma River Watershed Hydrology Model Development

SUBMITTED TO:

State Water Resources Control Board
1001 I Street, 14th Floor
Sacramento, CA 95814

PREPARED BY:

Paradigm Environmental
9320 Chesapeake Drive, Suite 100
San Diego, CA 92123

FINAL
AUGUST 21, 2025

THIS PAGE INTENTIONALLY LEFT BLANK

ACRONYMS

3DEP	3D Elevation Program
ALWU	Agricultural Land use & Water Use
ASCE-PM	American Society of Civil Engineers version of the Penman-Monteith equation
CAL FIRE	California Department of Forestry and Fire Protection
CDEC	California Data Exchange Center
CDFW	California Department of Fish and Wildlife
CDL	Cropland Data Layer
CDT	California Department of Technology
CIMIS	California Irrigation Management Information System
CoCoRaHS	Community Collaborative Rain Hail and Snow Network
COOP	Cooperative Obsrver Program
DEM	Digital Elevation Model
DWR	California Department of Water Resources
EOL	Earth Observing Laboratory
ET	Evapotranspiration
ET_0	Reference evapotranspiration
eWRIMS	Electronic Water Rights Information Management System
FEMA	Federal Emergency Management Agency
GHCN	Global Historical Climatology Network
GIS	Geographic Information System
GSA	Groundwater Sustainability Agency
GSP	Groundwater Sustainability Plan
HRU	Hydrologic Response Unit
HSG	Hydrologic Soil Group
HSPF	Hydrologic Simulation Program - Fortran
HUC	Hydrologic Unit Code
LCD	Local Climate Data
LSM	Land Surface Model
LSPC	Loading Simulation Program in C++
MODFLOW	USGS Modular Hydrologic Model
MRLC	Multi-Resolution Land Consortium
NCEI	National Centers for Environmental Information
NHD	National Hydrography Dataset
NID	National Inventory of Dams
NLCD	National Land Cover Database
NLDAS	North American Land Data Assimilation System
NOAA	National Oceanic and Atmospheric Administration

NRCS	Natural Resources Conservation Service
NSE	Nash-Sutclife Model Efficiency Coefficient
PBIAS	Percent Bias
PEVT	Potential Evapotranspiration
POD	Point of Diversion
PRISM	Parameter-elevation Regressions on Independent Slopes Model
RAWS	Remote Automated Weather Stations
SFEI	San Francisco Estuary Institute
SGMA	Sustainable Groundwater Management Act
SSURGO	Soil Survey Geographic Database
STATSGO2	State Soil Geographic Database
SWAT	Soil and Water Assessment Tool
SWRCB	State Water Resources Control Board
USDA	United States Department of Agriculture
USFS	United States Forest Service
USGS	United States Geological Survey
WBD	Watershed Boundary Dataset

Contents

Figures

Tables

1 INTRODUCTION

1.1 Project Objectives

In April 2021, Governor Gavin Newsom issued a state of emergency proclamation for specific watersheds across California in response to exceptionally dry conditions throughout the state. The April 2021 proclamation, as well as subsequent proclamations, directed the State Water Resources Control Board (Water Board) to address these emergency conditions to ensure adequate, minimal water supplies for critical purposes. To support Water Board actions to address emergency conditions, hydrologic modeling and analysis tools are being developed to contribute to a comprehensive decision support system that assesses water supply and demand and the flow needs for watersheds throughout California.

This work plan presents the available data and methodology that will be used to develop a hydrologic model of the Petaluma River watershed. This model will use historical records of precipitation, temperature, and evapotranspiration (ET) for simulation of processes associated with surface runoff, infiltration, interflow, and groundwater flow. The final calibrated model will be used to evaluate scenarios including current hydrologic conditions, water allocation, changes in demand, and the impact of extreme events such as droughts or atmospheric rivers.

1.2 Watershed Background

The Petaluma River is one of the major tributaries to the San Pablo and San Francisco Bays. The Petaluma River watershed, identified by a ten-digit hydrologic unit code (HUC), is part of the San Pablo Bay drainage area, which shares a boundary with the Tomales-Drake Bays watershed to the west, the Russian River watershed to the north, and the Sonoma Creek watershed to the east. The Petaluma River (HUC-10: 1805000206) is an estuarine watershed that drains approximately 198 square miles and is made up of seven main catchments, the largest being Petaluma River (HUC-12: 180500020601) (Figure 11). The Petaluma River flows south-east through the Sonoma and Marin County boundaries. The headwaters and tributaries of the Petaluma River originate on the southwest slopes of Sonoma Mountain, the southern slopes of Mecham Hill, and the eastern slopes of Weigand's Hill and Mount Burdell. The river flows through the city of Petaluma, then continues south-southwest into the northwest portion of San Pablo Bay (SRCD 2025).

The Petaluma River watershed ranges in elevation from below sea level at Point San Pedro in the south to over 700 meters at the northern most portion of the watershed near Mount Sonoma. The watershed has a Mediterranean climate with distinct wet and dry seasons with a mean annual precipitation total of 33.4 inches with higher rainfall amounts occurring in the higher elevations and lower amounts in the lower elevation regions (SFEI 2025). The watershed is dominated by grassland or herbaceous landcover at 48%. The second most prominent land use type is developed, which covers around 23% of the watershed.

The Petaluma River watershed represents an important habitat for native aquatic species and spawning ground for anadromous fish, especially chinook salmon and steelhead trout. However, there have been substantial declines in salmonid populations over time. The decline in anadromous fish populations within the Petaluma watershed has been linked to increases in sedimentation, habitat degradation, urbanization, channel modification, elevated stream temperatures, and the presence of warm-water predators (NOAA 2022). The Petaluma River has appeared on the Clean Water Act’s 303(d) list of impaired waters since at least 1975 due to elevated levels of fecal indicator bacteria, such as E. coli and enterococcus (CRWQB 2018). These bacteria indicate the presence of pathogens from endothermic animal waste, posing health risks to individuals engaging in water recreation activities. Additionally, the river has been identified as impaired by high nutrient concentrations, leading to eutrophication—a process that depletes oxygen levels and degrades aquatic habitats. These factors led to the development of a Total Maximum Daily Load (TMDL) for bacteria in 2019 (SFBRWQCB 2020).

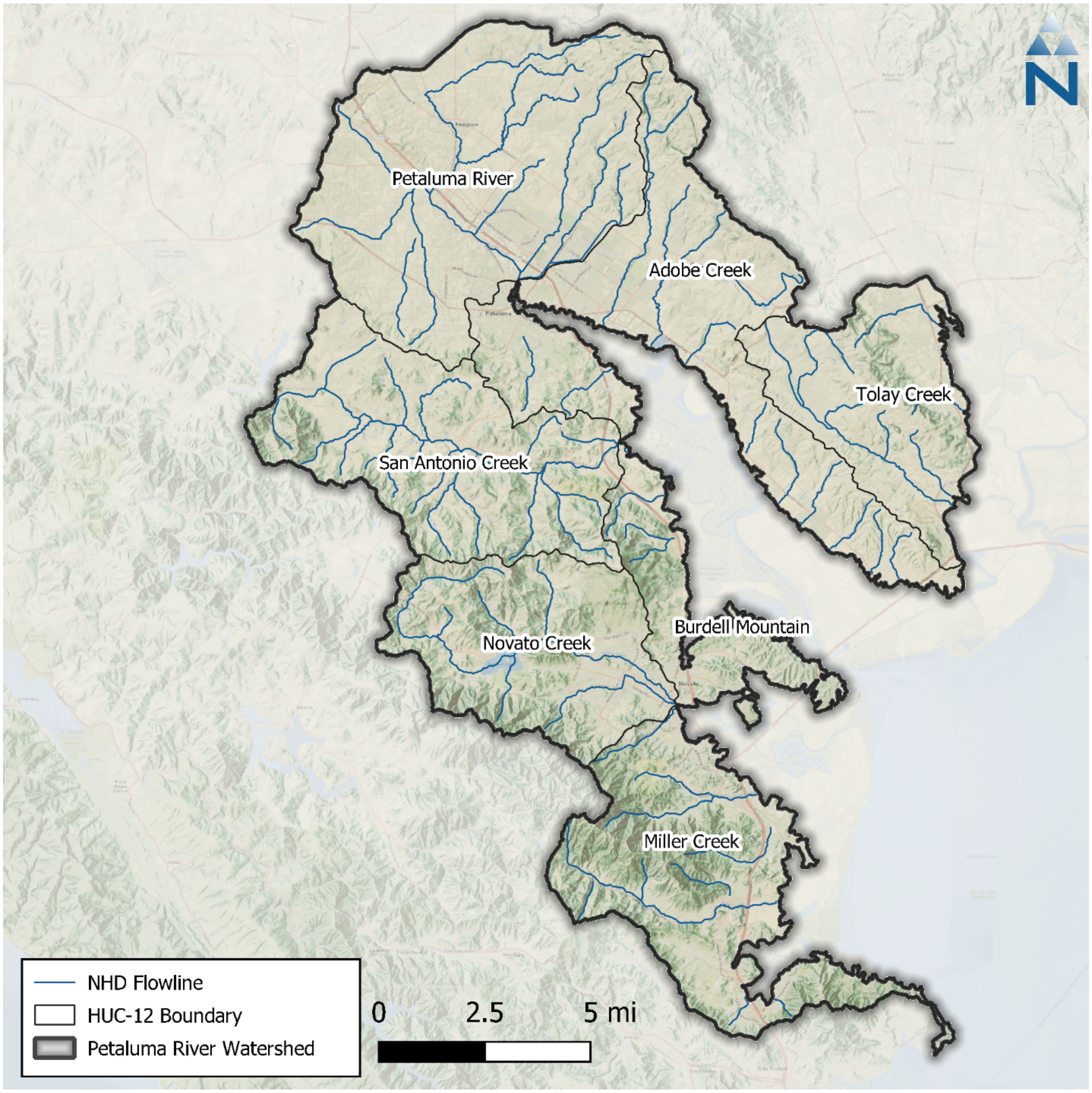

Figure 1-1. The Petaluma River watershed.

1.3 Leveraging Previous Modeling Efforts

The San Francisco Estuary Institute (SFEI) has performed hydrologic model development in the San Francisco Bay drainage area, which includes the Petaluma River watershed (Zi et al. 2021 and 2022). Figure 1-2 shows the location of the Petaluma River watershed within the larger extent of the SFEI model domain. SFEI's modeling work utilizes the Loading Simulation Program in C++ (LSPC) (Shen, Parker, and Riverson 2005) for long term continuous simulation with the objective of supporting management decisions regarding stormwater runoff and pollutant loads including mercury, PCBs, and sediment. This model was constructed at the regional scale to address San Francisco Bay-wide pollutant loading issues; therefore, the focus of model configuration and calibration differs slightly from the needs of this Petaluma River watershed-specific hydrologic study.

As part of the screening process of available data, the Petaluma River subset of the SFEI model was assessed to see how it could be used to support modeling for this study. Overall, the existing model offers a strong foundation, however, certain elements of that model will need some modifications to address the hydrology simulation accuracy and water budgeting needs of this project. The Petaluma Watershed boundary contains 15 catchments from the SFEI model. Section 7.2 provides additional discussion on proposed modifications to the SFEI model for this project.

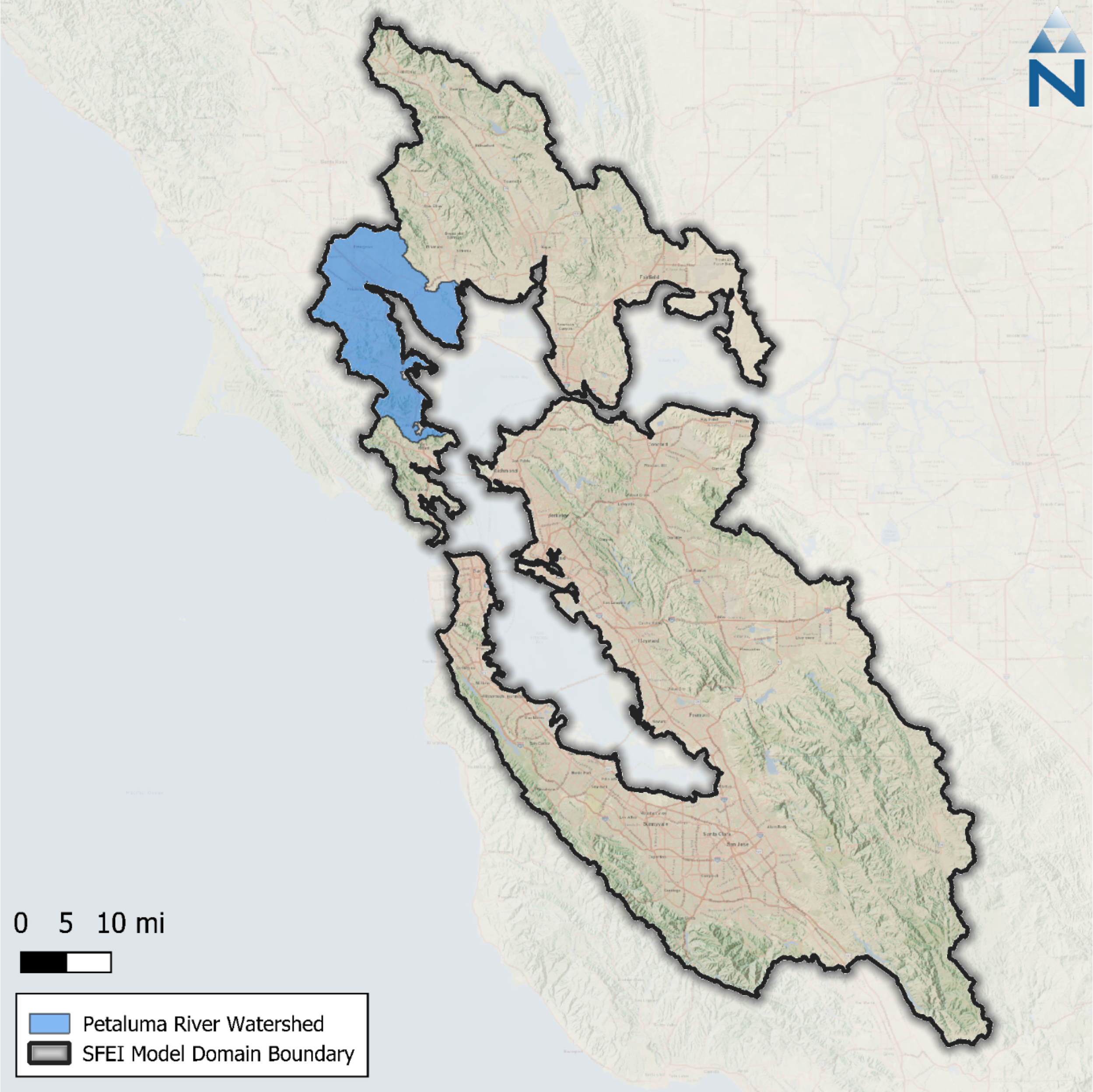

Figure 1-2. Map of the SFEI regional model domain, which includes the Petaluma River watershed (Zi et al. 2021).

1.4 Model Approach

The primary goal of this work plan is to outline an approach with sufficient robustness to support an analytical assessment of the Petaluma River watershed. This is presented first through a comprehensive inventory of available meteorological, hydrological, and geographic information system (GIS) data available for the Petaluma River watershed. The data compilation and assessment processes are outlined below and aim to highlight any existing data gaps that create limitations for the analysis. Based on the available data, any data gaps are identified that may be filled through additional outreach, data collection efforts, or noted as points of uncertainty in the model documentation.

This hydrologic analysis is based on a model development process that has been a tested platform for gaining valuable information and insight about hydrologic systems. The model development process proposed is an iterative and adaptive cycle that improves understanding of the system over time as better information becomes available. Figure 1-3 is a conceptual schematic of the proposed model development cycle, which is represented as circular as opposed to linear. The cycle is best summarized by the following six interrelated steps:

1. **Assess Available Data**: Data for source characterization, trends analysis, and defining modeling objectives.
2. **Delineate Model Domain**: Model segmentation and discretization needed to simulate streamflow at temporal and reach scales appropriate for assessing supply and demand.
3. **Set Required Model Inputs**: Spatial and temporal model inputs defining the appropriate hydrologic inputs and outputs.
4. **Represent Processes (Calibration)**: Adjustment of model rates and constants to mimic observed physical processes of the natural system.
5. **Confirm Predictions (Validation)**: Model testing with data not included in the calibration to assess predictive ability and robustness.
6. **Assess Applicability for Scenarios**: Sometimes the nature of modeled responses can indicate the influence of unrepresented physical processes in the modeled system. Sometimes that can be resolved with minor parameter adjustments, while other times the assessment exposes larger data gaps. A well-designed model can be adapted for future applications as new information about the system becomes available. Depending on the study objectives, data gaps sometimes provide a sound basis for future data collection efforts to refine the model. New information may require minor parameter adjustments affecting the configuration or calibration.

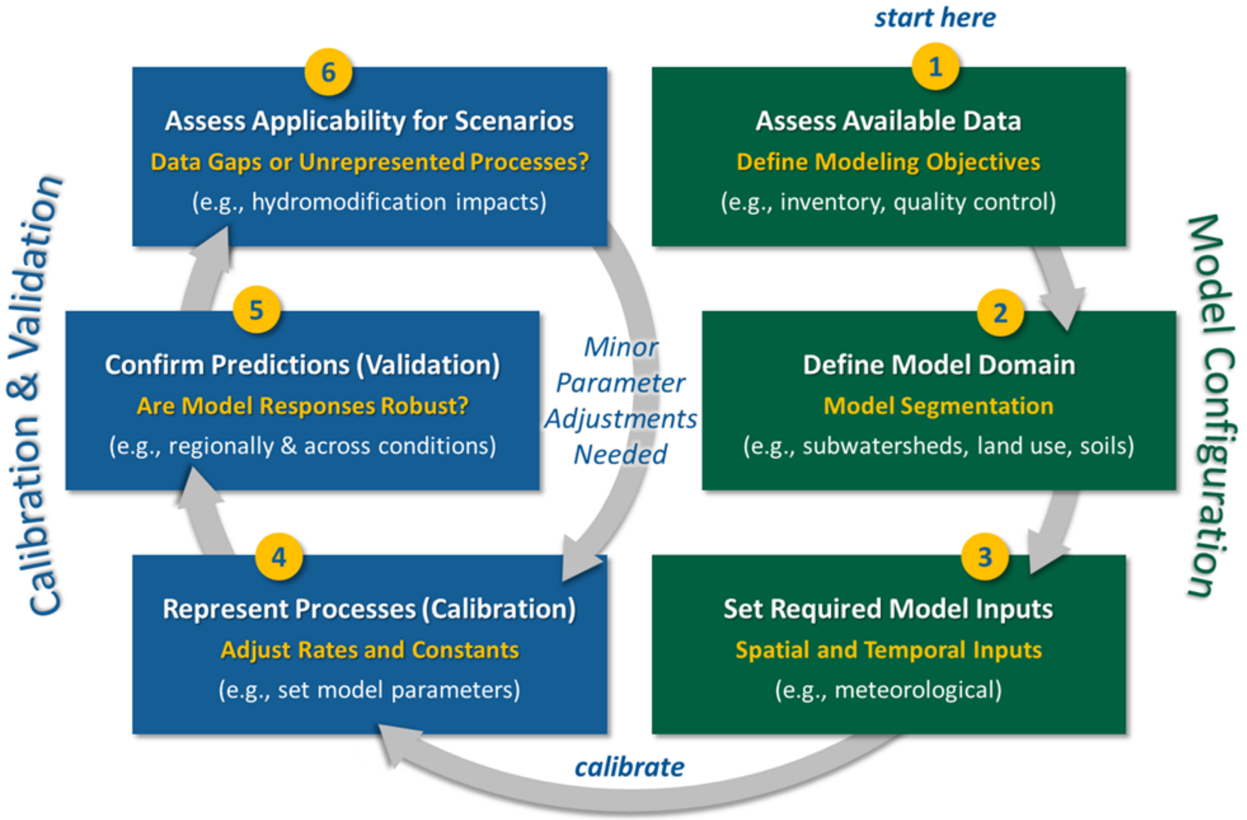

Figure 1-3. Conceptual schematic of model development cycle proposed for assessing instream flow needs in the Petaluma River watershed.

1.5 Data Availability

Table 1-1 through Table 1-4 present an inventory of the initial data collected that will form the basis of this modeling workplan These datasets were compiled from readily available sources, primarily those publicly available and published online by state and federal agencies. The data in the tables is organized by data type listed below and further described in the following sections.

