# Determining the Drivers of Public Participation in Environmental Governance Using AI

Jelena Jezdimirovic, Tyler Scott, Alvar Escriva-Bou, and Katrina Jessoe *

**Abstract**

Public comments comprise an essential and mandatory step in the rule making process for many environmental regulations. However, little is known about how notice-and-comments affect regulatory stringency and policy outcomes. Using large language models, we empirically investigate the features that drive participation in the environmental regulatory process, and the effects of participation on regulatory stringency. Our context is California's 2014 Sustainable Groundwater Management Act (SGMA), a state regulation that required groundwater basins throughout the state to create local governance agencies and develop sustainability plans to regulate groundwater use. We analyze more than 18,000 public comments on draft versions of local plans, classifying each comment by topic and sentiment. We then construct two sustainability metrics, groundwater overdraft and regulatory stringency, and evaluate their correlation with the subsequent volume and content of public comments. A first central result is that participation increases when stakeholders perceive negative and costly regulatory impacts, but is unaffected by current conditions. For agricultural stakeholders, a group reliant on groundwater for irrigation, the submission of negative comments increases with regulatory stringency. Environmental NGOs, a constituency that favors groundwater

*The authors gratefully acknowledge financial support from the National Science Foundation (NSF) Belmont Forum and the United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA). The authors have no additional financial relationships or conflicts of interest to disclose.

in-situ, exhibit the reverse pattern. However, the distribution of commenters does not systematically differ with current overdraft conditions. Despite extensive stakeholder participation in public comment process, a second central result is that participation does not appear to influence regulatory outcomes. Regulatory stringency changed in only a third of the plans between draft and final versions. Sentiment and participation were not systematically associated with whether the regulatory stringency changed, or with magnitude of the change. These findings suggest that stakeholder mobilization is driven by future expectations about resource scarcity as opposed to current conditions, but does not impact regulatory outcomes.

# 1 Introduction

Common-pool natural resources ranging from fisheries to forests to climate often suffer from over-extraction and degradation because of the absence of well defined property rights (Hotelling, 1931; Gordon, 1954; Stavins, 2011; Hartman, 2018). Nowhere is the mismanagement of common-pool resources more on display than in the context of groundwater. Groundwater is a a major source of global water supplies, accounting for 40% of irrigated agriculture (Siebert et al., 2010), and supplying over 50% of drinking water (UNESCO World Water Assessment Programme, 2022). However, the long-term availability of these supplies is in jeopardy, with recent estimates indicating that rapid overextraction is occurring in over a third of the world's largest aquifers (Richey et al., 2015; Jasechko et al., 2024). This decline is exacerbated by climate change, which is expected to increase water shortages in certain parts of the world (Carleton, Crews, and Nath, 2024).

The need to internalize external costs motivates the rationale for institutions and regulations to govern groundwater resources (Coase, 2013; Pigou, 2017). This has led to the use of technology standards, extraction limits, and markets to manage them (Bruno and Jessoe, 2021, 2024). Many of these policies are developed by administrative government agencies, which rely on rulemaking processes to arrive at specific regulatory parameters, requirements, and standards (Yackee, 2019; Smith and Licari, 2006). A key aspect of rulemaking includes notice and public comment procedures, which are typically required by state or federal law (Rosenblum-Sellers, 2025). These procedures involve publishing a draft of the regulation, soliciting public feedback, and then responding to comments deemed substantive in developing the final regulation (Yackee, 2019). This process is time consuming and leads to regulatory delays, but it also provides the public and affected parties an opportunity to weigh in on and influence regulations (Kerwin and Furlong, 2018).

Due to the prominence of notice and comment rulemaking in the United States, a substantial literature is devoted to understanding what drives stakeholder engagement and how this shapes regulatory outcomes. However, several obstacles limit the ability to empirically

quantify these relationships. The first challenge is that the existing literature frequently pools notice-and-comment processes from different policy domains, rule types, geographic contexts, and time frames (Crow, Albright, and Koebele, 2016, 2020; Yackee and Yackee, 2006). While this practice is frequently necessary to achieve sufficient statistical power for analysis, inherent differences in settings makes it challenging to systematically compare across cases. Second, studies that focus on how comments shape regulatory outcomes frequently measure difference between draft and final rule text, or whether an agency responded to comment (Haeder and Yackee, 2015; Ulibarri, Scott, and Perez-Figueroa, 2019; Costa, Desmarais, and Hird, 2019; Naughton et al., 2009). These studies do not systematically measure changes to actual regulatory stringency, which determines how comments shape regulatory outcomes. Finally, even with an ideal research design, until recently data processing limitations constrained the information that could be extracted on participation habits. Costa, Desmarais, and Hird (2019); Scott, Ulibarri, and Scott (2018, 2020); Hui, Ulibarri, and Cain (2018).

