Initial Results of Statewide Risk Assessment for California Wastewater Systems

ADVISORY GROUP SUMMARY – OCTOBER 2026 ADIVISORY GROUP MEETING

ADVISORY GROUP SUMMARY

Quantitative Risk Assessment

The California Wastewater Risk Assessment is a data-driven prioritization tool developed to support proactive oversight, outreach, and support of wastewater facilities and systems throughout the state. Three separate Risk Assessments are conducted for 1) National Pollutant Discharge Elimination System (NPDES) facilities, 2) Waste Discharge Requirements (WDRs) facilities, and 3) Sanitary Sewer System General Order (SSSGO) systems due to their unique operations, regulations, and data availability. These assessments are separate, but complementary, to the Inadequacy Assessment. The Risk Assessments assign each facility a composite risk score based on variables across three risk categories: operational, sociodemographic, and environmental.

The Risk Assessment is intended to flag facilities that may be struggling or at elevated risk of future compliance failure. It is designed as a forward-looking tool and is not a determination of current compliance. Additionally, the Risk Assessment includes 550 WDR facilities that could not be assessed for Inadequacy due to missing violation and enforcement data. For these WDR facilities, their risk score represents the best available statewide assessment of facility operations, community vulnerability, and environmental hazard exposure.

These Risk Assessments contain data spanning from the 2020 Census[1] and the 2019-2023 American Community Survey[2], to forward-looking climate projections by Cal-Adapt[3] and CoSMoS[4] extending to 2060. Core compliance and operational data is limited to January 1, 2021 – January 1, 2026, consistent with the Inadequacy Assessment.

Risk Assessment Weights, Thresholds, and Scores

The total risk score for each facility or system is calculated as a weighted average of three sub-index scores. Within each sub-index, variables are weighted differently based on signal strength as established through statistical modeling, and input from the WWNA project team, Advisory Group members, and Water Board staff.

---

[1] U.S. Census Bureau. 2020 Decennial Census of Population and Housing. U.S. Department of Commerce. https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-main.html
[2] U.S. Census Bureau. American Community Survey 5-Year Estimates, 2019–2023. U.S. Department of Commerce. https://www.census.gov/programs-surveys/acs
[3] California Energy Commission. Cal-Adapt Climate Data. University of California, Berkeley, Geospatial Innovation Facility. Accessed 2025–2026. https://cal-adapt.org
[4] Barnard, P.L., et al. Coastal Storm Modeling System (CoSMoS). U.S. Geological Survey. https://cmgds.marine.usgs.gov/data-releases/community/cosmos

Statistical Modeling for Variable Selection and Weighting

To identify which risk variables were most strongly associated with Inadequacy, we separately applied logistic regression with the LASSO and Elastic Net methods for SSSGO, NPDES, and WDR facilities. Modeling results informed variable selection and weighting by identifying variables with consistent positive associations with Inadequacy and conversely, understanding which variables showed inconsistent or negative associations. Statistical models did not determine final variable weights and thresholds alone. Final weighting decisions also incorporated conceptual and policy considerations, including cross-program consistency, data coverage, and a nuanced understanding that some risk variables will show a weak correlation with Inadequacy, such as missing regulatory data, but should be included in the Risk Assessment.

Variable Risk Thresholds

Each variable in the Risk Assessment is assigned 0.0, lower risk, 0.5, moderate risk, or 1.0, higher risk, based on where the facility falls relative to defined thresholds.

For continuous variables, two threshold schemes are used:

A. Scheme A: Thresholds set at the 50th and 75th percentiles of the dataset. These are variables where risk is spread more evenly across the variable.
B. Scheme B: Thresholds set at the 75th and 90th percentiles of the dataset. These are variables where risk concentrates in the upper tail.

A small number of variables use fixed thresholds rather than percentiles. For example, population change uses defined percentage bands (+/-6% and +/-9%), while permit age uses fixed year ranges such as under 20 years, 20-30 years, and over 30 years. Binary and categorical variables, such as missing SSMP (Sewer System Management Plan per SSSGO requirement) or agency governance type, are assigned scores based on their association with Inadequacy and on conceptual risk logic explored by the WWNA project team, Advisory Group members, and Water Boards staff.

Risk Tier Classification

Each facility receives a final composite risk score between 0.0 and 1.0. Fixed score cut points are used to categorize facilities and systems into Higher, Medium, and Lower Risk categories. The cut points differ across the NPDS, WDR, and SSSGO assessments because each assessment uses slightly different risk variables and thresholds, resulting in unique risk score distributions.