- **Meteorology Datasets**: Time series that represent water balance inputs and outputs to the watershed primarily from precipitation and ET. These time series are often used as forcing functions for hydrologic models.
- **Surface & Groundwater Datasets**: Datasets describing stream flow, groundwater, water use, and stream conditions for the Petaluma River. Time series observations of instream responses for the Petaluma River are often used as calibration and validation datasets for hydrologic models.
- **Geospatial Datasets**: Spatial datasets describing the landscape of the Petaluma River watershed. These datasets include physical properties (e.g., soil, land cover, elevation).

.

Table 1-1. Inventory of meteorology datasets

Source	Dataset	Data Range	Description	Model Use
National Centers for Environmental Information (NCEI)	Global Historic Climate Network (GHCN)	--	Daily precipitation and temperature data (varied data quantity/quality).	Rainfall input boundary time series.
NDEI	Local Climatological Data (LCD)	--	Hourly precipitation, temperature, wind speed, dewpoint, cloud cover.	Rainfall input boundary time series.
Remote Automated Weather Stations (RAWS)	Hourly Climate Data	--	Meteorological records for six stations.	Climate data boundary time series.
California Data Exchange Center (CDEC)	Precipitation, Temperature	--	Meteorological records are available for one station.	Rainfall input boundary time series.
PRISM Climate Group	AN81m Monthly	1900- Present	4-km grid resolution time series of precipitation (1900 – present).	Rainfall time series QA; address rainfall data gaps.
North American Land Data Assimilation System (NLDAS)	NLDAS-2 Forcing Data	1979 - Present	1/8th-degree grid resolution hourly time series of precipitation and other surface parameters (e.g., potential evapotranspiration (PEVT), and solar radiation).	Rainfall hourly distributions; address rainfall data gaps. Daily potential evapotranspiration totals × hourly solar radiation distributions.
Earth Observing Laboratory (EOL)	Daily/Hourly Gridded Precipitation	--	Various gridded precipitation time series; both daily and hourly time steps.	Rainfall hourly distributions; address rainfall data gaps.
California Irrigation Management Information System (CIMIS)	Reference ET	1990 – Present	Relative ET spatial zones and monthly scaling factors. There is also a grid-based model data product.	Deriving PEVT input forcing time series; estimation of irrigation demand.
OpenET	OpenET CONUS Ensemble Monthly ET	2016 - 2024	Satellite-based estimates (30-m res) of observed monthly ET for the CONUS; data is bias corrected against observational weather station networks.	Parameterization & evaluation of ET; estimation of irrigation demand.

Table 1-2. Inventory of surface water datasets

Category	Scale	Source	Dataset	Date Range	Description	Model Use	Link
Streamflow	Local	United States Geological Survey (USGS)	Stream Station Discharge	1929 – Current	Observed streamflow at one active location on Novato Creek.	Hydrology calibration.	LINK
		National Inventory of Dams (NID)	Dams of the United States	Current	Locations of the dams across the United States.	Hydrology calibration	LINK
Habitat	Local	CDFW	Total Maximum Daily Load for Bacteria in the Petaluma River Watershed	2020	Report that documents salmonid habitat and stream conditions under the bacteria TMDL.	Hydrology calibration & validation.	LINK
Water Budget	State	DWR	Well Completion Reports	Current	Well completion logs and reports.	Water budget.	LINK
			Interconnected Surface Water	2008	Two (2) river flow CDEC stations and six (6) rain CDEC stations identified as interconnected.		LINK
		SWRCB eWRIMS	Water Rights Points of Diversion	Current	Locations where water is being drawn from a surface water source such as a stream or river.		LINK
			Water Rights Overview Report	Current	This report will provide counts of various entities such as Applications, Registrations, Petitions etc. that will reflect the progress in processing such entities as of current date.		LINK
			Annual Water Use Report	1906 – 2023	Annual reports that provide monthly diversion data for various entities such as Applications, Registrations, Petitions, etc.		LINK
		DWR	Agricultural Land and Water Use Estimates	1998 – 2015	Water use estimates by various planning units.		LINK
		CDT	Water Districts	2022	Boundaries of all public water agencies in California.		LINK
			California Drinking Water System Area Boundaries	2023	Public California drinking water systems and state small drinking water system boundaries and information.		LINK

Table 1-3. Inventory of geospatial datasets

Category	Scale	Source	Dataset	Date Range	Description	Model Use	Link
Watershed Boundaries	National	USGS	Watershed Boundaries (WBD)	2023	Hydrologic unit boundaries to the 12-digit (6th level).	Model segmentation	LINK
Hydrology	National	USGS	National Hydrography Dataset (NHD) Plus High-Resolution National Release 1	2023	The NHDPlus HR combines the NHD, 3DEP DEMs, and WBD to create a stream network with linear referencing.		LINK
			NHD Best Resolution	2023	1:24,000; represents reaches and other network elements.		LINK
Soil	National	USDA NRCS	Grided Soil Survey Geographic Database (SSURGO)	2022	State-wide, 10-meter raster grid approximating the SSURGO vector dataset.	Represent infiltration process within land segments.	LINK
			Digital General Soil Map of the United States (STATSOGO2)	2016	State-wide, 10-meter raster grid approximating the SSURGO vector dataset.		LINK
Surficial Geology	National	USGS	The State Geologic Map Compilation (SGMC)	2017	1:1,000,000: Vector-based, state geologic map database.	As needed, hydrologic process with land segments.	LINK
Land Cover	National	MRLC	National Land Cover Dataset (NLCD) Land Cover	2021	Broad, 30 m grid-based land characterization. Differentiates developed land from coarse classifications of forest, cropland, wetlands, etc.	Land segment representation.	LINK
			NLCD Fractional Imperviousness	2021	Broad, 30-meter grid-based land characterization. Represent percent impervious area within raster cells.		LINK
Land Use	State	DWR	Statewide Crop Mapping	2020	Polygons attributed with DWR crop categories.	Identify crop distributions; estimate irrigation demand.	LINK
Vegetation	National	MRLC	Tree Canopy Cover	2021	Percent tree canopy estimates for each 30-meter pixel across all land covers and types.	Land segment representation.	LINK

Category	Scale	Source	Dataset	Date Range	Description	Model Use	Link
	State	USFS	Existing Vegetation	2018	1:24,000 to 1:100,000: Existing vegetation mapping.	As necessary, additional vegetation types for model land segments.	LINK
Agriculture & Crop Cover	National	USDA	Cropland Data Layer	2022	30-meter grid-based crop-specific land cover data layer.	Identify crop distributions; estimate irrigation demand.	LINK
Timber Harvesting	National	USDA	Timber Harvests	1820 - Present	Area planned and accomplished acres treated as a part of the timber harvest program of work.	Representing changes in land cover due to timber harvest activities.	LINK
	State	CAL FIRE	CAL FIRE Nonindustrial Timber Management Plans TA83	1991 - Present	Timber management plans.		LINK
			CAL FIRE Notices of Timber Operations TA83	1991 - Present	Notice of Timber Operations accepted by CAL FIRE.		LINK
			CAL FIRE Working Forest Management Plans TA83	2019 - Present	Working forest management plans approved by CAL FIRE.		LINK
Fire Perimeters & Burn Areas	State	CAL FIRE	California Fire Perimeters	1950 - Present	Wildfire perimeters.	Representing changes in land cover due to forest fire activities.	LINK
			Prescribed Burns	1950 - Present	Prescribed burns perimeters.		LINK
Elevation	National	USGS	USGS ten-meter resolution digital elevation model (DEM)	2020	10-meter resolution digital elevation model (DEM) produced through the 3D Elevation Program (3DEP).	Land segment representation.	LINK

Table 1-4. Inventory of groundwater datasets

Category	Scale	Source	Dataset	Date Range	Description	Model Use	Link
Groundwater Basin Boundaries	State	DWR	DWR's Bulletin 118	2020	Groundwater basin boundaries represent alluvial basins delineated by DWR.	Groundwater domain	LINK
Groundwater levels	State	DWR	Periodic Groundwater Level Measurements	2023	Groundwater levels	Model calibration	LINK
Geologic information	State	DWR	Well Completion Reports (OSWCR)	2023	Geologic information	Groundwater stratigraphy and properties	LINK
Groundwater Extraction	State	DWR	Groundwater Sustainability Plan (GSP) Annual Report Data	2023	Groundwater extraction	Groundwater extraction	LINK

2 METEOROLOGY

Precipitation and ET are key components of the water balance and critical inputs for developing a hydrologic model. The following subsections describe the primary data sources for precipitation and ET.

2.1 Precipitation

The primary source of precipitation data for the Petaluma River watershed will be the observed data from land-based stations within and in the vicinity of the watershed. However, any gaps in observed data from the land-based stations will be filled with grid-based data. This is referred to as the "hybrid" approach, which leverages both land-based and grid-based data to build a complete spatial and temporal meteorological dataset. Use of a hybrid approach preserves locally sampled station data while increasing the spatial and temporal quantity and quality of data over the watershed. This approach has been applied for large watershed-scale modeling applications including the County-wide model for Los Angeles County (LACFCD 2020).

Land-based observed precipitation data are primarily acquired from the National Centers for Environmental Information (NCEI), which maintains data from the National Oceanic and Atmospheric Administration's (NOAA) climate networks including the Global Historic Climate Network (GHCN), the Cooperative Observer Program (COOP), and the Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS). These networks provide quality-controlled hourly or daily observed precipitation data. Fourteen GHCN gauges were identified within or near the Petaluma River watershed, with each having a varied quantity and quality of data. In addition to the daily precipitation gauges, NCEI also maintains the Local Climatological Data (LCD) network, which reports hourly observations. There are three LCD stations within 15 km of the Petaluma River watershed. The California Data Exchange Center (CDEC) and the Remote Automated Weather Stations (RAWS) networks also report hourly precipitation. CDEC reports at one location and RAWS reports at six locations within and near the watershed.

In addition to rainfall data, snow data was also screened for the Petaluma River watershed model, which yielded eight GHCN stations within a 15 km buffer (reference Figure 2 1) that had data (e.g., snowfall measurements) overlapping with the model time period. These stations can be further evaluated if needed; however, snow simulation is not expected to be required because of the relatively low elevations in this watershed.

Table 2-1 is an inventory of the precipitation stations near the Petaluma River watershed with available data after 2000 with at least 75% data coverage during the modeling period. It also includes stations with 60% or more data coverage during the modeling period, provided that these stations have 90% or more station record coverage. Station record coverage is defined as the percentage of available records that are not missing throughout the duration of the record for the period overlapping with the modeling period. Additionally, all recently installed stations with at least 6 years of data (i.e., 30% or more model period availability) for the modeling period are included if they have 90% or more station record coverage for their reported period overlapping with the modeling period. Lastly, any stations with no neighboring stations within a 7 km buffer are included. However, these stations may not have sufficient data quality for use in model development.

Figure 2-1 shows the locations of the identified stations. There is one location where LCD and GHCN stations appear to be co-located (i.e., within a 500 m radius); these will be further assessed under Task 3.2 and the stations with the highest quality data will be chosen for use.

The primary source of the grid-based data for Petaluma River Watershed will be the Parameter-elevation Regressions on Independent Slopes Model (PRISM) (Daly et al. 2008; Daly, Neilson, and Phillips 1994; Gibson et al. 2002). PRISM is developed and maintained by the PRISM Climate Group at Oregon State University and provides gridded estimates of event-based climate parameters including precipitation, temperature, and dew point. The algorithm uses observed point data, a digital elevation model (DEM), and other spatial datasets to capture influences such as high mountains, rain shadows, temperature inversions, coastal regions, and other complex climatic regimes (Gibson et al. 2002). Because of its spatial and temporal resolution and consistency across the contiguous United States (4-km spatial resolution for the AN81d daily/monthly time series dataset and 800-m for the AN81m long term averages), PRISM is a commonly used and widely accepted source for meteorological data for hydrologic models (Behnke et al. 2016). The subset of the PRISM grid that covers the current study area is shown in

Figure 2-1. To disaggregate the PRISM data to hourly, North American Land Data Assimilation System (NLDAS) will be used. NLDAS is a quality-controlled land surface model (LSM) dataset of meteorological data designed specifically to support continuous simulation modeling activities (Cosgrove et al. 2003; Mitchell et al. 2004). NLDAS provides real-time hourly predictions of meteorological data required for LSPC at a 1/8th degree spatial resolution (about 8.625-mile intervals) for North America, with retrospective simulations beginning in January 1979. NLDAS has undergone rounds of refinement, extensive peer review, and performance validation through case study applications, all of which have demonstrated it to be a more robust predictor of variable meteorological conditions for continuous simulation modeling than using individual gauges (Xia et al. 2012).

Table 2-1. Summary of precipitation stations with observations available after 2000

Agency	Station ID[1, 3]	Name	Start Date	End Date	Lat.	Long.	Elevation (meters)	Model Period Coverage (%)[1]	Station Record Coverage (%)[2]
NOAA-LCD	WBAN:93227	NAPA CO AIRPORT, CA US	12/31/1999	Present	38.2075	-122.2804	4.7	100%	100%
	WBAN:00320	PETALUMA MUNICIPAL AIRPORT, CA US	7/30/2014	Present	38.2500	-122.6000	27.1	46%	100%
	WBAN:00135	GNOSS FIELD AIRPORT, CA US	7/30/2014	Present	38.1500	-122.5500	1.2	46%	100%
NOAA-GHCN	GHCND:USC00047880	SAN RAFAEL-CIVIC CTR	1/1/1894	Present	37.9983	-122.5372	36.6	80%	80%
	GHCND:USC00046027	MUIR WOODS	12/1/1940	Present	37.8978	-122.5689	67.1	83%	83%
	GHCND:USC00047414	RICHMOND	12/1/1950	Present	37.9192	-122.3772	6.1	84%	84%
	GHCND:USC00048351	SONOMA	1/1/1893	Present	38.2994	-122.4622	29.6	92%	92%
	GHCND:USC00044500	KENTFIELD	1/1/1902	Present	37.9567	-122.5447	44.2	100%	100%
	GHCND:USC00046826	PETALUMA AP	2/1/1893	Present	38.2578	-122.6078	6.1	100%	100%
	GHCND:USW00093227	NAPA CO AP	5/1/1998	Present	38.2075	-122.2803	4.9	100%	100%
	GHCND:US1CASN0049	SEBASTOPOL 1.1 SSE	2/1/2009	Present	38.3842	-122.8191	51.2	67%	92%
	GHCND:US1CASN0071	ROHNERT PARK 0.9 SW	2/1/2009	Present	38.3392	-122.7113	29.3	70%	96%
	GHCND:US1CAMR0009	PETALUMA 10.1 W	3/1/2009	Present	38.2475	-122.8119	36.3	72%	99%
	GHCND:US1CASN0080	GLEN ELLEN 1.5 N	4/1/2009	2/28/2025	38.3778	-122.5343	102.7	72%	100%
	GHCND:US1CAMR0026	NOVATO 1.2 SW	10/1/2014	Present	38.0790	-122.5698	69.5	41%	93%
	GHCND:US1CAMR0025	SAN ANSELMO 2.0 NNW	9/1/2013	Present	38.0096	-122.5801	62.2	50%	99%

Agency	Station ID[1, 3]	Name	Start Date	End Date	Lat.	Long.	Elevation (meters)	Model Period Coverage (%)[1]	Station Record Coverage (%)[2]
	GHCND:US1CA NP0005	NAPA 2.0 WNW	3/1/2009	5/31/2011	38.3128	-122.3312	41.5	9%	84%
CDEC	RHL	RICHMOND CITY HALL	10/1/2015	Present	37.9369	-122.3427	16.8	40%	96%
RAWS	MDEC1	MIDDLE PEAK	5/17/2004	Present	37.9278	-122.5872	712.9	78%	81%
	WDAC1	WOODACRE	5/23/2003	Present	37.9906	-122.6447	426.7	82%	82%
	NBRC1	BIG ROCK	9/29/2003	12/27/2024	38.0394	-122.5700	457.2	97%	97%
	BBEC1	BARNABY	2/13/1997	Present	38.0281	-122.7022	246.9	99%	99%
	NVHC1	NOVATO FIRE - ROBINHOOD	1/21/2016	Present	38.1125	-122.5498	146.9	38%	100%
	OVYC1	OLEMA VALLEY	5/23/2005	Present	38.0425	-122.7958	11.3	73%	81%

1. Data coverage for LCD stations are LCD portal-reported values which are reflective of data availability between the reported start date and end date in this table, not proportionally scaled to the modeling period (i.e. 10/2003 to 9/2023) when LCD's coverage period does not fully overlap with the modeling period. Data coverage for CDEC, GHCN, and RAWS are based on data availability for the modeling period (10/2003 to 9/2023). Data completeness will be further assessed under Task 3.2 and additional stations may be considered as needed.
2. Station record coverage is defined as the percentage of available records that are not missing from the Start Date to the End Date overlapping with the modeling period.
3. Stations marked with an asterisk were included because there are no neighboring stations within a 7 km buffer; data quality at these stations may be insufficient for use in model development and will be further assessed under Task 3.2.

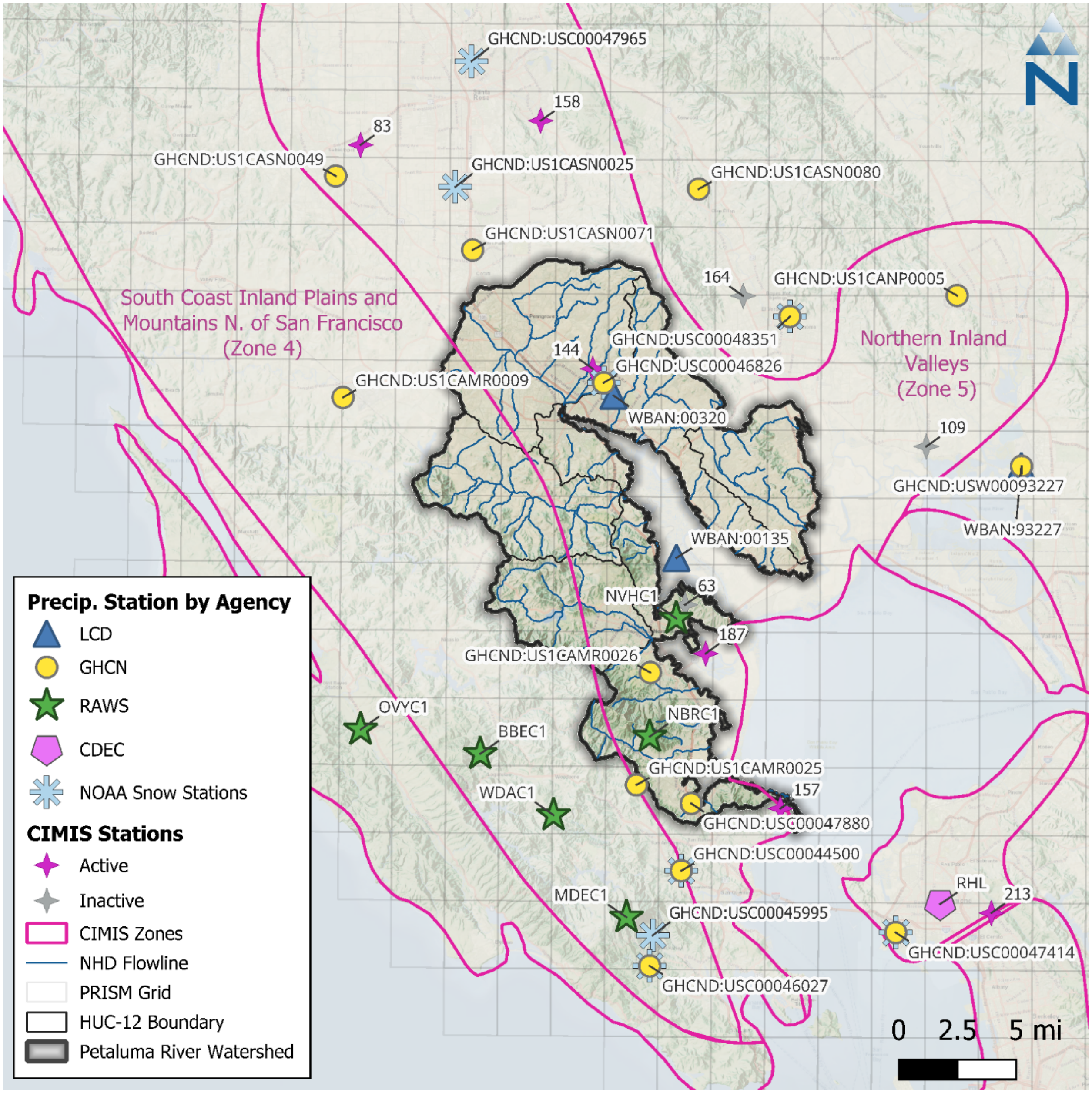

Figure 2-1. Identified rainfall stations and CIMIS ET Zones near the Petaluma River watershed.