We overcome these limitations by using recently developed large language models to study stakeholder participation and the notice and comment process in California's 2014 Sustainable Groundwater Management Act (SGMA), a state-level policy that required local decentralized governance of close to 100 common groundwater basins at a single moment in time. This state law required local water basins to define their own management targets and plans, and currently, 117 groundwater plans have been developed. Our analysis benefits from SGMA's unique context: multiple agencies working simultaneously within a common regulatory framework to develop individual plans for managing groundwater. This creates significant cross-sectional variation within a common regulatory framework.

SGMA also mandated stakeholder involvement procedures, including notice-and-comment periods. A crucial advantage of our setting is the ability to observe plans before and after public comment, allowing us to directly measure if public participation affected regulatory stringency in the adopted plans. In total, more than 18,000 public comments comprising roughly 1.8 million words were submitted on draft plans. The sheer scale of the text makes

manual coding infeasible. Traditional natural language processing tools could automate analysis, but they have limitations: some methods require pre-labeled training data, others produce unlabeled topic clusters that require interpretation, and most struggle with nuanced tone and context for sentiment analysis. Instead, we use a large language model to turn the public comment text corpus into a structured database cataloging stakeholder concerns and positions.

In this paper we construct a new groundwater database on stakeholder comments and identities, groundwater overdraft and regulatory stringency and apply it to study what drives stakeholder participation in the regulatory process and how participation affects regulatory outcomes. We leverage a large language model Claude Sonnet 4 to extract topic and sentiment from the public comments submitted on the groundwater sustainability plans. This allows us to assess stakeholder perception of the proposed regulation in draft plans. Since there is substantial flexibility in defining overdraft and pathway to sustainability, we construct two consistent sustainability metrics: excess historical pumping in each basin, and the future regulatory stringency, which measures how much additional groundwater decline each plan permits. We collect data on future regulatory stringency in draft and adopted groundwater sustainability plans and calculate the difference between the two. We use these data to analyze how stakeholder response varies with regulation stringency and existing overdraft conditions. Finally, we test whether public comment is associated with meaningful changes in regulatory stringency between draft and adopted plans.

We find that participation in groundwater management increases when stakeholders anticipate being adversely affected by regulation. Agricultural stakeholders submit a higher share of more negative comments as stringency increases, while the environmental NGOs submit fewer, less negative comments with increasing stringency. In contrast, participation does not vary systematically with the scale of the overdraft problem in a basin. Taken together, these results point to stakeholders responding to the regulatory stringency set forth in the plans, and acting strategically to further their interests.

While there is active stakeholder participation in the process, we observe limited changes in regulatory stringency between draft and final plans. Roughly two thirds of the plans permit the same groundwater decline in draft and final versions. Where regulatory stringency differs, the change is modest: on average the plans allow 1.4 feet of additional groundwater level decline in the final plans. Further, we find no difference in stakeholder participation and sentiment between plans with and without changes. These results indicates that stakeholder participation in the notice and public comment process had a limited effect on the regulatory outcome.

# 2 The Sustainable Groundwater Management Act

California's SGMA was developed in response to a historic drought that led to thousands of wells running dry, widespread land subsidence due to groundwater overdraft, and significant declines in fish populations (Legislative Analyst's Office, 2021; Lund et al., 2018). A total of 94 groundwater basins, accounting for more than 89% of statewide groundwater use, fall under its purview (California Department of Water Resources, 2019). Instead of a top-down approach, SGMA delegated authority to local water users to develop individual plans for sustainable management. The decentralized approach resulted in substantial variation in local planning, stakeholder engagement, and plan stringency.

As a starting point, agencies and users overlying groundwater basins were required to create groundwater sustainability agencies (GSAs) and assume local regulatory responsibility. To date, 232 unique GSAs have formed. Panels A and B of Figure 2.1 provide a map of the basins regulated under SGMA and the concentration of GSAs in the state.[1]

Under SGMA, the California Department of Water Resources (CDWR) required each regulated basin to submit a groundwater sustainability plan (GSP). The law defines sus-

[1]While other states have implemented groundwater policies, they were either centralized, or involved a limited number of management entities created over time (Escriva-Bou et al., 2016). In contrast, SGMA covered dozens of basins and created hundreds of new agencies in a short timespan, providing a large set of comparative cases for analyzing the drivers of public participation.