The risk tier cut points were set through a combination of statistical analysis, validation with the Inadequacy Assessment, and Water Boards policy judgment. The final cut points

were chosen to flag 10-14% of facilities or systems as Higher Risk across program types. The tiers are simply a prioritization tool rather than a definitive judgment about facility or system condition.

Overall Preliminary Results

Program	Total Facilities[5]	At-Risk	Potentially At-Risk	Not At-Risk
NPDES	265	35 (13%)	35 (13%)	195 (74%)
WDR	971	132 (14%)	158 (16%)	681 (70%)
SSSGO	946	93 (10%)	120 (13%)	733 (78%)
Total	**2,182**	**260 (12%)**	**313 (14%)**	**1,609 (74%)**

Known Limitations

The validation of the Risk Assessments relies on the current Inadequacy status as a performance measure. This is an imperfect proxy, as some facilities may be struggling despite not being identified as Inadequate[6], while others may have been classified based on limited or inconsistent data. However, when validated with facilities and systems with the top 10% of risk scores, almost 100% of all High Risk facilities and systems are in the top 10% of overall risk scores; only a handful are slightly below this percentile. This validates that the current cut points are identifying the same High Risk facilities and systems regardless of whether you use the cut points established by the Inadequacy rates or if you use a purely percentile-based approach.

A substantial number of WDR facilities, 550 of 971, lack any violation and enforcement records available through CIWQS at the time of Inadequacy Assessment completion (June

[5] Total number of facilities analyzed within the WWNA project scope. Corresponds to the same number and list of facilities analyzed for Inadequacy. For WDR facilities, 971 were included in the WWNA project scope, but only 471 had sufficient violation and enforcement data to be fully assessed for Adequacy/Inadequacy. All 971 WDR facilities can be assessed in the Risk Assessment.

[6] Inadequate Definition (per WWNA scope): Wastewater system or facility does not adequately treat and/or dispose of wastewater, leading to environmental, and/or public health degradation.

2026). These facilities could not be assessed in the Inadequacy Assessment, but they do receive risk scores based on other available data sources.

Appendix

NPDES – Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Operational (50% overall weight)	Flow Capacity Stress	25%	<p75	p75-p90	>p90
	Agency Type	20%	Not Small Special District, Privately-Owned	Not small Muni/City, Small Special District, Small County/State/Fed/Govt Combination	Small Muni/City, Small Mobile Home Park
	Vacant Supervisor	20%	0 (filled)		1 (vacant)
	5-year exceedance slope nearing violation threshold	10%	0 (not flagged)		1 (flagged)
	Facility discharging to an impaired waterbody with no current limit	10%	0 (not flagged)		1 (flagged)
	3-year population change	10%	< ±6%	±6–9%	> ±9%; > 200% capped
	Agricultural Land Use	5%	0 (not flagged)		1 (flagged)
Socio-demographic (35% overall weight)	Percent Poverty	50%	<25.1% (p<50)	25.1 – 34.2% (p50-p75)	>34.2% (>p75)
	DAC Status	33%	0.0 (Non-DAC)	0.5 (DAC)	1.0 (SDAC)
	Hispanic/Latino % of residents served by facility	17%	<26.7% (p50)	26.7 – 42.1% (p50 – p75)	>42.1% (p75)

NPDES – Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Environmental (15% overall weight)	Hectares projected to burn annually by wildfire (2030 – 2060)	20%	<20.2 (p75)	20.0 – 34.1 (p75 – p90)	>34.1 (p90)
	Average number of days with extreme precipitation (2030 – 2060)	20%	<9.3 (p75)	9.3 – 17.8 (p75 – p90)	>17.8 (p90)
	Composite Seismic Score	20%	<97.7 (p75)	97.7 – 112.0 (p75 – p90)	>112.0 (p90)
	Number of days per year exceeding the extreme heat threshold and daily maximum temperature during a synthetic mega-drought scenario (2030 – 2060)	20%	Both <p75	Average of component scores p75-p90	Either component >p90
	Hours of inundation and centimeters of inundation under a climate scenario of 1 meter of sea level rise and 100-year storm	20%	0 (no exposure)	Any exposure	Either component >p90