The hybrid approach mentioned above entails three main steps. First, impaired intervals (i.e., missing, or accumulated) at observed stations will be patched with grid-based data. Second, the PRISM grid cells, and patched observed stations will be mapped to the NLDAS grid cells to disaggregate the monthly PRISM and daily station data using normalized hourly data from NLDAS. Third, the disaggregated gridded meteorological data from PRISM will be used to fill any remaining spatial gaps in the observed station network as needed. It should be noted that while PRISM gridded data also provides estimates of precipitation on a daily time step, using monthly PRISM totals for disaggregating with hourly observed data, as opposed to daily PRISM totals, eliminates the need to estimate distributions for instances where an hourly distribution does not coincide with a daily total.

Figure 2-2 presents a summary of the hybrid approach to blend observed precipitation with gridded meteorological products. Observed data and gridded products are to be processed in parallel to: (1)

create a temporally complete set of hourly distributions and (2) identify spatial gaps in coverage to be supplemented with disaggregated gridded data. Assuming a 10 km buffer around observed stations for this approach, the coverage shown in the lower right map in Figure 2-2 also shows what a hybrid dataset of observed time series, supplemented by gridded products, would look like.

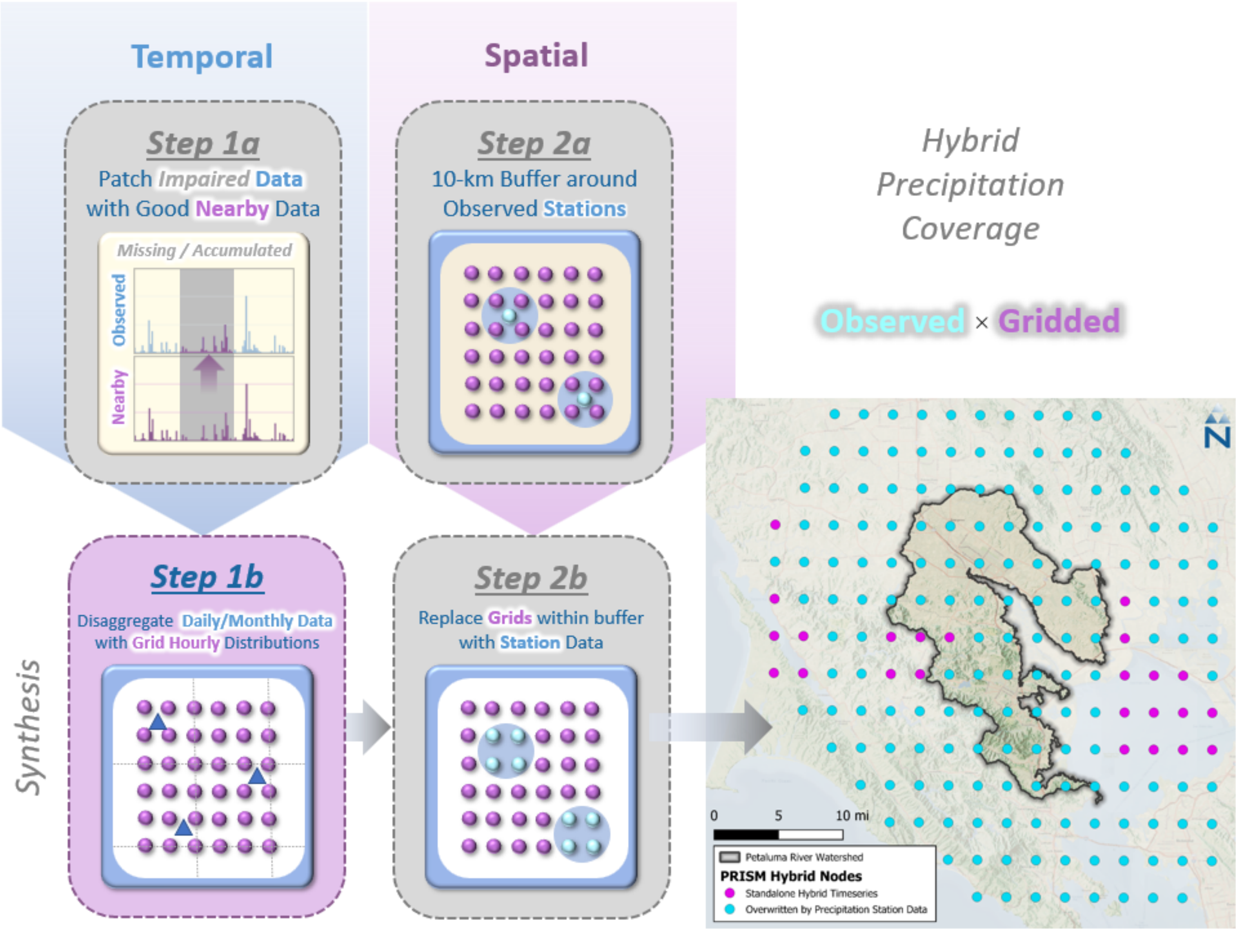

Figure 2-2. Hybrid approach to blend observed precipitation with gridded meteorological products.

2.2 Evapotranspiration

The primary ET dataset identified for consideration is California Irrigation Management Information System (CIMIS). CIMIS was developed in 1982 by the California Department of Water Resources (DWR) and the University of California, Davis. The network is composed of over 145 automated weather stations throughout California where primary weather data including temperature, relative humidity, wind speed, and solar radiation are monitored and quality controlled. Observations are measured over standardized reference surfaces (e.g., well-watered grass or alfalfa) and are used to estimate reference evapotranspiration (ET_o) using versions of the Penman and Penman-Monteith equations. CIMIS has divided California into 18 zones based on long-term monthly average ET_o values calculated using data from CIMIS weather stations.

CIMIS operates nine stations within ten miles of the Petaluma River watershed, including: Santa Rosa (ID 83), Petaluma (ID 144), Point San Pedro (ID 157), Bennett Valley (ID 158), Black Point (ID 187), El Cerrito (ID 213), Carneros (ID 109), Novato (ID 63) and Valley of the Moon (ID 164). The Valley of the Moon, Carneros, and Novato stations are no longer operating, but their collective historical time series data covers the period from July 1986 through January 2022. Of the active stations within the 10-mile watershed buffer, the Petaluma station is within the northern watershed region and contains data from August 1999 through the present, and the Santa Rosa station, situated northwest of the watershed region, contains data from January 1990 through the present. As shown in Figure 2 1, the Petaluma River watershed intersects two CIMIS zones, with nearly 79% of the watershed area in Zone 5 (Northern Inland Valleys), and 21% of the watershed area in Zone 4 (South Coast Inland Plains and Mountains North of San Francisco). These zones experience average annual reference evapotranspiration levels from 43.9 inches per year in Zone 5 to 46.6 inches per year in Zone 4.

CIMIS also has a newly derived gridded product, CIMIS Spatial, that expresses daily ET_o estimates calculated at a statewide 2-km spatial resolution using the American Society of Civil Engineers version of the Penman-Monteith equation (ASCE-PM) (Allen et al. 2005). The ASCE-PM method calculates ET_o using solar radiation, air temperature, relative humidity, and wind speed at two meters height. This product provides a consistent spatial estimate of ET_o that is California-specific, implicitly captures macro-scale spatial variability and orographic influences, is available from 2003 through Present, and is routinely updated within a couple of days.

Representative potential evapotranspiration (PEVT) time series can be estimated for the Petaluma River watershed from daily data from CIMIS Spatial by disaggregating the hourly time series using hourly distributions from land observation stations (e.g., RAWS, NCEI) or hourly distributions from NLDAS. PEVT is reported at 3-hour intervals; however, the hourly distributions of solar radiation from NLDAS, which have sinusoidal patterns over daylight hours, provide a sound basis for disaggregating the daily CIMIS depths while maintaining the overall annual water budget reflected in CIMIS.

For LSPC, the user provides PEVT rates as model input. The LSPC model then uses these values along with other model parameters to estimate actual ET. Sometimes ET_o is provided instead, and HRU-specific coefficient multipliers are used to stratify those inputs based on physical HRU properties such as vegetation density. Additionally, for applications where the study area has significant irrigation practices, the user can provide irrigation water usage rates to represent additional water beyond precipitation that is added to the system—that water would also be available for ET.

The actual ET estimated by an LSPC model can be validated through comparison with data from OpenET. The OpenET project is an operational system for generating and distributing ET data at a field scale using an ensemble of six well-established satellite-based approaches for mapping ET (Melton et al. 2022). OpenET has undergone extensive intercomparison and accuracy assessment conducted using ground measurements of ET; results of these assessments demonstrate strong agreement between the satellite-driven ET models and observed flux tower ET data. Within California, OpenET has data beginning in 2016 and uses CIMIS meteorological datasets to compute ET_o. In addition to LSPC ET validation, OpenET data can be used to help inform irrigation estimation and parameterization.

In the SFEI LSPC model containing the Petaluma River watershed, PEVT is represented using NLDAS and no irrigation is simulated. To assess the performance of this model, simulated ET was compared to the model input PEVT timeseries, CIMIS Zone 8 observed ET_0, and gridded ET estimates from OpenET for the years 2016-2020. Figure 2-3 summarizes the distribution of spatial and temporal variation of monthly ET for these datasets for the adjacent Napa River watershed that is expected to have similar ET signatures (except for the CIMIS dataset, which shows the spatial-temporal average of Zone 8). Findings from the Napa River watershed shown in Figure 2-3 are expected to be applicable for the Petaluma River watershed due to its proximity. However, a similar analysis will be performed for the Petaluma River watershed to confirm.

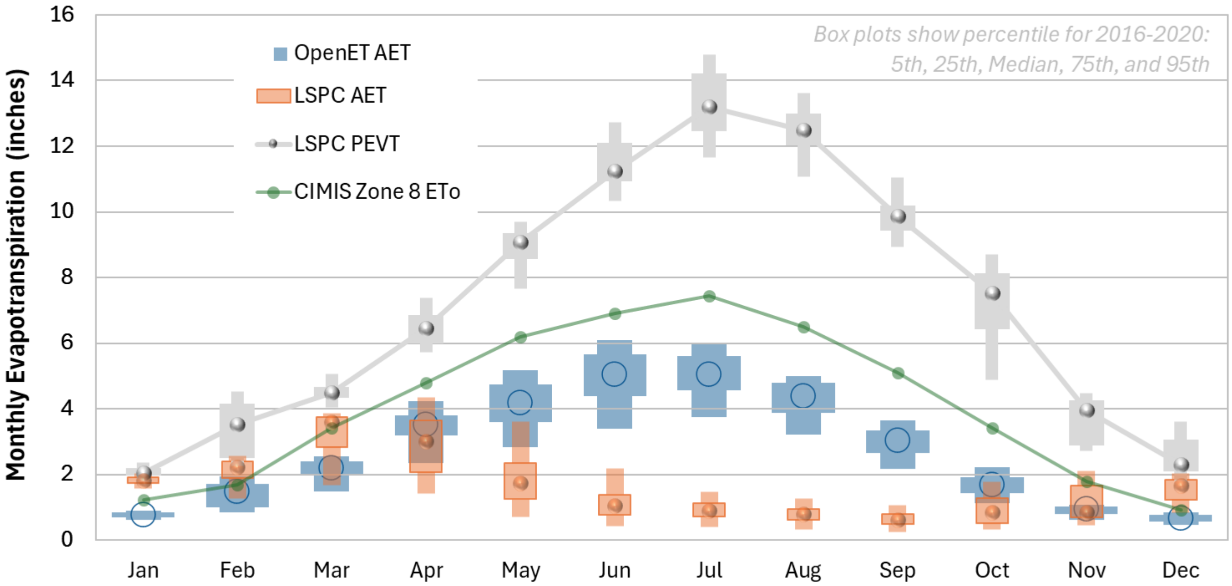

Figure 2-3. Evaluation of monthly evapotranspiration in existing SFEI model. Note that AET is used as an abbreviation for actual ET.

This analysis yielded three key findings for the SFEI model:

1. Modeled input for PEVT based on NLDAS is significantly greater than observed ET_0 from CIMIS.
2. During non-growing season months, modeled ET has a positive bias relative to ET from OpenET. This indicates overestimation of the system energy balance.
3. During growing season months, modeled ET is significantly lower than observed ET from OpenET. While the extent of agricultural practice in the Petaluma River watershed is relatively small, without including irrigation in the model, along with any necessary adjustments to subsurface storage, there is not enough water available to make up for this deficit, which could then result in misleading inferences about the overall hydrologic budget.

Enhancing PEVT accuracy and integrating Petaluma River watershed irrigation rates are critical steps to improve ET predictions and representation of the overall hydrologic budget. The existing PEVT dataset used in the SFEI model (from NLDAS) results in overestimated rates of simulated ET compared to in situ observations (Xia et al. 2015). To remove this positive bias and smooth diurnal cycles of ET, the hourly sinusoidal distribution of NLDAS short-wave radiation would be used to disaggregate daily ET_o from CIMIS. The disaggregated ET_o CIMIS data could then be used as PEVT input data to the model. Additionally, irrigation water usage should be represented with observed rates where data are available. The observed irrigation rates should be evaluated against estimated monthly irrigation derived by calculating differences in observed ET (OpenET) and model-simulated actual evapotranspiration (AET) with irrigation turned off.

3 SURFACE HYDROLOGY

3.1 Watershed Segmentation

The United States Geological Survey (USGS) delineates watersheds nationwide based on surface hydrological features and organizes the drainage units into a nested hierarchy using HUCs. These

HUCs have a varying number of digits to denote scale ranging from 2-digit HUCs (largest) at the region scale to 12-digit HUCs (smallest) at the subwatershed scale. The Petaluma River watershed is defined by a HUC-10 watershed that comprises of seven HUC-12 subwatersheds.

For units smaller than HUC-12, catchment and tributary boundaries, flowlines, outlet points, and related attribute information will rely on the National Hydrography Dataset (NHD), HUC, and catchment delineations. This analysis will primarily use readily available data to define the outer watershed boundary. Any available local data will be used to supplement and refine the understanding of tributary boundaries and reach geometry. The NHD Plus v2 (NHDPlus) further discretizes the watershed into 147 catchments ranging in size between 1.8 acres to nearly 5,300 acres. Table 3-1. presents summary statistics of NHDPlus catchment sizes by HUC-12 subwatershed. Figure 3-1 is a map of HUC-12 and NHDPlus catchments within the Petaluma River watershed (HUC-10).

The NHDPlus dataset provides a good foundation for model segmentation at a spatial scale that is suitable for representing the watershed for the purposes of modeling daily, seasonal, and annual streamflow. The NHDPlus catchment boundaries will be aggregated and/or adjusted only as necessary to align with any selected points of interest (e.g., flow monitoring sites) to allow for direct output of model results for comparison and analysis.

Table 3-1. Summary of NHDPlus catchment sizes (acres) within the Petaluma River HUC-10

HUC-12 Name	Count	Catchment Size (acres)			
		Minimum	Mean	Median	Maximum
Tolay Creek	16	1.8	745.8	620.9	2,520.4
Burdell Mountain	17	9.9	891.3	327.5	5,270.2
Petaluma River	21	14.5	1,337.9	1,110.1	4,463.5
Miller Creek	25	10.0	711.2	527.3	3,008.7
Adobe Creek	25	39.6	750.2	421.0	2,322.5
San Antonio Creek	30	5.1	660.6	638.1	1,833.6
Novato Creek	13	20.2	1,152.2	584.5	3,292.4

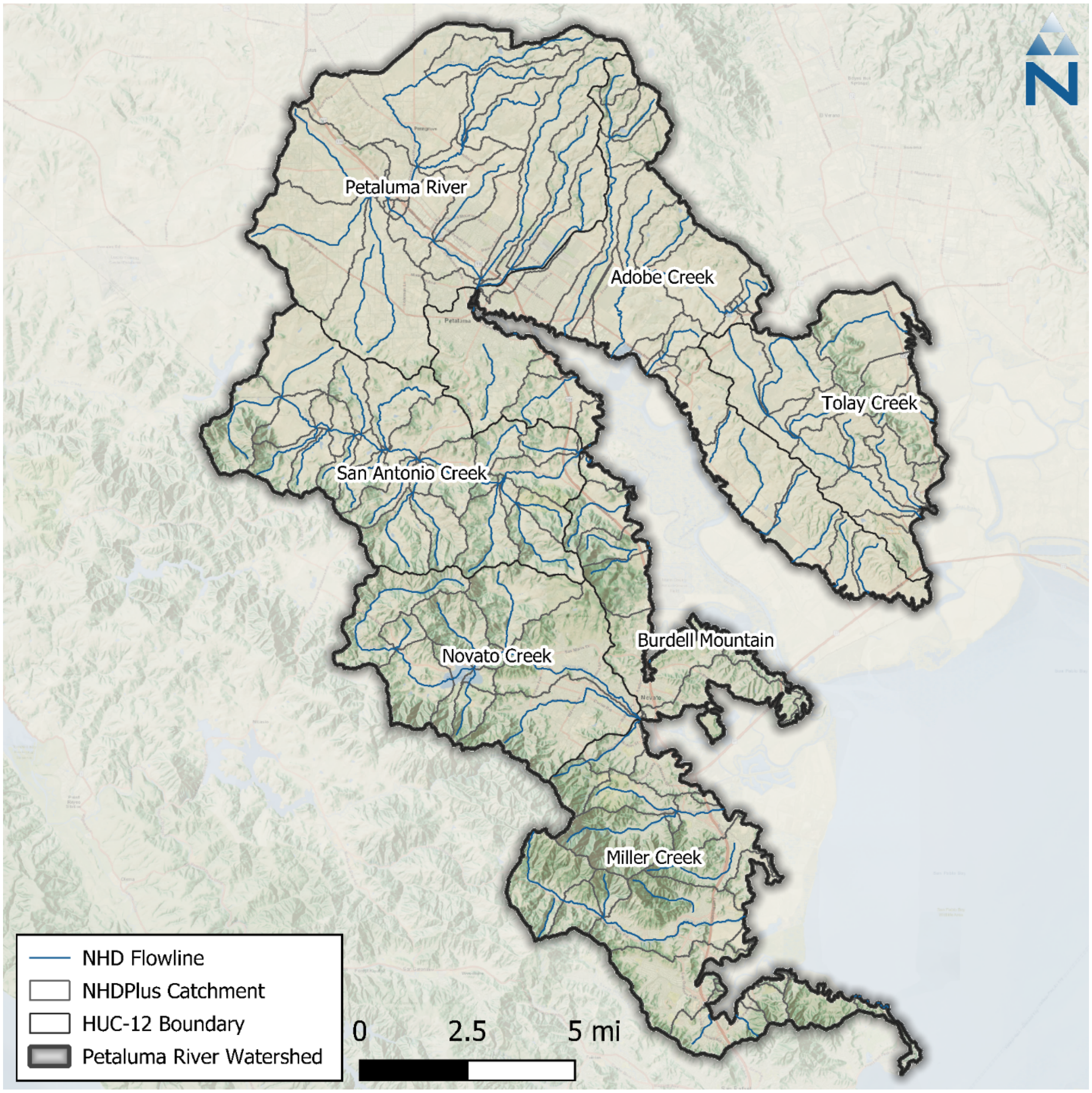

Figure 3-1. Initial catchment segmentation for the Petaluma River watershed.

3.2 Streams and Channels

As described above, the hydrographic characteristics of the streams and rivers within the Petaluma River watershed are primarily derived from NHDPlus. This dataset depicts primary flow paths based on a nation-wide 10m DEM and includes additional attributes such as hydrologic sequence and flow line slope. These characteristics will be important for creating representative reach segments within the hydrologic model. Reference Figure 3-1 for the location of the watershed's major tributaries ("NHD Flowline").

3.3 Streamflow

The primary source of streamflow data is from the USGS, which includes one current long-term station: the Novato Creek at Novato, CA (USGS 11459500). There are also two historical streamflow stations located upstream on tributaries, but these do not have data after 1983 and therefore, are not expected to be used during model calibration and validation. Table 3-2 presents a summary of the available USGS streamflow data after 2000. Figure 3-2 shows the location of the USGS flow station within the Petaluma River watershed.

Table 3-2. Summary of USGS daily streamflow data after 2000.

Station Description	Station ID	Drainage Area (mi^2)	Start Date	End Date	Station Active?
NOVATO C A NOVATO CA[1]	11459500	17.6	10/1/1946	Present	Yes

1. USGS notes "Flow regulated by Stafford Lake beginning Dec. 1, 1951, capacity, 4,500 acre-ft, since Oct. 18, 1954. Diversion from Stafford Lake for municipal water supply began Apr. 25, 1952." (https://waterdata.usgs.gov/nwis/wys_rpt/?site_no=11459500&agency_cd=USGS#adr).