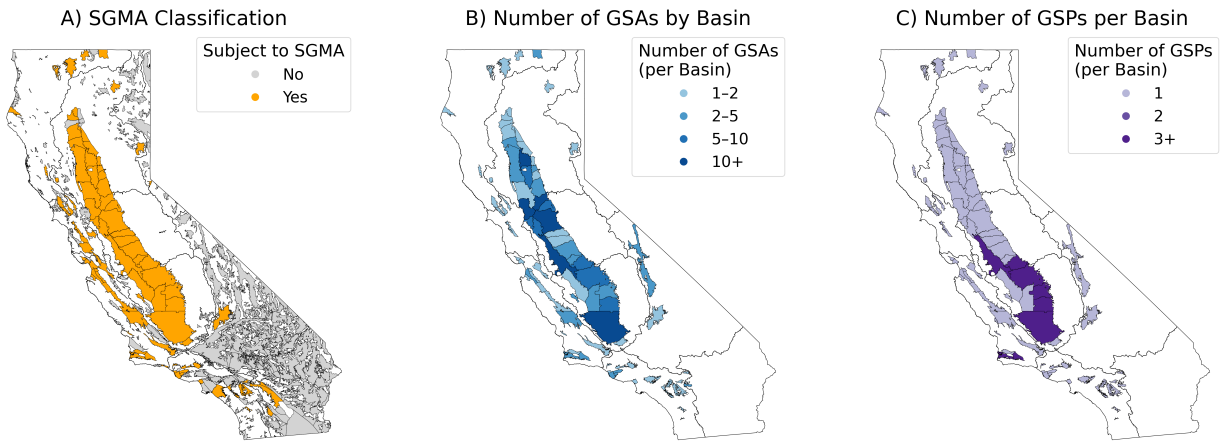



Figure 2.1: Groundwater basins subject to SGMA; number of GSAs and GSPs per basin

tainable management as the avoidance of six undesirable results: (1) chronic lowering of groundwater levels, (2) reduction of groundwater storage, (3) seawater intrusion, (4) land subsidence, (5) degradation of water quality and (6) depletion of interconnected surface water. Plans must establish quantifiable thresholds for each of the undesirable results, and agencies must monitor to ensure that these thresholds are not exceeded. Additionally, plans must describe current and historical groundwater conditions, develop a comprehensive water budget tracking inflows and outflows of water, quantify sustainable pumping levels, and detail the projects and management actions needed to achieve groundwater sustainability. While plans differ substantially in how they choose to address each requirement, this standard framework enables systematic comparison.

SGMA required a public notice-and-comment procedure to support stakeholder participation. Between 2019 and 2022, GSAs were required to publish draft GSPs, public comments, and consider comments received prior to submitting final plans for state review (and include all comment letters as a separate section in the plan). This process allowed for public input and participation, and created a detailed record of public input.

# 3 Evidence from Sustainability Plans

Our final sample comprises 80 basins for which we have both complete regulatory stringency data (minimum thresholds from draft and final plans) and public comment records. Single plans were submitted in 74 basins and multiple plans submitted in 6 basins in our sample. Typically, multiple GSAs collaborate on a single plan. These plans include information, summarized in Table 1, on the number of public comments and sustainability metrics.

Descriptive statistics reveal substantial heterogeneity in water use and plan management 1. Historic pumping ranges from zero to over 1.5 million-acre feet annually, and groundwater dependence varies from 2% to 100%. Basins report that sustainable yield pumping, or pumping quantities that would allow them to meet sustainability targets, are on average 140 TAF. These are lower than historic pumping quantities of 145 TAF. To meet sustainable yield targets, average pumping would need to reduce by 8%, though in some basins reductions would be as high as 79%. Differences in pumping reductions across basins underscores substantial variation in groundwater sustainability challenges across basins.

In total, 18,061 public comments comprising over 1.8 million words were submitted on the the draft groundwater sustainability plans. There is a substantial variation in stakeholder engagement across basins: while some basins received a handful of comments, others attracted more than 1,000.

# 4 Measuring Overdraft and Regulatory Stringency

Groundwater sustainability agencies have substantial flexibility in defining the six undesirable results listed under SGMA. This poses a measurement challenge, since basins differ in their definition of undesirable results. To address this issue, we construct two metrics, defined in Table 2, for each groundwater basin. These metrics allow for a comparison in overdraft and regulatory stringency across basins.