WDR – Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Operational (50% overall weight)	Agency Type	20%	Not Small Special District, Privately-Owned	Not small Muni/City, Small Special District, Small County/State/Fed	Small Muni/City, Small Mobile Home Park
	Vacant Supervisor	15%	0 (filled)		1 (vacant)
	No Regulatory Data	20%	0 (not flagged)		1 (flagged)
	Effluent Exceedance Risk	20%	0 (not flagged)		1 (flagged)
	Agricultural Land Use	15%	0 (not flagged)		1 (flagged)
	3-year population change	5%	< ±6%	±6–9%	> ±9%; > 200% capped
	Permit Age	5%	<20 years old	20 -30 years old	>30 years old
Socio-demographic (35% overall weight)	Percent Poverty	50%	<28.1% (p<50)	28.1 – 38.8% (p50-p75)	>38.8% (>p75)
	Hispanic/Latino % of residents served by facility	33%	<23.5% (p50)	23.5 – 42.4% (p50 – p75)	>42.4% (p75)
	DAC Status	17%	0.0 (Non-DAC)	0.5 (DAC)	1.0 (SDAC)

WDR – Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Environmental (15% overall weight)	Hectares projected to burn annually by wildfire (2030 – 2060)	20%	<33.1 (p75)	33.1 – 42.7 (p75 – p90)	>42.7 (p90)
	Average number of days with extreme precipitation (2030 – 2060)	20%	<11.75 (p75)	11.75 – 19.2 (p75 – p90)	>19.2 (p90)
	Composite Seismic Score	20%	<79.4 (p75)	79.4 – 99.9 (p75 – p90)	>99.9 (p90)
	Number of days per year exceeding the extreme heat threshold and daily maximum temperature during a synthetic mega-drought scenario (2030 – 2060)	20%	Both <p75	Average of component scores p75-p90	Either component >p90
	Hours of inundation and centimeters of inundation under a climate scenario of 1 meter of sea level rise and 100-year storm	20%	0 (no exposure)		Any exposure

SSSGO - Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Operational (50% overall weight)	SSMP Status	30%	Current SSMP		No SSMP on record
	Agency Type	20%	Not Small Special District, Privately-Owned	Not small Muni/City, Small Special District, Small County/State/Fed/Govt Combination	Small Muni/City, Small Mobile Home Park
	Cumulative System Age pre-1940	10%	0% (none)	>0	>50%
	Cumulative System Age pre-1960	10%	<25%	25-50%	>50%
	Pct Spills Attempted Recovery	10%	>66.7% (p50)	42.5-66.7% (p25-p50)	<42.5% (p25)
	3-year population change	5%	< ±6%	±6–9%	> ±9%; > 200% capped
	Pct System Cleaned	5%	>32.8% (p50)	14.4-32.8% (p25-p50)	<14.4% (p25)
	Pct System Inspected	5%	>7.5% (p50)	1.4-7.5% (p25-p50)	<1.4% (p25)
	Agricultural Land Use	5%	0 (not flagged)		1 (flagged)
Socio-demographic (35% overall weight)	Percent Poverty	50%	<25.5% (p<50)	25.1 – 36.6% (p50-p75)	>36.6% (>p75)
	DAC Status	33%	0.0 (Non-DAC)	0.5 (DAC)	1.0 (SDAC)
	Hispanic/Latino % of residents served by facility	17%	<27.9% (p50)	27.9 – 49.5% (p50 – p75)	>49.5% (p75)

SSSGO – Risk Variable Weights & Thresholds					
Sub-Index	**Variable**	**Weight**	**Low Score (0.0)**	**Moderate Score (0.5)**	**High Score (1.0)**
Environmental (15% overall weight)	Hectares projected to burn annually by wildfire (2030 – 2060)	20%	<19.7 (p75)	19.7 – 36.1 (p75 – p90)	>36.1 (p90)
	Average number of days with extreme precipitation (2030 – 2060)	20%	<8.4 (p75)	8.4 – 15.3 (p75 – p90)	>15.3 (p90)
	Composite Seismic Score	20%	<89.8 (p75)	89.8 – 105.5 (p75 – p90)	>105.5 (p90)
	Number of days per year exceeding the extreme heat threshold and daily maximum temperature during a synthetic mega-drought scenario (2030 – 2060)	20%	Both <p75	Either p75 – p90	Either >p90
	Hours of inundation and centimeters of inundation under a climate scenario of 1 meter of sea level rise and 100-year storm	20%	0 (no exposure)		Any non-zero