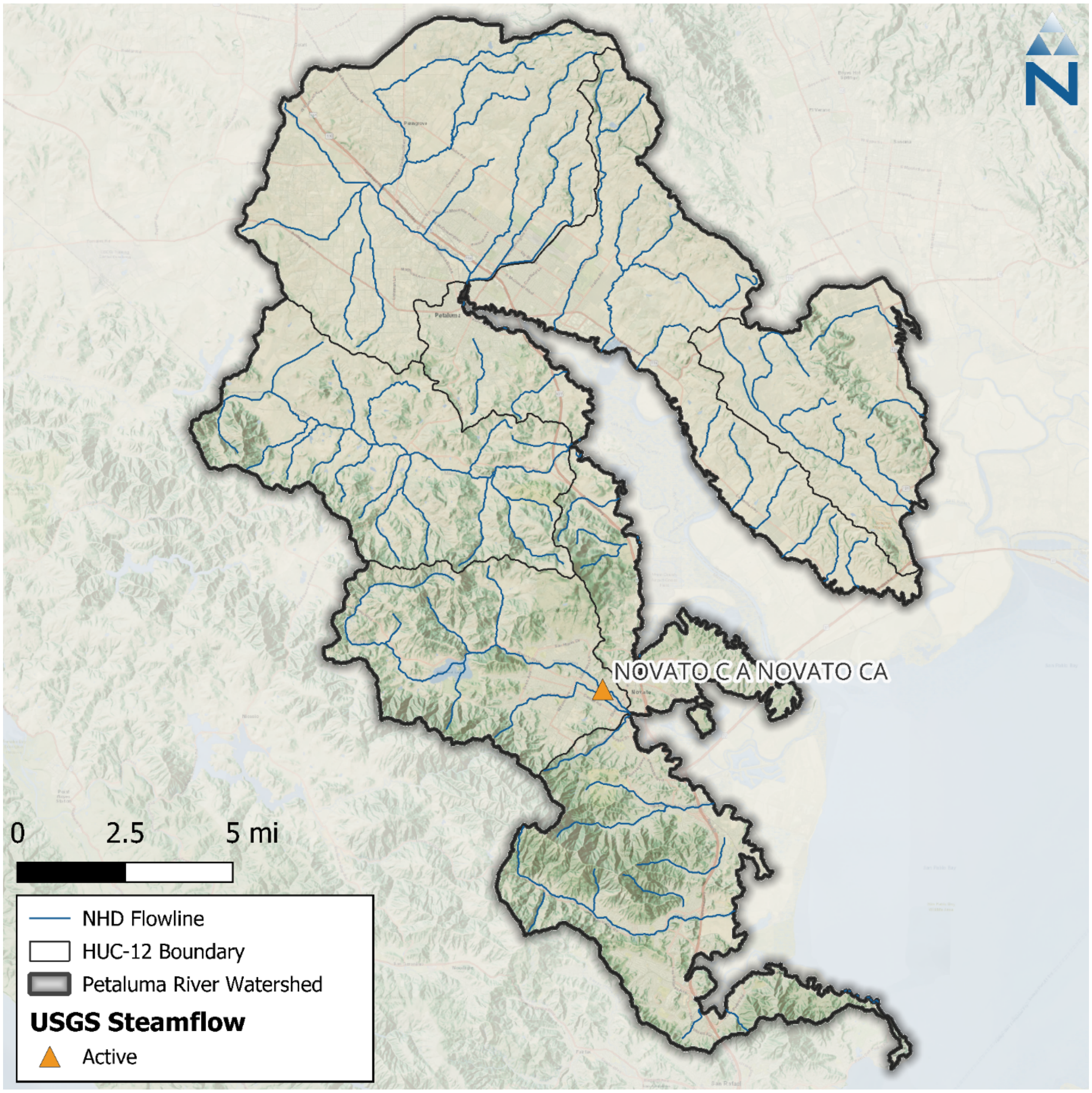

Figure 3-2. USGS streamflow stations in the Petaluma River watershed.

3.4 Dams, Reservoirs, and Impoundments

The Petaluma River watershed contains several small reservoirs that may require representation within the hydrology model, as shown in Figure 3-3. Information for the six dams within the watershed is from the National Inventory of Dams (NID) and these are also listed in Table 3-3. Lawler is owned by the City of Petaluma, and Novato Creek is a public utility owned by North Marin County Water District. Sonoma Hills, Pinheiro, Sleepy Hollow No. 2, and Vonsen are privately owned. All dams are used for water supply, irrigation, or recreation. Key metrics such as the average flow volume and average storage of each dam are currently unknown, as no readily accessible data was found and the NID database does not report this information. Capturing the operation of these features will be important to accurately represent the movement of water throughout the watershed; however, most are relatively small and expected to have a negligible impact. Having stage-storage relationships for reservoirs, and any other outflow rates or operating conditions, will be needed for any waterbodies in the model.

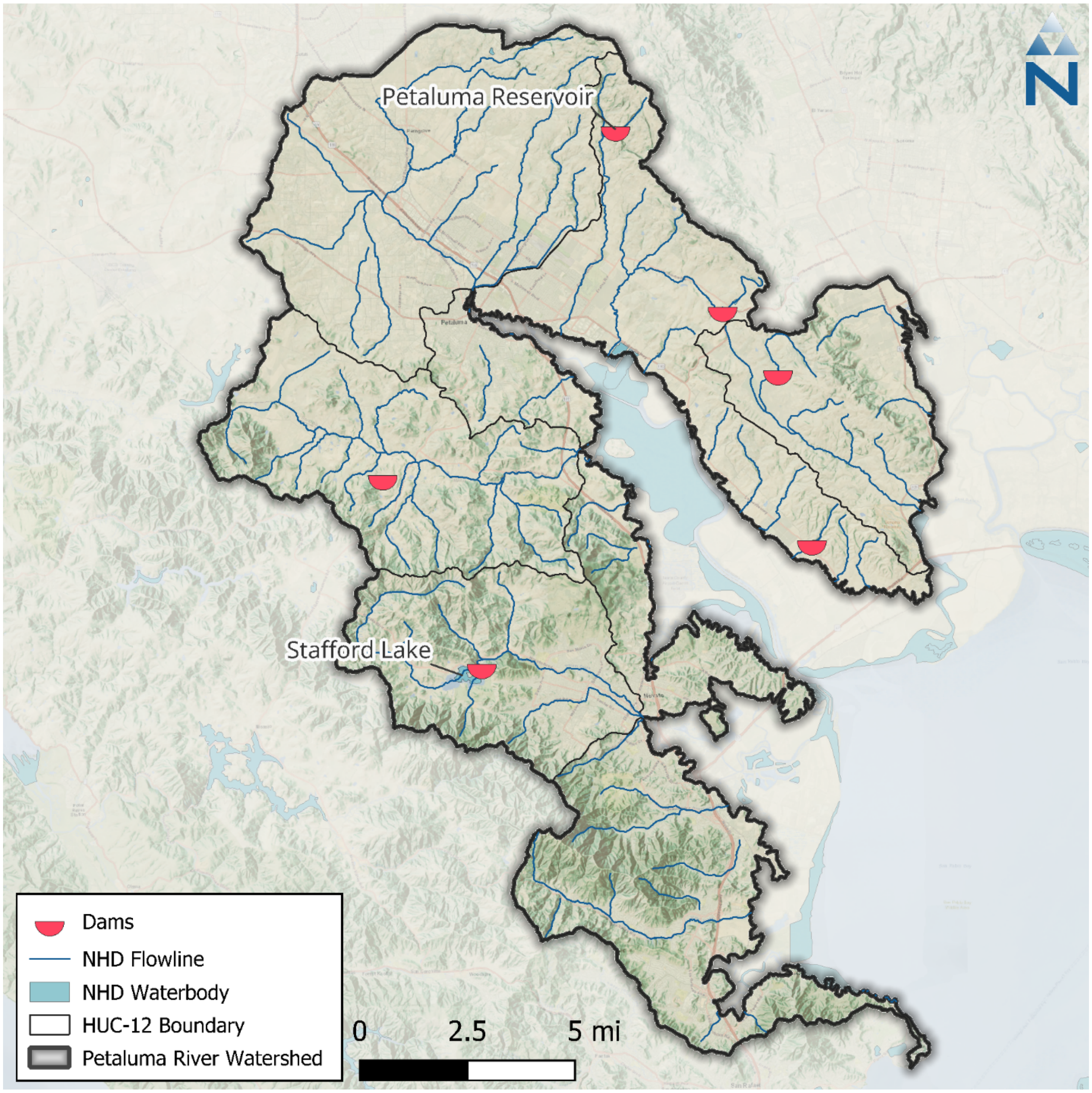

Figure 3-3. Dams and waterbodies in the Petaluma River watershed (NID 2025).

Table 3-3. Waterbodies and Dams within the Petaluma River watershed

NHD Waterbody	Dam Name	NID ID	Drainage Area (ac)	Area (ac) [1]	Storage Capacity (ac-ft)	Average Flow Volume (ac-ft)	Average Storage (ac-ft)
Petaluma Reservoir	Lawler	CA00834	320	10	227	Unknown[3]	Unknown[3]
Stafford Lake	Novato Creek	CA00321	5,530	195	4,430	Unknown[3]	Unknown[3]

NHD Waterbody	Dam Name	NID ID	Drainage Area (ac)	Area (ac) [1]	Storage Capacity (ac-ft)	Average Flow Volume (ac-ft)	Average Storage (ac-ft)
Unknown	Sonoma Hills	CA01321	832	29	240	Unknown[3]	Unknown[3]
Unknown	Pinheiro	CA01062	-[2]	11	83	Unknown[3]	Unknown[3]
Unknown	Sleepy Hollow No. 2	CA00590	449	5	104	Unknown[3]	Unknown[3]
Unknown	Vonsen	CA00592	128	9	70	Unknown[3]	Unknown[3]

1. The primary source of data for area is NID. If NID is not available because of the absence of dams associated with the waterbody, then NHD data is used.
2. Data not listed from NID or NHD.
3. Data may exist but were not readily accessible.

3.5 Surface Water Withdrawals

Datasets related to water rights, points of diversion (PODs), and surface withdrawals (i.e., wells and irrigation) were identified through searches of the Water Board’s Electronic Water Rights Information Management System database (eWRIMS) and the DWR Agricultural Land and Water Use Estimates database (ALWU). These datasets can be used to represent diversions, withdrawals, and irrigation practices in the watershed model. The volumes quantified in those datasets can be compared to annual and seasonal water budget estimates in the Petaluma River watershed to assess the relative impacts based on observed precipitation, ET, and streamflow data. The impact of diversions or water usage may be localized along specific tributaries; however, the temporal resolution of the data determines the resolution of those impacts in the model. Additionally, the extent of modeled irrigation will depend on land-use classification, and its water usage rates will be corrected against spatial variations in the observed evaporative deficit where necessary.

Figure 3-4 provides an overview of primary sources of drinking water. Water systems distributed throughout the watershed include a mixture of both surface water diversions from the Petaluma River and its primary tributaries, as well as groundwater withdrawals for the Petaluma River watershed. For 35 out of the 38 drinking water systems, the water source is listed as groundwater, and 3 have surface water listed as the source. The number of active surface water PODs and withdrawal volumes will be evaluated from eWRIMS reports during model development.

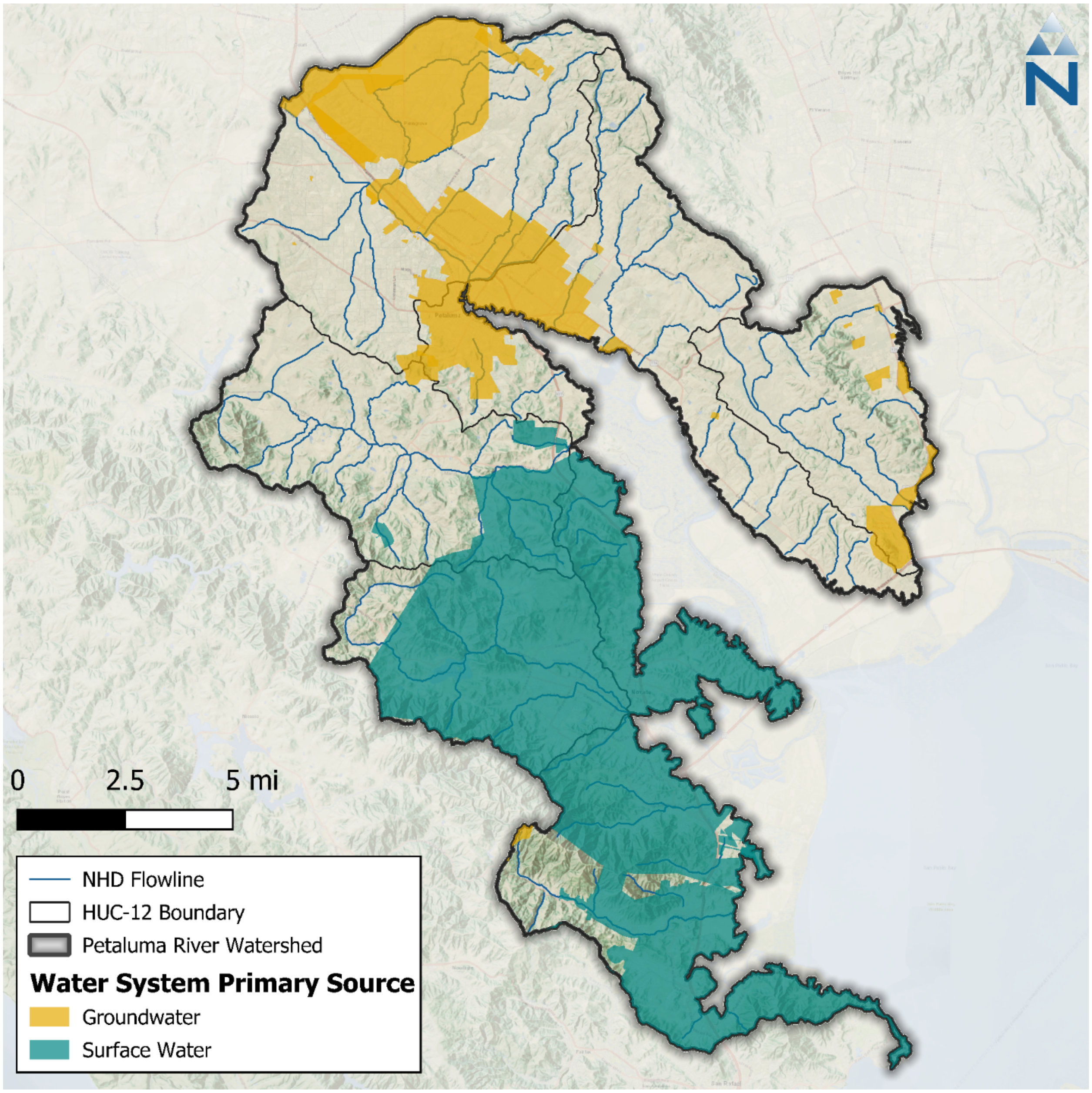

Figure 3-4. Water system primary sources in the Petaluma River watershed.

4 SUBSURFACE HYDROLOGY

The Petaluma River watershed overlaps with four groundwater basins as delineated by Bulletin 118 (DWR 2020). The watershed contains most of the Petaluma Valley basin (# 2-001) and also overlaps with the Novato Valley basin (# 2-030), the Wilson Grove Formation Highlands (# 1-059), and the Sonoma Valley sub-basin (# 2-002.02), with a sliver containing negligible overlap at the border of the Santa Rosa Plain sub-basin (# 1-055.01). The watershed does not fully contain any basins. Approximately 39% of the watershed area falls within the groundwater basins delineated by Bulletin 118 and the remaining 41% consists of andesitic and basaltic volcanic rock to the east and marine mudstones to the west.

As per the respective basin priority details (Sustainable Groundwater Management Act (SGMA) Basin Prioritization Dashboard), the Petaluma Valley basin is Medium priority due to declining groundwater levels and a large number of total wells. Approximately 70% of the Petaluma Valley basin falls within the Petaluma River watershed. The Novato Valley basin is Low priority; approximately 40% of the Novato Valley basin falls within the Petaluma River watershed. The Wilson Grove Formation Highlands basin is Very Low priority. Approximately 10% of the Wilson Grove Formation Highlands basin falls within the Petaluma River watershed. The Sonoma Valley sub-basin is High priority due to declining water levels, saline intrusion, and a large number of water supply and total wells. Approximately 5% of the Sonoma Valley sub-basin falls within the Petaluma River watershed. The Sonoma Valley Groundwater Sustainability Agency (GSA) operates within the Sonoma Valley sub-basin, and the Petaluma Valley GSA operates within the Petaluma Valley basin (DWR 2025a and 2025b).

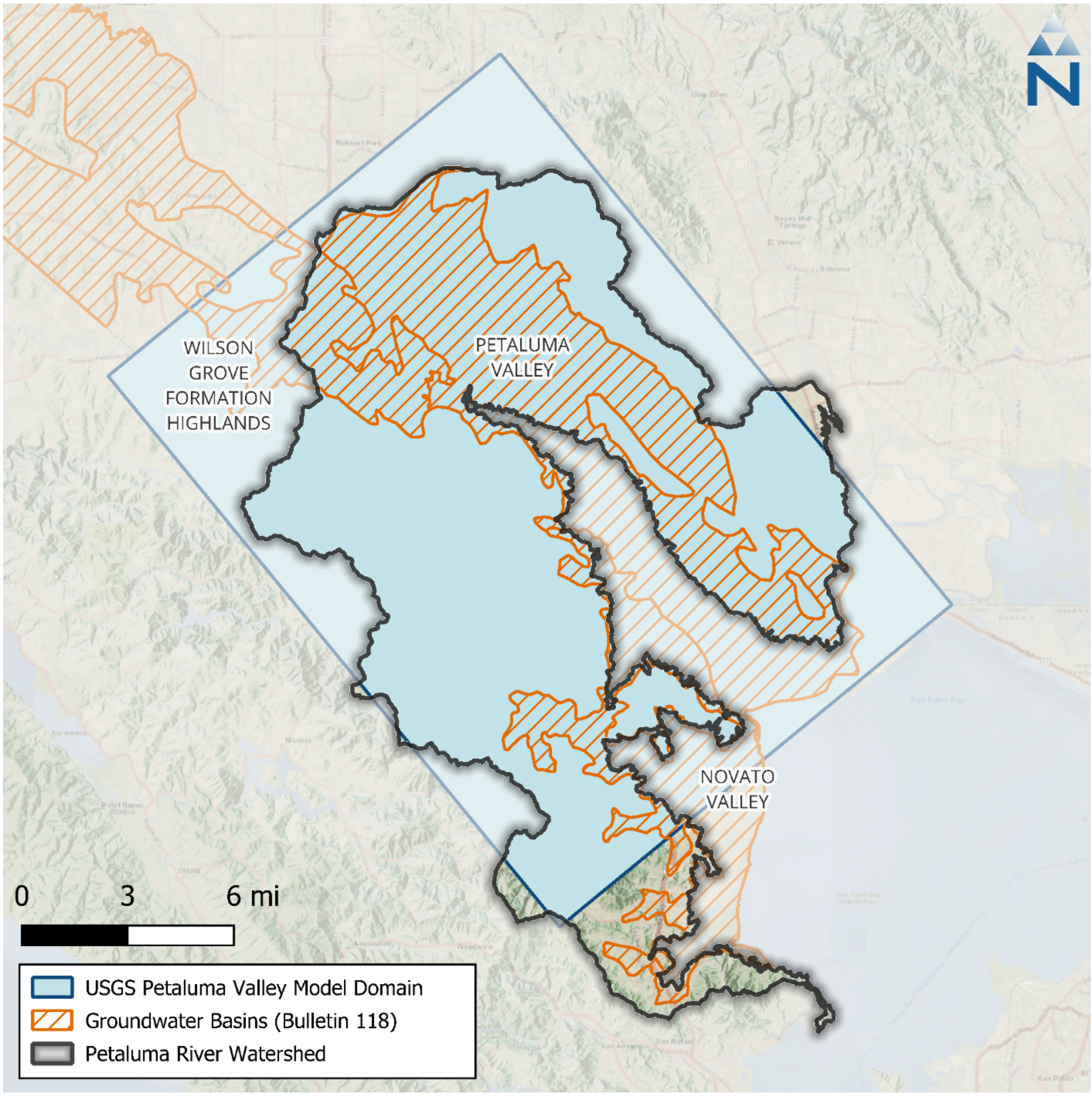

Figure 4-1. Outlines of the Petaluma River Watershed Study Area, groundwater basin boundaries (Bulletin 118), and USGS Petaluma Valley Model Boundary.

4.1 Water Budget Components

The publicly available USGS Petaluma River model overlaps the watershed (USGS 2025). The average annual groundwater budget presented in the USGS model report (Traum et al. 2022) indicated that stream seepage to groundwater makes up 83% of groundwater inflows, with the remaining inflow coming from upland and San Pablo Bay boundaries. Agricultural groundwater pumping in the simulation makes up 4% of groundwater outflows, with urban and rural pumping accounting for less than 1%. The remaining modeled outflow is predominantly groundwater discharge to streams.

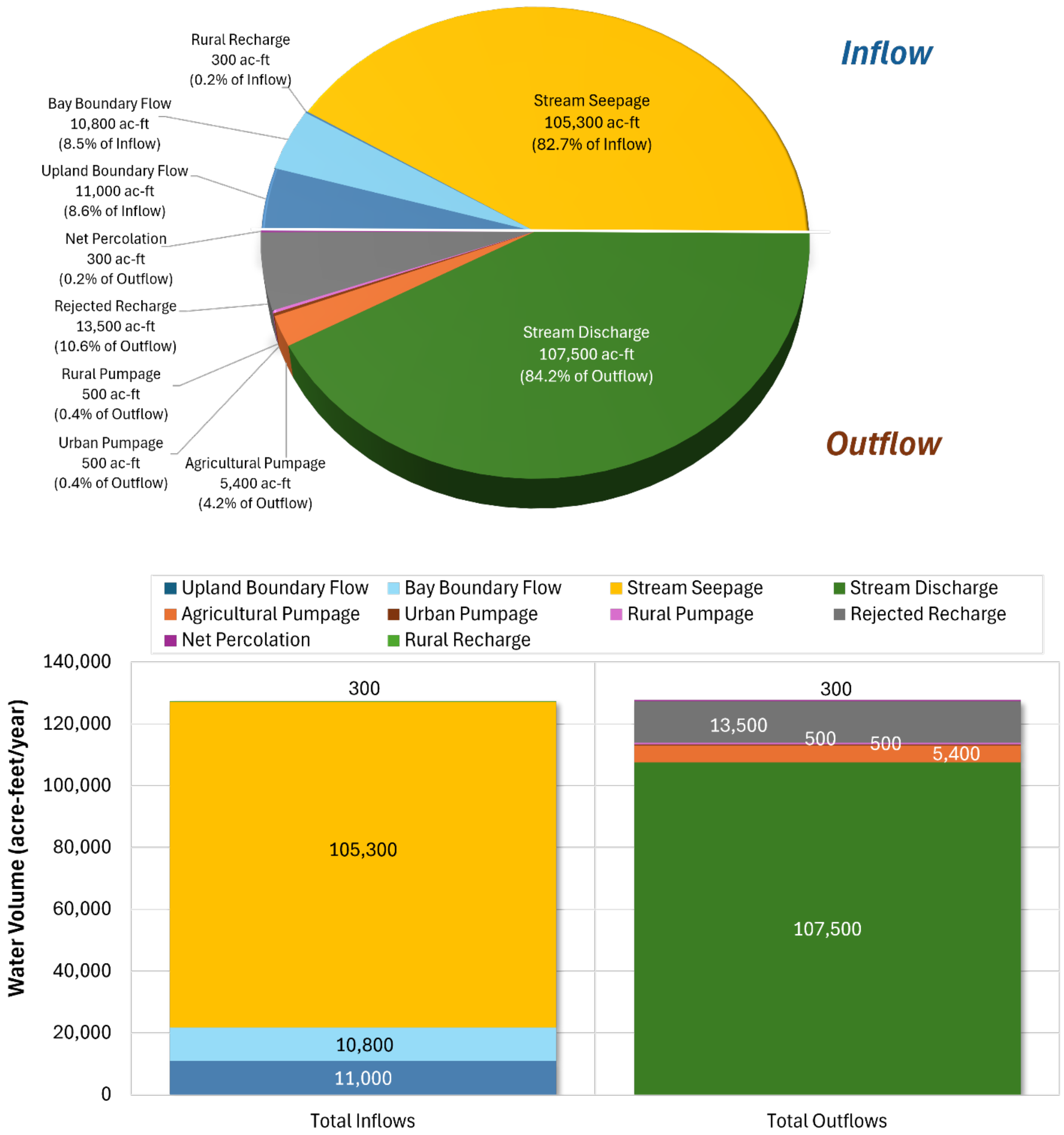

Figure 4-2. Annual Groundwater Budget Components as per the USGS Petaluma Model Report (Traum et al. 2022: Table D17).

4.2 Geology

The foregoing references provide coverage primarily within the groundwater basins delineated as per Bulletin 118. The water bearing units within the Petaluma Valley basin include quaternary alluvium and the Plio-Pleistocene Merced Formation, a massive fine grained marine sandstone. The water bearing units within the Novato Valley basin are alluvial deposits. The water bearing units within the Wilson Grove Highlands basin are alluvial deposits and the Pliocene Wilson Grove formation, a massive sand deposition within a subsiding embayment. The water bearing units within the Sonoma

Valley sub-basin include Quaternary alluvial deposits, with occasional wells also completed in the St. Helena rhyolite member of the Sonoma Volcanics. In the Petaluma River watershed outside the delineated basins, the bedrock is composed of the Sonoma Volcanics to the east (a group of andesitic to basaltic rocks) and the Cretaceous Great Valley Sequence to the west (a group of marine mudstones) as shown by the California Geological Survey in their 1982 regional Santa Rosa map (California Division of Mines and Geology 1982). The Bulletin 118 delineations do not account for any potential sources of 'non-basin' water within weathered bedrock formations, fractures, or other void spaces outside or underneath the designated basins.

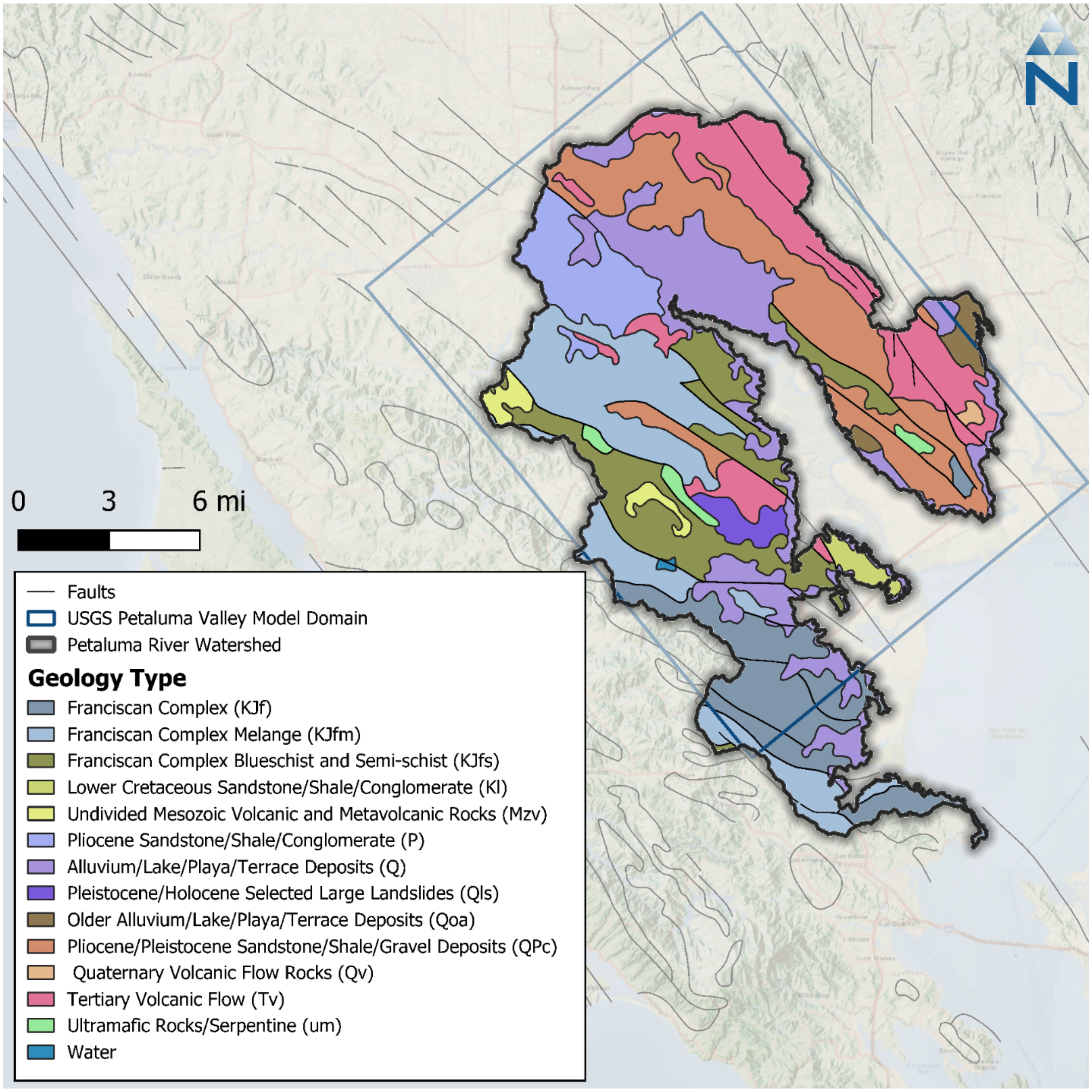

Figure 4-3. Geology map of the Petaluma River Watershed including the USGS Petaluma Valley Model boundary.

5 LANDSCAPE CHARACTERIZATION

Landscape characterization describes the physical characteristics of the landscape including the types of soils and geology, topography, land cover, land use, and other physical properties that can be represented within the hydrological model. Hydrologic Response Units (HRUs) are the core landscape unit in a watershed model. Each HRU represents areas of similar physical characteristics attributable to certain hydrologic processes. Spatial or geological characteristics such as land cover, soils, geology, and slopes are typically used to define HRUs. The spatial combinations of these various characteristics ultimately determine the number of meaningful HRU categories considered for the model. The following sections describe the component layers available to derive HRUs for the Petaluma River watershed.

5.1 Elevation & Slope

The USGS publishes DEMs expressing landscape elevation through a raster grid data product with 30-meter resolution. The Petaluma River watershed ranges in elevation from near sea level along the coasts and valleys and to over 700 meters at Sonoma Mountain in the north. It should be noted that the lowest elevation values (-103m) correspond to the San Rafael Rock Quarry at the southernmost tip of the watershed. This area will be further evaluated under Task 3.2 but may not be represented as a quarry in the model; given the quarry's location, it is not expected to impact representation of the watershed's hydrology. As a geoprocessing input, the DEM can be used to derive both slope and aspect as data inputs to a model. Figure 5-1 shows the change in elevation across the Petaluma River watershed.

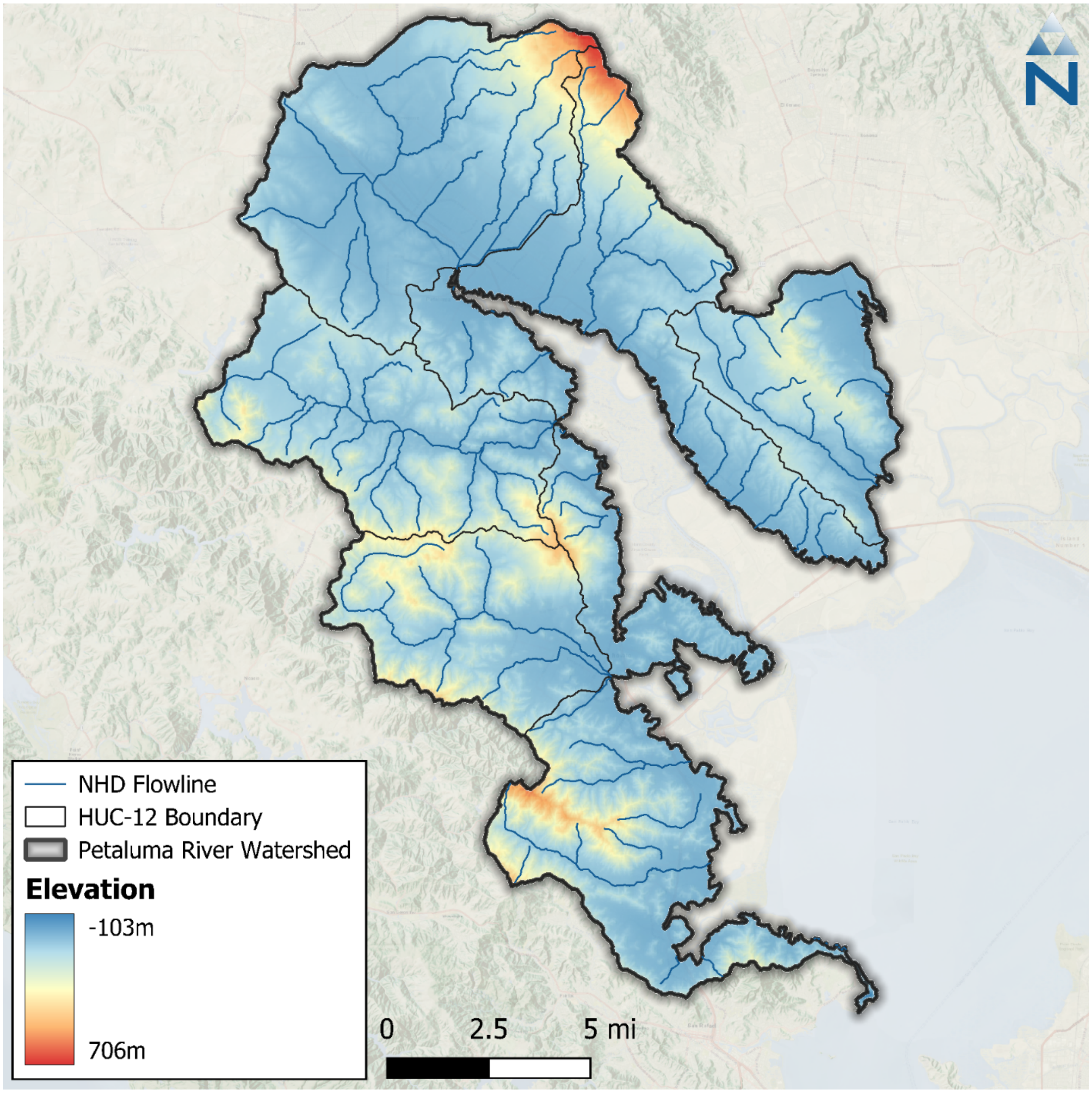

Figure 5-1. Digital Elevation Model of the Petaluma River watershed.

5.2 Soils

Soils data for the Petaluma River watershed were obtained from the Soil Survey Geographic Database (SSURGO) (USDA 2024a) and State Soil Geographic Database (STATSGO2) (USDA 2024b); both are published by the Natural Resource Conservation Service (NRCS).

There are four primary hydrologic soil groups (HSG) used to characterize soil runoff potential. Group A generally has the lowest runoff potential (e.g., sands) whereas Group D has the highest runoff potential (e.g., clays). Both SSURGO and STATSGO2 soils databases are composed of a GIS polygon layer of map units and a linked database with multiple layers of soil property. Soil characteristics for predominant hydrologic soil groups are described in Table 5-1.

Table 5-1. NRCS Hydrologic soil group descriptions

Hydrologic Soil Group	Description
A	Sand, Loamy Sand, or Sandy Loam
B	Silt, Silt Loam or Loam
C	Sandy Clay Loam
D	Clay Loam, Silty Clay Loam, Sandy Clay, Silty Clay, or Clay

Source: Natural Resource Conservation Service (NRCS), Technical Release 55 (TR-55) (USDA 1986) .

Table 5-2 provides a summary of areas occupied by each SSURGO HSG, and Figure 5-2 shows the spatial distribution of these groups throughout the Petaluma River watershed. The dominant soil group in the watershed is Group C (35%), containing sandy clay loams that typically have low infiltration rates. Group B (22%) is the next most common soil group in the watershed, containing moderately well to well-drained silt loams and loams. Group D also makes up 22% of the watershed, with the lowest infiltration rates, containing clay loam, silty clay loam, sandy and silty clay, and clay. Group A, containing well-draining sand, loamy sand, and sandy loam, makes up only 0.5%. Twelve percent of the watershed areas have mixed soil groups. For modeling purposes, mixed soils will be combined with the nearest primary group as follows: A/D → B, B/D → C, and C/D → D. Finally, approximately 8% of the watershed HSG area is classified as unknown in the SSURGO database and reside primarily within developed areas or correspond with waterbodies. For these areas, the corresponding HSG from the STATSGO2 dataset will be used to supplement the data gaps.

Table 5-2. NRCS Hydrologic soil groups in the Petaluma River watershed

Hydrologic Soil Group	Area (acres)	Percent Area
A	624.16	0.49%
B	27,872.81	22.01%
C	44,950.58	35.49%
C/D	15,734.91	12.42%
D	27,422.29	21.65%
N/A	10,042.50	7.93%
Total	**126,647.24**	**100.0%**

Source: State Soil Geographic and Soil Survey Geographic Database (STATSGO2/SSURGO)

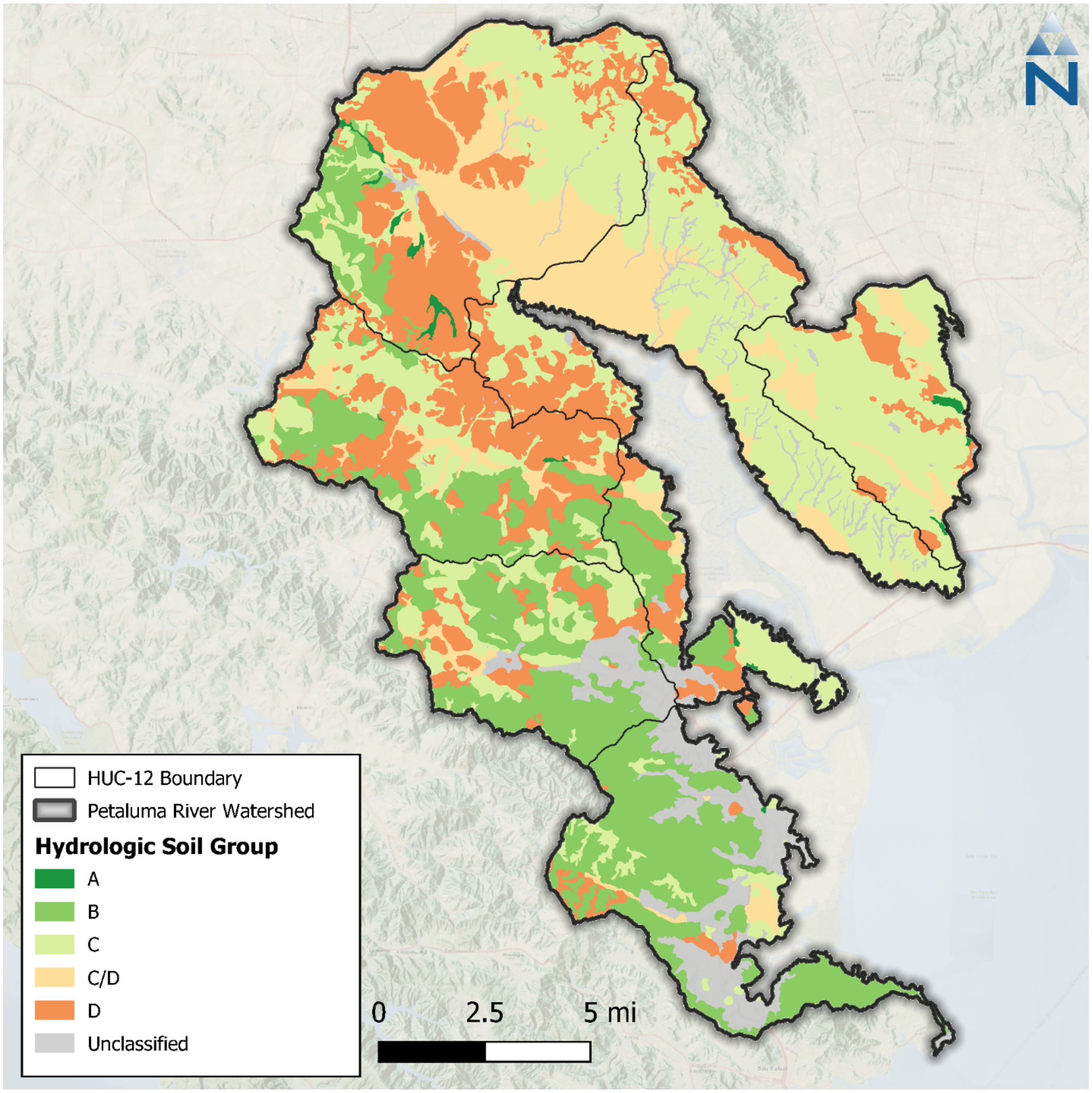

Figure 5-2. SSURGO hydrologic soil groups within the Petaluma River watershed.

5.3 Land Cover

Land cover data are a key layer for HRUs. The primary source of land cover data identified for this effort is the 2021 National Land Cover Database (NLCD) maintained by the Multi-Resolution Land Consortium (MRLC), a joint effort between multiple federal agencies. The primary objective of the MRLC NLCD is to provide a current data product in the public-domain with a consistent characterization of land cover across the United States. The first iteration of the NLCD dataset was in 1992. Since the 2001 NLCD version, a consistent 16-class land cover classification scheme has been adopted nationwide. The 2021 NLCD adopted this 16-class scheme at a 30-meter grid resolution.

Table 53. summarizes areal coverage of land use classes from a subset of the 2021 NLCD dataset that covers the Petaluma River watershed, and Figure 53 shows the spatial distribution of these

classifications. Grassland/Herbaceous is the dominant land cover class covering approximately 48% of the watershed. When combined, deciduous forest, evergreen forest, mixed forest, shrub/scrub, and grassland/herbaceous account for 73% of the total watershed area. Developed land cover makes up approximately 23% of the total watershed area and is roughly evenly split between Open Space, Low Intensity Development, and Medium Intensity Development. Less than 2% of the total watershed area is cultivated crop land, which potentially underestimates the true cultivated area because many individual cultivated areas in the watershed may be smaller than the NCLD's 2.7-acre minimum mapping unit.

Table 5-3. National Land Cover Database 2021 land cover summary in the Petaluma River watershed

NLCD Class	Classification Description	Area (acres)	Percent
11	Open Water	431.47	0.34%
21	Developed, Open Space[1]	8,361.85	6.60%
22	Developed, Low Intensity[1]	8,229.07	6.50%
23	Developed, Medium Intensity[1]	10,100.63	7.98%
24	Developed, High Intensity[1]	2,150.46	1.70%
31	Barren Land (Rock/Sand/Clay)	111.20	0.09%
41	Deciduous Forest	1,033.30	0.82%
42	Evergreen Forest	9,368.24	7.40%
43	Mixed Forest	9,006.83	7.11%
52	Shrub/Scrub	12,429.02	9.81%
71	Grassland/Herbaceous	61,275.91	48.38%
81	Pasture/Hay	921.21	0.73%
82	Cultivated Crops	2,311.92	1.83%
90	Woody Wetlands	142.79	0.11%
95	Emergent Herbaceous Wetlands	777.76	0.61%
TOTAL[2]		**126,651.67**	**100%**

Source: 2021 National Land Cover Database

1. Imperviousness: Open Space (<20%); Low Intensity (20-49%); Medium Intensity (50-79%); High Intensity (≥80%).
2. Note that because of the raster resolution, this total is approximately 4 acres more than the model domain.

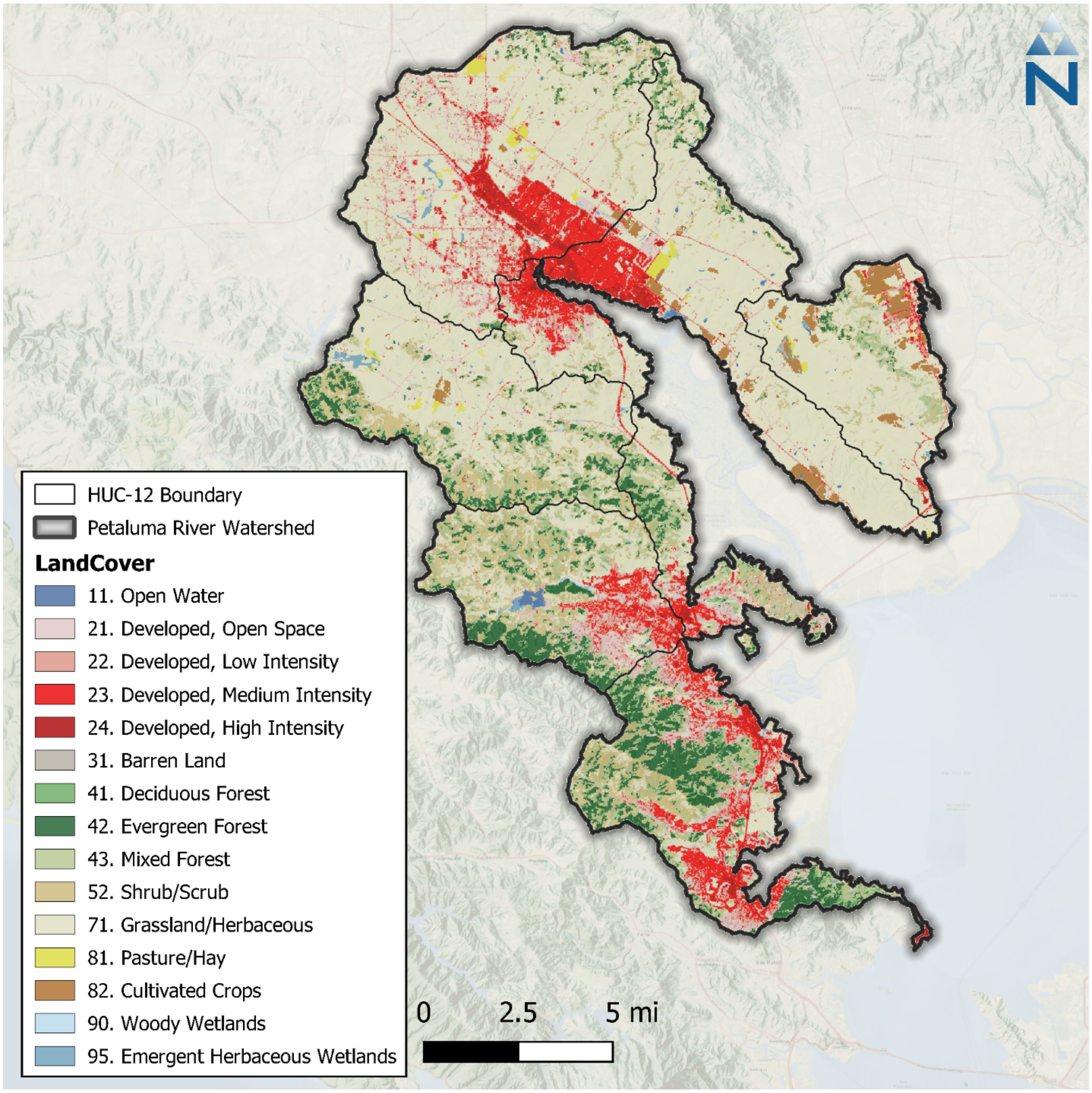

Figure 5-3. NLCD 2021 land cover within the Petaluma River watershed.

MRLC publishes a developed impervious cover dataset as a companion to the NLCD land cover; this dataset is also provided as a raster with a 30-meter grid resolution. Impervious cover is expressed in each raster pixel as a percentage of total area ranging from 0 to 100 percent. Because this dataset provides impervious cover estimates for areas classified as *developed,* non-zero values closely align with developed areas (NLCD classification codes 21 through 24). Review of the Petaluma River watershed using this dataset shows that just under 23% of the area is developed. The developed area is classified further into open space and low, medium, and high intensity development. Of those subcategories, open space and low intensity development make up most of the total developed area.

Because land cover can vary significantly over time due to anthropogenic changes (e.g., development, timber harvest) or naturally occurring events (e.g., forest fires, landslides), it may be necessary to also time-vary land cover through the model simulation or, at a minimum, align the dataset used to represent land cover with the same time period as streamflow data used for model calibration. The NLCD 1992, 2001, 2006, 2011, and 2021 snapshots are all available for representing land cover changes within the model depending on the period, or multiple periods, or time selected for model calibration and validation. Land use change in the Petaluma River watershed will be assessed as part of the model development, and a decision will be made based on the results as to whether land use change is represented explicitly, or a single land use snapshot is used.

Furthermore, the California Department of Forestry and Fire Protection (CAL FIRE) maintains databases of timber harvest plans and fire perimeters (see Table 1-3) which may be used in conjunction with the basic NLCD land cover snapshots to vary the land cover representing dynamic processes like timber harvests or episodic fire-related activities.

5.4 Tree Canopy Cover

MRLC publishes a tree canopy dataset as a companion to the NLCD land cover dataset that estimates the percentage of tree canopy cover spatially. The underlying data model was developed by the United States Forest Service (USFS) and is available through their partnership with the MRLC. This dataset is also provided as a raster with a 30-meter grid resolution. Like the impervious cover dataset, each raster pixel expresses the percent of the total area covered by tree canopy with values ranging from 0 to 100 percent. The percent tree canopy cover layer was produced by the USFS using a Random Forests regression algorithm (Housman et al. 2023). Across the Petaluma River watershed, an average of 12% of the total watershed area is covered by tree canopy. Tree canopy cover data can be used to estimate model parameters like interception storage and lower-zone ET rates.

5.5 Agriculture & Crops

Land cover data for the Petaluma River Watershed (see Section 5.3) was analyzed to identify predominant cropland vegetation classes. This analysis revealed that about 2.6% of the Petaluma River watershed area is classified as Pasture/Hay (class 81) or Cultivated Crops (class 82), and 58% of the watershed was classified as either Shrub/Scrub (class 52) or Grassland/Herbaceous (class 71); of the area that is classified as shrub or grassland, a portion may include areas of cultivated crops that were not automatically recognized through processing of the remote sensing data or include cultivated crops on a rotating schedule. To reflect these situations, supplemental information published by the United States Department of Agriculture (USDA) can be used. The USDA Cropland Data Layer (CDL) (USDA 2024) is an annual updated raster dataset that geo-references crop-specific land use. The dataset comes as 30-meter resolution raster with a linked lookup table of 85 standard crop types which can be used to classify agricultural land. The purpose of the CDL dataset is to provide a supplemental estimate of annual acreage used for major crop commodities. Figure 5-4 shows the spatial distribution of these classes through the study area, and Table 5-4 summarizes their areal coverage. Additionally, a large-scale crop and land use identification dataset for the year 2020 is made available by the California DWR (2022) and could be used to supplement data gaps if necessary. This dataset is intended to quantify crop acreage statewide and was constructed by analyzing remote sensing data gathered at the field scale.

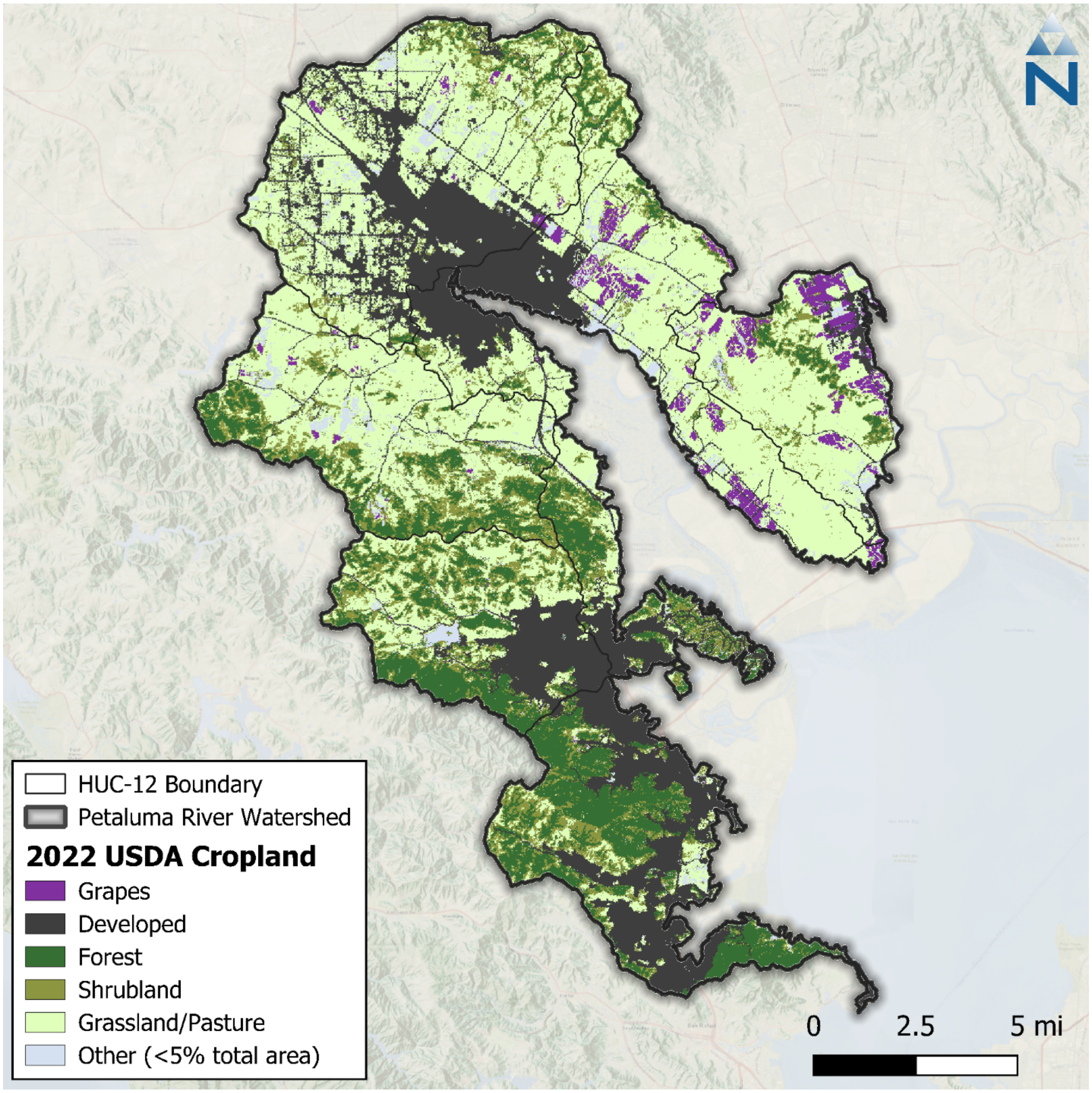

Figure 5-4. USDA 2022 Cropland Data within the Petaluma River watershed.

Table 5-4. USDA 2022 Cropland Data summary within the Petaluma River watershed

Crop Type	Area (ac)	Area (%)
Developed	28,788.64	22.73%
Forest	17,355.78	13.70%
Shrubland	16,993.93	13.42%
Grassland/Pasture	54,538.74	43.06%
Grapes	3,557.41	2.81%
Other (<5% Total Area)	5,412.06	4.27%
Total	**126,646.56**	**100.00%**

6 DATA GAPS AND LIMITATIONS

Based on a review of the hydrology datasets presented in Section 3.3, one potential limitation is the spatial extent of available daily streamflow data to support model calibration. USGS operates only one active gauge with daily streamflow records within the HUC-12 in the Novato Creek watershed. The calibration would need to be extrapolated outside of these tributaries to other waterbodies of interest unless other local gauge data is available.

Based on review of the datasets presented in this report, a potential limitation is the availability, quality, and temporal resolution of data for surface water diversions within the watershed. The eWRIMS database will be queried to identify major surface water diversions that are likely to have data to integrate into the model; however, other surface water diversions, such as water use to support agriculture, may not be mapped or have available data. These diversions may need to be mapped, and assumptions could be needed to represent water demand in the model if these demands are needed for model calibration purposes.

Another potential data gap is that observed streamflow records are only available in one HUC-12 (Novato Creek). HRU parameters above this station will be calibrated to match these observations but may not be reflective of the watershed as a whole. Any additional streamflow gauges, potentially operated by county or municipal governments or regional water quality control boards, could be incorporated to improve model calibration and validation. Similarly, at the time of writing it is unknown what stage-storage relationships or operational records or rules may be available for the six reservoirs identified in Section 3.4; obtaining additional information or deriving estimates may be necessary to achieve an acceptable calibration.

Some of the irrigation data gaps can be estimated from related data sources. For example, as previously mentioned in Section 2.2, the difference between OpenET (reasonable large-scale gridded estimate for observed) and model estimates of AET could be used to estimate water volume differences attributable to irrigation activity. Figure 6-1 shows an example of the monthly variation of OpenET and LSPC AET (based on the SFEI LSPC model) for the Napa River watershed which is also within the SFEI model domain. The green shaded area between OpenET and LSPC AET represents excess evapotranspiration volume—that volume would not have naturally occurred had it not been added to the system; this excess overlaps with the summer months where there is little to no rainfall, as shown in Figure 6-2. Irrigation volumes will most likely need to be slightly larger than that estimated difference because only a portion of irrigation volume would ultimately be expressed as AET depending on assumptions about the source of irrigation water, irrigation method, efficiency, and excess runoff from irrigated lands.

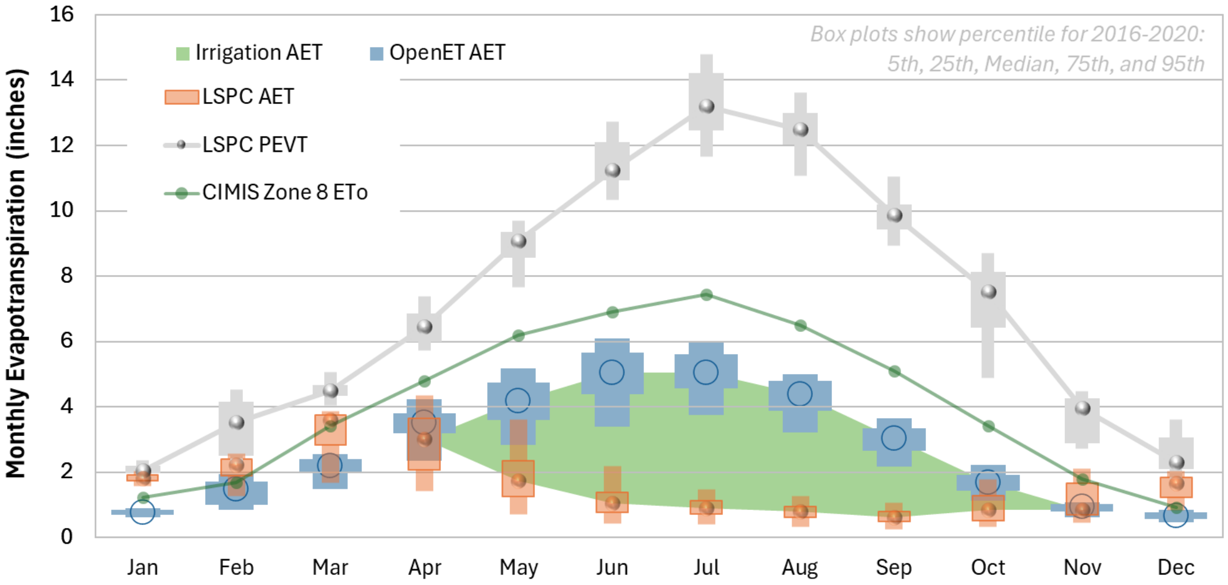

Figure 6-1. Estimating actual irrigation evapotranspiration (AET) from OpenET and modeled AET (data from the Napa River watershed).

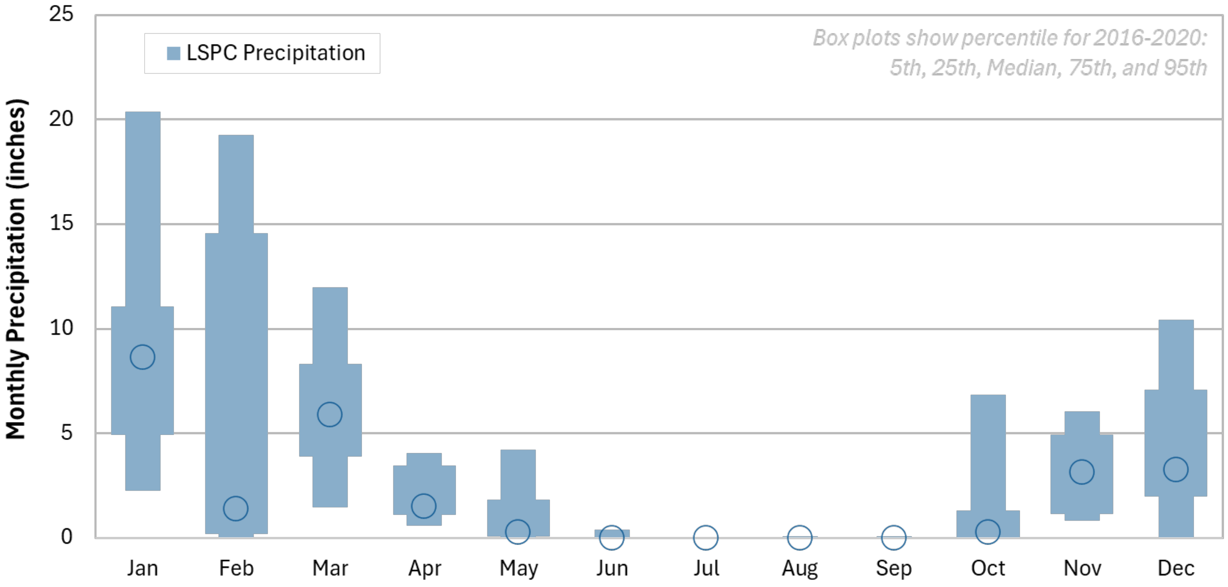

Figure 6-2. Monthly precipitation variability (data from the Napa River watershed).

7 MODEL CONFIGURATION

Model configuration encompasses model selection and data integration. Model selection considered not only available data and the ability of available models to address key study objectives, but also, considered how existing or on-going modeling efforts such as the SFEI watershed modeling efforts could be leveraged to address the specific objectives of this study (Section 1.3). This section elaborates further on model selection and model reconfiguration.

7.1 Model Selection

This modeling study's objectives influence hydrologic model selection and technical approach development. The available data presented in this work plan for characterizing the watershed also influence model selection. The key study objectives to be addressed with the selected hydrologic model are summarized below:

- Representation of unimpaired flows and baseline flows (e.g., water use and other human activities that impact instream flows and how they affect the water balance)
- The model simulation period should be long enough to capture variability between water years to represent conditions such as dry and wet year flows, environmental flows, drought curtailment, and other hydrological impacts.

To simulate streamflow, the model must be able to represent seasonal variability on the landscape and be responsive to both natural changes (e.g., meteorological conditions, vegetation cycles) and anthropogenic/hydromodification impacts (e.g., stream diversions, impoundments, groundwater pumping, timber harvest). An ideal platform should also be adaptable for simulating (1) spatial changes like those associated with representing pre-developed/unimpaired land cover states, (2) temporal changes like those associated with modeling climate change impacts, or (3) catastrophic impacts like those associated with extreme events such as 100-year storms and wildfires.

Public-domain models that can address those study objectives include the Hydrologic Simulation Program – Fortran (HSPF) (Barnwell and Johanson 1981), LSPC (Shen, Parker, and Riverson 2005), the Precipitation-Runoff Modeling System (PRMS) (Markstrom et al. 2015), and Soil and Water Assessment Tool (SWAT) (Neitsch et al. 2011). LSPC has been used extensively throughout California to model the unique hydrologic characteristics of the State's watersheds and to inform regulatory decisions (i.e., development of TMDLs and associated amendments to Water Quality Control Plans), watershed management, or climate change analyses. Watersheds in California where LSPC modeling has been conducted include those in the San Francisco Bay region (SCVURPPP 2019; SMCWPPP 2020; Zi et al. 2021 and 2022). the Clear Lake watershed in the Central Valley Region (CVRWQCB 2006), the Lake Tahoe watershed in the Lahontan Region (LRWQCB and NDEP 2010; Riverson et al. 2013), all coastal watersheds of Los Angeles County (LACFCD 2020; LARWQCB 2010, 2012, 2013a, 2013b, and 2015; LARWQCB and USEPA 2005a, 2005b, 2006, and 2011; Tariq et al. 2017), the San Jacinto River watershed in the Santa Ana Region (SAWPA 2003 and 2004), and most coastal watersheds of the San Diego Region (City of San Diego and Caltrans 2016; City of Vista 2008; Los Peñasquitos Responsible Agencies 2015; San Diego Bay Responsible Parties 2016; SDRWQCB 2008, 2010, and 2012). These efforts have included comprehensive peer review processes and public comment, requiring demonstration of model accuracy based on standard practices for quantifying and documenting model performance. All the modeling documentation and reports cited here have withstood peer review and have supported amendments to Water Quality Control Plans or the approval of watershed plans submitted to the Water Board or Regional Water Quality Control Boards to demonstrate regulatory compliance. Additionally, the Water Board recently utilized LSPC to perform hydrology analyses within the South Fork Eel River and Shasta River watersheds.

LSPC is a modernized version of the HSPF platform that is now organized around a Microsoft Access relational database; otherwise, the LSPC model is functionally identical to the HSPF model. The relational database provides efficient data management, model maintenance, and development of alternative scenarios. The LSPC model runs using hourly input boundary conditions and can be sufficiently configured using the meteorological datasets discussed in Section Meteorology. LSPC also has a feature that can vary land use over time when needed to explicitly represent dynamic processes such as timber harvests and wildfires—that feature needs supporting spatial and temporal data to represent dynamic land use changes. Additionally, LSPC was the selected modeling platform for two other Water Board studies performed for the South Fork Eel River and Shasta River watersheds. Those two watershed models utilize data from many of the same sources compiled in this work plan for the Petaluma River watershed. Based on the extensive history of successful LSPC model applications and its strengths and flexibility for potential coupling with a groundwater model (e.g., MODFLOW), LSPC is recommended as the watershed model for this study.

As described in Section 1.3, SFEI has developed an LSPC model for watersheds draining to San Francisco Bay with a spatial domain that includes the Petaluma River watershed (Zi et al. 2021 and 2022). The SFEI modeling objective was to support management decisions regarding stormwater runoff and pollutant loads including mercury, PCBs, and sediment. Because this model was constructed at the regional scale; the focus of model configuration and calibration differs slightly from the needs of this study; however, elements such as initial process parameter values provide a good foundation for building upon for this study.

7.2 Model Reconfiguration

The SFEI LSPC model will be reconfigured using the data sets presented in this work plan. A hydrologic analysis shall be developed with the primary goal of simulating instream flow time series for a minimum of 20 years through Water Year 2023 (10/1/2003 – 9/30/2023) and capable of representing both current/managed flow conditions and natural (pre-development) conditions. The following describes how major elements of the model will be structured using the available data sets. Further details about each process and underlying assumptions will be documented in a modeling report:

- **Climate Boundaries:** Climate forcing inputs to the model will include both precipitation and evapotranspiration. As was done for the SFEI model, precipitation will be represented using the 4-km gridded PRISM monthly precipitation, which provides an accurate representation of the long-term water balance. Monthly PRISM precipitation totals will be disaggregated using daily and hourly NCDC, RAWS, and CDEC observed timeseries. The SFEI model used area-weighted average monthly PRISM totals per subcatchment; however, for this effort, precipitation data will simply be assigned based on the grid cell with the largest areal coverage. Eliminating the step of area-weighting is not expected to have a significant impact given that the 4-km spatial resolution is already finely resolved compared to the average NHDPlus subcatchment size (~2-km^2). Another benefit of this reconfiguration is that it streamlines processing effort when (1) extending the timeseries in the future or (2) making updates to model subcatchment boundaries. It also establishes a consistent methodology across other modeled watersheds.

  Evapotranspiration will be represented using the CIMIS daily reference evapotranspiration 2-km gridded data set and disaggregated to hourly based using NLDAS. It is important to note that NLDAS potential evapotranspiration is reported at a 3-hour interval as shown in Figure 7-1, panel A. However, the hourly *distribution* of solar radiation from NLDAS, which has a sinusoidal pattern over daylight hours, offers a suitable alternative for disaggregating daily CIMIS depths (Figure 7-1, panel B). Daylight hours in NLDAS solar radiation also exhibit natural seasonal variation with latitude. The hourly distribution is derived by dividing the hourly solar radiation values for each day by the corresponding total solar radiation for the day. Those distributions are then multiplied by the total CIMIS evapotranspiration to disaggregate CIMIS daily to hourly.

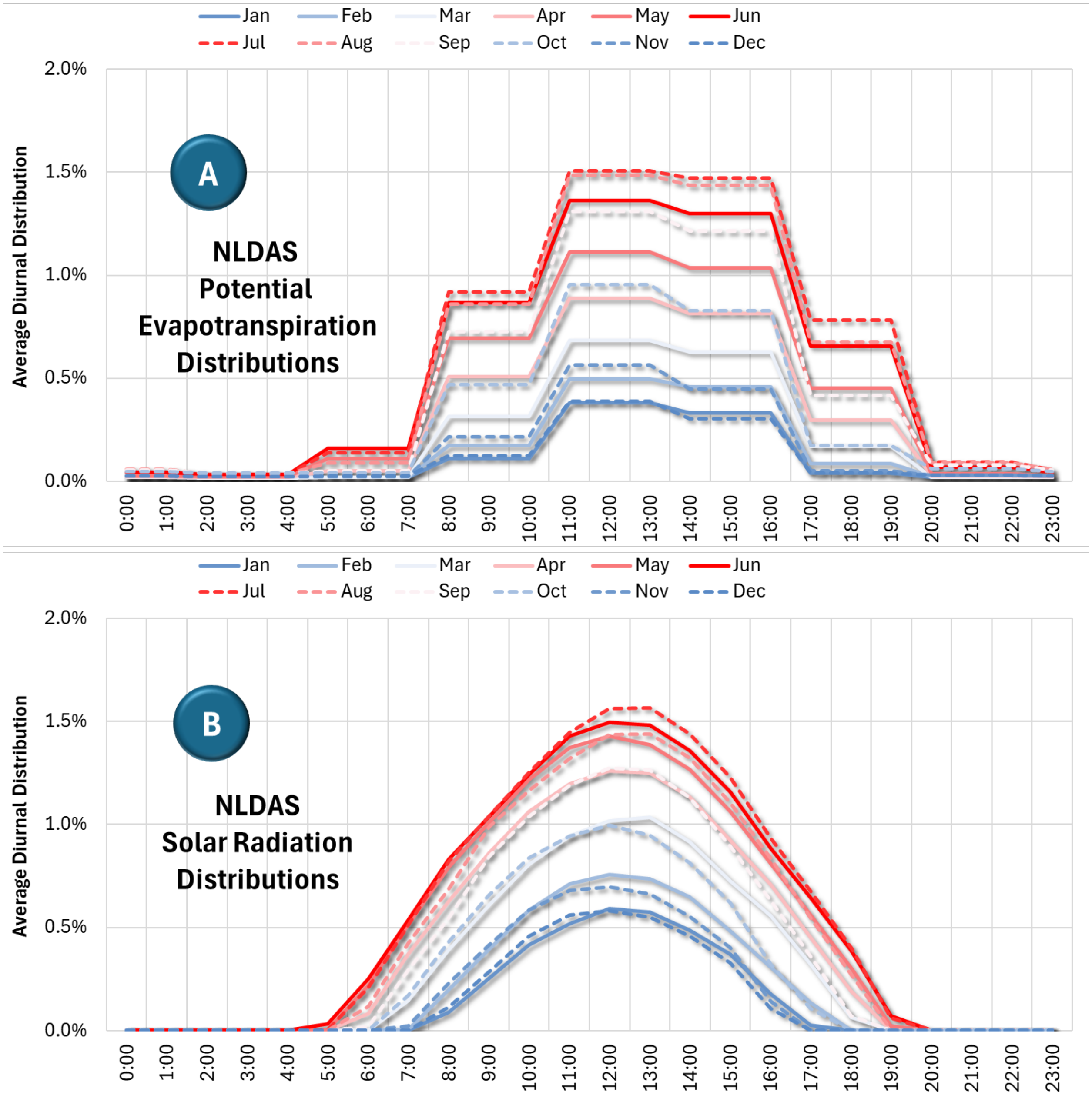

Figure 7-1. Comparison of monthly diurnal NLDAS potential evapotranspiration (A) and solar radiation (B) (data from NLDAS grids for the Napa River watershed).

- **Model Segmentation**: The SFEI subcatchment delineations were at a coarser resolution as part of a larger model of San Francisco Bay watersheds. As previously noted, area weighting PRISM grids by subcatchment is reasonable at the larger subcatchment scale, but is not as beneficial when the average subcatchment size is *smaller* than a PRISM grid. For this effort, subcatchment delineations will be based on HUC-12 boundaries and use NHDPlus catchment boundaries to subdivide the HUC-12 boundaries to represent key points of interest in the network (e.g., confluence of tributaries, station locations, points of diversion, and other points of interest). Up to one primary reach segment will be represented per subcatchment and will use a cross-section calculated using trapezoidal geometry as a function of cumulative upstream drainage area (Bent and Waite 2013; McCandless 2003; McCandless and Everett 2002). If additional cross-sectional information is available, these geometries can be updated per reach in the model.

- **Hydrologic Response Units**: HRUs represent unique combinations of landscape characteristics that will be derived by overlaying GIS data sets describing land cover, hydrologic soil group, and slope. The unique combinations of these three elements will form a set of HRUs that will be configured within the LSPC model. When crop type is known, this will be used to override the land cover data. As described in Section 5.5, "Grapes" are by far the dominant crop type in the Petaluma River watershed. In the final model configuration, some HRUs may be reclassified and grouped when appropriate for model parameterization (e.g., multiple types of forest may be grouped into a single "Forest" HRU category unless there is a reason to represent different responses in the model for each type). Because of the density of developed areas along the coastal regions of the Petaluma River watershed, the acreage of mapped impervious area (MIA) may be adjusted for calibration to effective impervious area (EIA), or to the portion of MIA which is directly connected to the conveyance network, using the Sutherland Equations. This refinement is necessary to avoid an initial overestimation of impervious surfaces contributing to runoff before initiating process-based model calibration (Sutherland 2000).
- **Water Budget**: To the extent that major sources of water use (e.g., irrigation, groundwater pumping, surface diversions) or inter-basin transfers are known, these volumes will be included as withdrawals or inputs to the model. Because grapes are the dominant crop type in the Petaluma River watershed, assumptions about irrigation will be configured accordingly. In cases where specific data are not available, reasonable assumptions may need to be made and documented for some sources/sinks, while others may need to be excluded entirely if the impact(s) on the model prediction cannot be quantified in a representative way. Priority will be given to features that directly influence predictions at points where the model is being compared to observed data for calibration purposes.

Table 7-1 summarizes key components of the SFEI LSPC model, proposed changes that would make it more suitable for this application, as well as reasoning for these changes. Given that accurate flow simulation and associated water budgeting are the overarching objectives of this workplan, the primary purpose of these proposed changes is to improve the model's ability to predict key water balance components (streamflow, evapotranspiration, interflow, and groundwater flow) to increase confidence in the output of potential water management scenarios that are simulated.

Table 7-1. Summary of SFEI model and proposed modifications

Data Group	Data Category	Reference Section[1]	Description	Data Set(s)[2]	Proposed Changes	Reasoning
Climate Boundaries	Precipitation (hourly)	2	Daily 4 km gridded precipitation (PRISM) is spatially bias corrected and temporally disaggregated using observed station networks and hourly 12 km gridded NLDAS-2 data; area weighted aggregation by subcatchment is conducted before temporal disaggregation.	PRISM (daily), NLDAS-2 (hourly) Observed[3]	Disaggregate monthly PRISM using Observed[3] data, then assign single grid cell to each subcatchment based on areal coverage.	Disaggregating raw PRISM grid cells (rather than area averages for subcatchments) improves efficiency when modifying meteorological forcing; consistent with our other modeling efforts.
	PET (hourly)	2.2	12 km resolution gridded dataset of hourly potential evapotranspiration.	NLDAS-2 (3-hourly)	Disaggregate daily ET_0 for CIMIS using hourly NLDAS-2 short-wave radiation.	NLDAS-2 overestimates evapotranspiration (Xia et al. 2014).
Model Segmentation	Watershed Boundaries	3	Subcatchments manually delineated from HUC-8 drainage boundaries.	WBD HUC-8	Use NHDPlus watershed boundaries and stream network.	Increased spatial resolution will improve routing accuracy; consistent with our other modeling efforts.
	Stream Network	3.2	Composite of stream segmentation datasets at varying spatial resolutions; intended to fill spatial gaps in stream network.	BAARI, NHD, WBD		
HRUs	Land use & Landcover	5.3	Composite of proprietary (ABAG) and open-source (NLCD) grid-based land characterization; Differentiates developed land from coarse classifications of forest, cropland, and wetlands; Identifies the age of development of existing infrastructure (ABAG).	ABAG, NLCD	Only use NLCD.	ABAG is a proprietary and requires special permissions whereas NLCD is open source; ABAG includes age of development data for infrastructure, which is useful for identifying contaminant point sources, but outside of the scope of this study.
	Soil	5.2	Composite of state-wide 10 m gridded and 1:1,000,000 vector-based datasets that identify primary hydrologic soil groups	SSURGO, STATSGO 2	-	-
	Elevation & Slope	0	10-m resolution gridded digital elevation map.	NED	-	-
	Impervious Areas	5.3	Broad, 30-meter grid-based land characterization; Represent percent impervious area within raster cells.	NLCD	-	-

Data Group	Data Category	Reference Section[1]	Description	Data Set(s)[2]	Proposed Changes	Reasoning
	Vegetation & Disturbances	5.4	Existing tree canopy cover	Missing	Include available open-source datasets as they fit within the project scope. Some potential sources include: MRLC, USFS, USDA, USGS, and CAL FIRE.	Improved accuracy of HRU construction and model parameterization; represents changes in land cover from disturbances.
		5.5	Existing vegetation and cropland classification			
		-	Timber harvesting records and GIS mapping			
Water Budget	Streamflow	3.3	Observed average daily streamflow.	USGS Streamflow (daily)	-	-
	Reservoir Operations	3.4	Geolocation and withdrawal time series of major reservoirs.	USGS, CDWR, SFPUC, EBMUD, SCVWD	-	-
	Diversions	3.5	Distributary and flood diversion channel routing	Missing	Include available open-source datasets as they fit within the project scope. Some potential sources include: DWR, SWRCB eWRIMS, and CDT.	Better inform parameterization to enhance accuracy of water budget; consistent with other modeling efforts.
	Withdrawals	3.5	Irrigation water usage			
			Wells and other groundwater withdrawal rates			
	Hydrogeology	4	Geospatial identification and description of geologic features used for groundwater routing.	UNK	Utilize blend of DWR datasets to identify key groundwater properties required for MODFLOW.	Data source is unknown; additional parameters are required for MODFLOW implementation (if necessary); consistent with other modeling efforts.

1. Points to the section in this workplan that the dataset is category is discussed.
2. For the full name of these acronyms refer to the *Acronyms* section of this report; for more information about how this source will be used and links to the dataset, refer to *Table 1-1*, *Table 1-2*, *Table 1-3*.
3. Observed precipitation station data reportedly used by SFEI includes: HPD, ISD, CIMIS, SCVWD, and GHCND. For more information on their usage, refer to Zi et al. 2021.

8 MODEL CALIBRATION

A combination of visual assessments and computed numerical evaluation metrics will be used to assess model performance during calibration. Model performance will be assessed using graphical comparisons or modeled vs. observed data (e.g., time-series plots, flow duration curves, cumulative distribution plots, and others) quantitative metrics and qualitative thresholds recommended by Moriasi et al. (2015) and Duda et al. (2012), which are considered highly conservative. Moriasi et al. (2007 and 2015) assign narrative grades for hydrology and water quality modeling to the percent bias (PBIAS), the ratio of the root mean square error to the standard deviation of measured data (RSR), and the Nash-Sutcliffe model efficiency (NSE). These metrics are defined as follows:

- The percent bias (PBIAS) quantifies systematic overprediction or underprediction of observations. A bias towards underestimation is reflected in positive values of PBIAS while a bias towards overestimation is reflected in negative values. Low magnitude values of PBIAS indicate better fit, with a value of 0 being optimal.
- The ratio of the root mean square error to the standard deviation of measured data (RSR) provides a measure of error based on the root mean square error (RMSE), which indicates error results in the same units as the modeled and observed data but normalized based on the standard deviation of observed data. Values for RSR can be greater than or equal to 0, with a value of 0 indicating perfect fit. Moriasi et al. (2007) provides narrative grades for RSR.
- The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance compared to the measured data variance (Nash and Sutcliffe 1970). NSE indicates how well the plot of observed versus simulated data fits the 1:1 line. Values for NSE can range between $-\infty$ and 1, with NSE = 1 indicating a perfect fit.

Other metrics can also be computed and used to assess calibrated model performance, including the Kling-Gupta Efficiency (KGE). This metric can provide additional or complementary information on model performance to the three metrics listed above and is defined as follows:

- The Kling-Gupta Efficiency (KGE) metric is based on the Euclidean Distance between an idealized reference point and a sample's bias, standard deviation, and correlation within a three-dimensional space (Gupta et al. 2009). KGE attempts to address documented shortcomings of NSE, but the two metrics are not directly comparable. A KGE value of 1 indicates perfect fit, with agreement becoming worse for values less than 1. Knoben, Freer, and Woods (2019) have suggested a KGE value > - 0.41 as a benchmark that indicates a model has more predictive skill than using the mean observed flow.

Both modeled time series and observed data will be binned into subsets of time to highlight seasonal performance and different flow conditions. Hydrograph separation will be performed to assess stormwater runoff vs. baseflow periods to isolate model performance on stormflows and low flows. Table 8-1 is a summary of performance metrics that will be used to evaluate hydrology calibration; as shown in this table, "All Conditions" (i.e., annual interval) for R-squared and NSE is the primary condition typically evaluated during model calibration. For sub-annual intervals, the pattern established in the literature for PBIAS when going from "All Conditions" to sub-annual intervals is to shift the qualitative assessment by one category (e.g., use the "good" range for "very good," "fair" for "good," and so on). This pattern will also be followed for R-squared and NSE qualitative assessments of sub-annual intervals.

The LSPC calibration performance in the Petaluma River watershed will be assessed to see if linkage of the LSPC model with a groundwater model (e.g., MODFLOW) could improve performance and process interactions. This could be manifested through a significant mismatch between the simulated and observed baseflow during dry periods. Other indicators include the mismatch between the simulated and observed hydrograph shape, demonstrating significant flow timing and magnitude differences.

Table 8-1. Summary of performance metrics used to evaluate hydrology calibration

Performance Metric	Hydrological Condition	Performance Threshold for Hydrology Simulation			
		Very Good	Good	Fair	Poor
Percent Bias (PBIAS)	All Conditions[1]	<5%	5% - 10%	10% - 15%	>15%
	Seasonal Flows[2]	<10%	10% - 15%	15% - 25%	>25%
	Highest 10% of Daily Flow Rates[3]				
	Days Categorized as Storm Flow[4]				
	Days Categorized as Baseflow[4]				
RMSE – Std Dev Ratio (RSR)	All Conditions[1]	≤0.50	0.50 - 0.60	0.60 - 0.70	>0.70
	Seasonal Flows[2]	≤0.40	0.40 - 0.50	0.50 - 0.60	>0.60
Nash-Sutcliffe Efficiency (NSE)	All Conditions[1]	>0.80	0.70 - 0.80	0.50 - 0.70	≤0.50
	Seasonal Flows[2]	>0.70	0.50 - 0.70	0.40 - 0.50	≤0.40
Kling-Gupta Efficiency (KGE)	Monthly Aggregated[5]	≥0.90	0.90 - 0.75	0.75 - 0.50	<0.50

1. All Flows considers all daily time steps in the model time series.
2. Seasonal Flows consider daily flows during a predefined, seasonal period (e.g., Wet Season and Dry Season). The Wet Season includes the months of October through April. The Dry Season includes the months of May through September.
3. Highest 10% of Flows considers the top 10% of daily flows by magnitude as determined from the observed flow duration curve.
4. Baseflows and Storm flows were determined from analyzing the daily model time series by applying the USGS hydrograph separation approach (Sloto and Crouse 1996).
5. KGE evaluated using thresholds for monthly aggregated time series (Kouchi et al. 2017).

9 SUMMARY & NEXT STEPS

This work plan presented the available data and proposed methods for developing a hydrologic model of the Petaluma River watershed. Once this work plan is finalized, the data sets described in this memo will be used to develop an LSPC model as described in Section 7. After finalizing the work plan, the first step of that process will be to present and finalize watershed boundaries and subcatchment delineations that capture key points of interest in the watershed (e.g., tributary confluences, streamflow station locations, etc.). Once built, this model will be calibrated using the metrics presented in Section 8 and documented in a model development report. Table 9-1. presents a summary of the deliverables planned for the Petaluma River watershed.

Table 9-1. Proposed schedule and summary of deliverables

Task	Subtask	Deliverable	Due Date
2	2.2	Draft Work Plan	8/1/2025
	2.3	Final Work Plan	Two (2) weeks after receiving comments
3	3.1	Subbasin delineation and stream GIS files	Two (2) weeks after completing Task 2.3
	3.2	LSPC database, model inputs, and GIS files[1]	Eighteen (18) weeks after completing Task 3.1
4	4.1	Draft Calibration Slide Deck	Twelve (12) weeks after completing Task 3.2
		Final Calibration Slide Deck	Six (6) weeks after receiving comments on Draft Calibration Slide Deck
5	5.1	Partial Draft Model Development Report[1]	Twelve (12) weeks after completing Task 3.1
		Draft Model Development Report	Six (6) weeks after completing Task 3.2
	5.2	Final Model Development Report	Four (4) weeks after receiving comments on Task 5.1 Draft MDR
	5.3	Final LSPC Model Code & Software	Two (2) weeks after Task 5.2
	5.4	Final Model Files including LSPC executable, LSPC database, LSPC model inputs, final GIS files	Two (2) weeks after Task 5.2

1. Partial Draft Model Development Report under Task 5.1 will be delivered in conjunction with Task 3.2 to document the model configuration.

10 REFERENCE

Allen, R. G., Walter, I. A, Elliott, R., Howell, T., Itenfisu, D., Jensen, M. 2005. The ASCE Standardized Reference Evapotranspiration Equation.

Barnwell, T.O., and Johanson. R. 1981. HSPF: A Comprehensive Package for Simulation of Watershed Hydrology and Water Quality.

Behnke, R., Vayrus. S., Allstadt. A., Allbright. T., Thogmartin. W. E., and Radeloff. V. C. 2016. Evaluation of Downscaled, Gridded Climate Data for the Conterminous United States. Ecological Applications 26(5): 1338–51. https://doi.org/10.1002/15-1061.

Bent, G. C., and Waite, A. M. 2013. Equations for Estimating Bankfull Channel Geometry and Discharge for Streams in Massachusetts. U.S. Geological Survey Scientific Investigations Report 2013–5155: 62. https://doi.org/10.3133/sir20135155.

California Division of Mines and Geology. 1982. Geologic Map of California, Santa Rosa Sheet. California Department of Conservation.

City of San Diego and Caltrans. 2016. Mission Bay Watershed Management Area Water Quality Improvement Plan. Submitted to the San Diego Regional Water Quality Control Board by the City of San Diego and Caltrans. San Diego, CA.

City of Vista. 2008. Agua Hedionda Watershed Management Plan – Final. Prepared by Tetra Tech for the City of Vista, CA.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. 2003. Real-Time and Retrospective Forcing in the North American Land Data Assimilation System (NLDAS) Project. Journal of Geophysical Research: Atmospheres 108(22): 8842. doi:10.1029/2002jd003118.

CRWQB (California Regional Water Quality Control Board). 2018. Total Maximum Daily Load for Bacteria and Nutrient Impairment Analysis in Petaluma River.

CVRWQCB (Central Valley Regional Water Quality Control Board). 2006. Amendment to the Water Quality Control Plan for the Sacramento River and San Joaquin River Basins for The Control of Nutrients in Clear Lake. California Water Quality Control Board, Central Valley Region. Rancho Cordova, CA.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology, 28(15), 2031–2064. https://doi.org/10.1002/joc.1688.

Daly, C., Neilson, R. P., & Phillips, D. L. 1994. A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. Journal of Applied Meteorology, 33(2), 140–158. https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Duda, P. B., Hummel, P. R., Donigian Jr., A. S., & Imhoff, J. C. 2012. BASINS/HSPF: Model Use, Calibration, and Validation. Transactions of the ASABE, 55(4), 1523–1547. https://doi.org/10.13031/2013.42261.

DWR (California Department of Water Resources). 1982. California Irrigation Management Information System.

DWR (California Department of Water Resources). 2020. California's groundwater—Update 2020: Sacramento, Calif., California Department of Water Resources Bulletin 118, 485 p. Accessed September 2023 from https://water.ca.gov/Programs/Groundwater-Management/Bulletin-118.

DWR (California Department of Water Resources). 2022. Statewide Crop Mapping. Accessed April 2024 from https://data.cnra.ca.gov/dataset/statewide-crop-mapping.

DWR (California Department of Water Resources). 2025a. Petaluma Valley GSA. Accessed July 2025 from https://sgma.water.ca.gov/portal/gsa/print/350

DWR (California Department of Water Resources). 2025b. Sonoma Valley GSA. Accessed July 2025 from https://sgma.water.ca.gov/portal/gsa/print/263

Gibson, W.P., Daly, C., Kittel, T., Nychka, D., Johns, C., Rosenbloom, N., McNab, A., and Taylor, G. 2002. Development of a 103-Year High-Resolution Climate Data Set for the Conterminous United States.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. 2009. Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. Journal of Hydrology 377(1–2): 80–91. doi:10.1016/j.jhydrol.2009.08.003.

Housman, I., K. Schleeweis, J. Heyer, B. Ruefenacht, S. Bender, K. Megown, W. Goetz, and S. Bogle. 2023. National Land Cover Database Tree Canopy Cover Methods. Salt Lake City, Utah. https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/docs/TCC_v2021-4_Methods.pdf (April 28, 2024).

Knoben, W. J. M., Freer, J. E. and Woods, R. A. 2019. Technical Note: Inherent Benchmark or Not? Comparing Nash–Sutcliffe and Kling–Gupta Efficiency Scores. Hydrology and Earth System Sciences 23(10): 4323–31. https://doi.org/10.5194/hess-23-4323-2019.

LACFCD (Los Angeles County Flood Control District). 2020. WMMS Phase I Report: Baseline Hydrology and Water Quality Model. Prepared for the Los Angeles County Flood Control District by Paradigm Environmental. Alhambra, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board). 2010. Los Angeles River Watershed Bacteria Total Maximum Daily Load. California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board). 2012. Reconsideration of Certain Technical Matters of the Malibu Creek and Lagoon Bacteria TMDL. California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board). 2013a. Reconsideration of Certain Technical Matters of the Santa Monica Bay Beaches Bacteria TMDLs; Marina Del Rey Harbor Mothers' Beach and Back Basins TMDL; and the Los Angeles Harbor Inner Cabrillo Beach and Main Ship Channel Bacteria TMDL. California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board). 2013b. Reconsideration of Certain Technical Matters of the TMDL for Bacteria Indicator Densities in Ballona Creek, Ballona Estuary, and Sepulveda Channel. California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board). 2015. Total Maximum Daily Loads for Indicator Bacteria in San Gabriel River, Estuary and Tributaries. California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board) and USEPA (U.S. Environmental Protection Agency). 2005a. Total Maximum Daily Load for Metals in Ballona Creek. California Regional Water Quality Control Board, Los Angeles Region, and U.S. Environmental Protection Agency Region 9. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board) and USEPA (U.S. Environmental Protection Agency). 2005b. Total Maximum Daily Loads for Metals Los Angeles River and Tributaries. U.S. U.S. Environmental Protection Agency, Region 9 and California Regional Water Quality Control Board, Los Angeles Region. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board) and USEPA (U.S. Environmental Protection Agency). 2006. Total Maximum Daily Loads for Metals and Selenium, San Gabriel River and Tributaries. California Regional Water Quality Control Board, Los Angeles Region, and U.S. Environmental Protection Agency, Region 9. Los Angeles, CA.

LARWQCB (Los Angeles Regional Water Quality Control Board) and USEPA (U.S. Environmental Protection Agency). 2011. Dominguez Channel and Greater Los Angeles and Long Beach Harbor Waters Toxics Total Maximum Daily Loads. California Regional Water Quality Control Board, Los Angeles Region, and U.S. Environmental Protection Agency Region 9. Los Angeles, CA.

Los Peñasquitos Responsible Agencies. 2015. Los Peñasquitos Watershed Management Area Water Quality Improvement Plan and Comprehensive Load Reduction Plan. Submitted to the San Diego Regional Water Quality Control Board by the City of San Diego, County of San Diego, City of Del Mar, Caltrans, and City of Poway. San Diego, CA.

LRWQCB (Lahontan Regional Water Quality Control Board) and NDEP (Nevada Division of Environmental Protection). 2010. Lake Tahoe Total Maximum Daily Load Technical Report. California Regional Water Quality Control Board, Lahontan Region. South Lake Tahoe, CA.

Markstrom, S. L., Regan, R. S., Hay, L. E., Viger, R. J., Webb, R. M., Payn, R. A., and LaFontaine, J. H. 2015. PRMS-IV, the Precipitation-Runoff Modeling System, Version 4.

McCandless, T.L. 2003. Maryland Stream Survey: Bankfull Discharge and Channel Characteristics in the Allegheny Plateau and the Valley and Ridge Hydrologic Region. Annapolis, MD.

McCandless, T.L., and Everett, R.A. 2002. Maryland Stream Survey: Bankfull Discharge and Channel Characteristics in the Piedmont Hydrologic Region. Annapolis, MD.

Melton, F. S., Huntington, J., Grimm, R., Herring, J., Hall, M., Rollison, D., Erickson, T., et al. 2022. OpenET: Filling a Critical Data Gap in Water Management for the Western United States.JAWRA Journal of the American Water Resources Association 58(6): 971–94. doi:10.1111/1752-1688.12956.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., et al. 2004. The Multi-Institution North American Land Data Assimilation System (NLDAS): Utilizing Multiple GCIP Products and Partners in a Continental Distributed Hydrological Modeling System. Journal of Geophysical Research: Atmospheres 109(7). doi:10.1029/2003jd003823.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50(3): 885–900. doi:10.13031/2013.23153.

Moriasi, D. N., Gitau, M. W., Pai, N., and Daggupati, P. 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Transactions of the ASABE 58(6): 1763–85. doi:10.13031/trans.58.10715.

Nash, J. E., and Sutcliffe, J. V. 1970. River Flow Forecasting through Conceptual Models Part I - A Discussion of Principles. Journal of Hydrology 10(3): 282–90. doi:10.1016/0022-1694(70)90255-6.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., and Williams, J.R. 2011. Soil and Water Assessment Tool Theoretical Documentation.

NID (National Inventory of Dams). 2025. USACE (United States Army Corps of Engineers): Dams of the Nations. Accessed April 2025 from https://nid.sec.usace.army.mil/#/.

NOAA (National Oceanic and Atmospheric Administration). 2022. National Marine Fisheries Service. Endangered Species Act (ESA) Section 7(a)(2) Biological Opinion for the Issuance of an ESA Section 10(a)(1)(A) Scientific Research and Enhancement Permit associated with the Rescue and Rearing Management Plan for Petaluma River Steelhead.

Riverson, J., Coats, R., Costa-Cabral, M., Dettinger, M., Reuter, J., Sahoo, G., and Schladow, G. 2013. Modeling the Transport of Nutrients and Sediment Loads into Lake Tahoe under Projected Climatic Changes. Climatic Change 116(1): 35–50. doi:10.1007/s10584-012-0629-8.

San Diego Bay Responsible Parties. 2016. San Diego Bay Watershed Management Area Water Quality Improvement Plan. Submitted to the San Diego Regional Water Quality Control Board by the San Diego Bay Responsible Parties. San Diego, CA.

SAWPA (Santa Ana Watershed Project Authority). 2003. Lake Elsinore and Canyon Lake Nutrient Source Assessment. Prepared by Tetra Tech for the Santa Ana Watershed Project Authority. Riverside, CA.

SAWPA (Santa Ana Watershed Project Authority). 2004. San Jacinto Nutrient Management Plan. Prepared by Tetra Tech for the Santa Ana Watershed Project Authority. Riverside, CA.

SCVURPPP (Santa Clara Valley Urban Runoff Pollution Prevention Program). 2019. Santa Clara Valley Reasonable Assurance Analysis Addressing PCBs and Mercury: Phase II Green Stormwater Infrastructure Modeling Report. Prepared by Paradigm Environmental for the Santa Clara Valley Urban Runoff Pollution Prevention Program. Sunnyvale, CA.

SDRWQCB (San Diego Regional Water Quality Control Board). 2008. Total Maximum Daily Loads (TMDLs) for Copper, Lead, and Zinc in Chollas Creek. California Regional Water Quality Control Board, San Diego Region. San Diego, CA.

SDRWQCB (San Diego Regional Water Quality Control Board). 2010. Revised Total Maximum Daily Loads for Indicator Bacteria Project I – Twenty Beaches and Creeks in the San Diego Region (Including Tecolote Creek). California Regional Water Quality Control Board, San Diego Region. San Diego, CA.

SDRWQCB (San Diego Regional Water Quality Control Board). 2012. Los Peñasquitos Lagoon Sediment/Siltation TMDL. California Regional Water Quality Control Board, San Diego Region. San Diego, CA.

SFEI (San Francisco Estuary Institute). 2025. Future Climate Impacts to Riparian Condition & Functions in the Petaluma River Watershed. Publication #1235. San Francisco Estuary Institute, Richmond, CA.

SFBRWQCB (San Francisco Bay Regional Water Quality Water Quality Control Board). 2020. Total Maximum Daily Load for Bacteria in Petaluma River Watershed. California Regional Water Quality Control Board, San Francisco Region. San Francisco, CA.

Shen, J., Parker, A., and Riverson, J. 2005. A New Approach for a Windows-Based Watershed Modeling System Based on a Database-Supporting Architecture. Environmental Modelling & Software 20(9): 1127–38. doi:10.1016/j.envsoft.2004.07.004.

Sloto, R.A., Crouse, and M.Y. 1996. HYSEP: A Computer Program for Streamflow Hydrograph Separation and Analysis: U.S. Geological Survey Water-Resources Investigations Report 1996–4040. doi:10.3133/wri964040.

SMCWPPP (San Mateo Countywide Water Pollution Prevention Program). 2020. San Mateo County-Wide Reasonable Assurance Analysis Addressing PCBs and Mercury: Phase II Green Infrastructure Modeling. Prepared by Paradigm Environmental and Larry Walker Associates for the San Mateo Countywide Water Pollution Prevention Program. Redwood City, CA.

SRCD (Sonoma Resource Conservation District). 2025. Watershed Overview. Accessed April 2025 from https://sonomarcd.org/district-watersheds/petaluma-river/.

Sutherland, R. C. 2000. Methods for Estimating the Effective Impervious Area of Urban Watersheds, Technical Note 58. In The Practice of Watershed Protection, eds. T. R. Scueler and H.K. Holland. Ellicot City, MD: Center for Watershed Protection, 193–95.

Tariq, A., Lempert, R. J., Riverson, J., Schwartz, M., and Berg, N. 2017. A Climate Stress Test of Los Angeles' Water Quality Plans. Climatic Change 144(4): 625–39. doi:10.1007/s10584-017-2062-5.

Traum, J.A., Teague, N.F., Sweetkind, D.S., and Nishikawa, T. 2022, Hydrologic and geochemical characterization of the Petaluma River watershed, Sonoma County, California: U.S. Geological Survey Scientific Investigations Report 2022–5009, 217 p. https://doi.org/10.3133/sir20225009.

USDA (United States Department of Agriculture). 1986. Urban Hydrology for Small Watersheds, Technical Release 55 (TR-55).

USDA (United States Department of Agriculture). 2024a. Digital General Soil Map of the United States (SSURGO). Accessed April 2025 from https://nrcs.app.box.com/v/soils.

USDA (United States Department of Agriculture). 2024b. Digital General Soil Map of the United States (STATSGO2). Accessed April 2025 from https://www.nrcs.usda.gov/resources/data-and-reports/description-of-statsgo2-database.

USDA (United States Department of Agriculture). 2024c. Cropland Data Layer. Accessed April 2025 from https://croplandcros.scinet.usda.gov/.

USGS (United States Geological Survey). 2025. Sustainable Groundwater: Groundwater Modeling. Accessed May 2025 from https://ca.water.usgs.gov/sustainable-groundwater-management/california-groundwater-modeling.html.

Xia, Y., Hobbins, M. T., Mu, Q., and Ek, M. B. 2015. Evaluation of NLDAS-2 Evapotranspiration against Tower Flux Site Observations. Hydrological Processes 29(7): 1757–71. doi:10.1002/HYP.10299.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., et al. 2012. Continental-Scale Water and Energy Flux Analysis and Validation for the North American Land Data Assimilation System Project Phase 2 (NLDAS-2): 1. Intercomparison and Application of Model Products. Journal of Geophysical Research Atmospheres 117(3): 3109. doi:10.1029/2011JD016048.

Zi, T., Braud, A., McKee, L., and Foley, M. 2022. San Francisco Bay Watershed Dynamic Model (WDM) Progress Report, Phase 2. Report Prepared for the Sources Pathways and Loadings Workgroup of the Regional Monitoring Program for Water Quality. SFEI Contribution #1091. Richmond, CA. Accessed May 2023 from https://www.sfei.org/documents/san-francisco-bay-watershed-dynamic-model-wdm-progress-report-phase-2.

Zi, T., Mckee, L, Yee, D., and Foley, M. 2021. San Francisco Bay Regional Watershed Modeling Progress Report, Phase 1. Report Prepared for the Sources Pathways and Loadings Workgroup of the Regional Monitoring Program for Water Quality. SFEI Contribution #1038. Richmond, CA. Accessed May 2023 from https://www.sfei.org/documents/san-francisco-bay-regional-watershed-modeling-progress-report-phase-1.