Our first metric, labeled 'pumping reduction,' is defined as the percent reduction

Table 1: Summary Statistics by Basin

| **Measure** | **Obs** | **Mean (SD)** | **Min** | **Max** |
|---|---|---|---|---|
| GSAs by Basin | 80 | 3.21 | 1 | 23 |
| | | (4.13) | | |
| GSPs by Basin | 80 | 1.29 | 1 | 7 |
| | | (1.10) | | |
| Groundwater % of Total Water Use | 80 | 0.65 | 0.02 | 1.00 |
| | | (0.28) | | |
| Overdraft by Basin (1000 acre-feet) | 80 | 18.28 | 0 | 369.89 |
| | | (51.08) | | |
| Historical Pumping (1000 acre-feet) | 80 | 145.08 | 0 | 1,590.37 |
| | | (269.51) | | |
| Sustainable Yield Pumping (1000 acre-feet) | 80 | 140.42 | 0 | 1,313.00 |
| | | (234.65) | | |
| Pumping Reduction | 79 | 0.08 | 0 | 0.79 |
| | | (0.14) | | |
| Future Permitted Groundwater Decline in Draft Plans (median, feet) | 80 | −33.91 | −236.70 | 38.77 |
| | | (41.52) | | |
| Future Permitted Groundwater Decline in Final Plans (median, feet) | 80 | −35.35 | −236.70 | 38.77 |
| | | (44.66) | | |
| Number of Public Comments by Basin | 80 | 207.14 | 17 | 868 |
| | | (221.57) | | |

in historical groundwater pumping required to achieve sustainable yield. SGMA defines sustainable yield as the maximum quantity of groundwater that can be pumped annually without causing undesirable results (California State Legislature, 2014). The law does not require agencies to eliminate historical overdraft, or to return the basin to pre-development conditions (Smith, 2021). Since our metric quantifies the gap between historical sustainable yield and pumping rather than enforceable cuts, we use it as an indicator of the size of the problem faced in each basin.[2]

| **Metric** | **Formula** |
|---|---|
| Pumping Reduction (PR) | $1 - \left(\frac{\text{Sustainable Yield}}{\text{Average Historical Pumping}}\right)$ |
| Future Permitted Groundwater Decline | median (2015 Depth to GW − MT Depth to GW) |

Table 2: Sustainability Metrics Formulas

Our second metric focuses on the stringency of thresholds that agencies set to monitor compliance with SGMA. In the "future permitted groundwater decline" we calculate the difference between 2015 water table readings and the minimum groundwater depth thresholds defined in the plans. We focus on thresholds because they are the benchmarks GSAs must use to track progress toward sustainability. Minimum thresholds are groundwater elevations, set at the well level, that if exceeded will likely lead to undesirable results (Department of Water Resources, 2017). To compare minimum thresholds across basins, we choose 2015 water depth as a baseline measure and subtract thresholds from it. A negative difference indicates agencies allow water tables to decline below 2015 levels while still meeting sustainability goals, while a positive difference shows levels are not allowed to fall below 2015 conditions. We take the median of the individual well groundwater depth differences to create a basin-level metric. This metric follows the approach in Bostic et al. (2023), which compared minimum thresholds to 2019 water levels. To measure how public comment affected regulatory stringency, we calculate this metric for both draft and final plans.

[2]While we frame it as 'pumping reduction,' agencies can address the gap through demand reduction (pumping cuts) and/or supply expansion (enhanced recharge, imported water).

We use the two sustainability metrics to examine whether stakeholders react more to the scale of the overdraft challenge (pumping reduction) or to how agencies define thresholds to manage it (future permitted groundwater decline).

# 5 Public Sentiment and GSPs

Using AI we construct a database comprised of the topic and sentiment of each comment, and classify commenter identities into sectors (e.g. agriculture). As shown in Figure 5.1 the most common comment categories are line edits and on the water budget/groundwater model, followed by supply enhancement projects, monitoring network and surface water-groundwater interaction . Together, these top five categories account for 8,750 comments or nearly 48% of comments. The prominence of line edits, where a commenter requests specific editorial changes, indicates a deep engagement and close reading of the groundwater sustainability plans. Strong representation of water budget and groundwater model concerns reflects a scrutiny of the science underpinning management decisions.

A mix of individuals, agencies, and NGOs submitted comments on the groundwater sustainability plans. As shown in Figure 5.2, while non-affiliated individuals (35%), water districts (19.9%) and agriculture (19.5%) comprise the three sectors with the greatest number of commenters, environmental NGOs are responsible for submitting most comments (27.5 % of the total), with groups often submitting multiple comments. Figure 5.3 reveals that there is substantial heterogeneity across groups in dominant comment themes. Comments submitted by environmental groups focus on environmental health, specifically ecosystems and surface water-groundwater interactions. In contrast, the focus of agricultural entities is on actions intended to reduce extraction (i.e. demand management actions) or augment supplies (i.e. supply enhancement actions). Water utilities and government agencies most often submit clarifying questions, followed by concerns about water budgets and groundwater models.

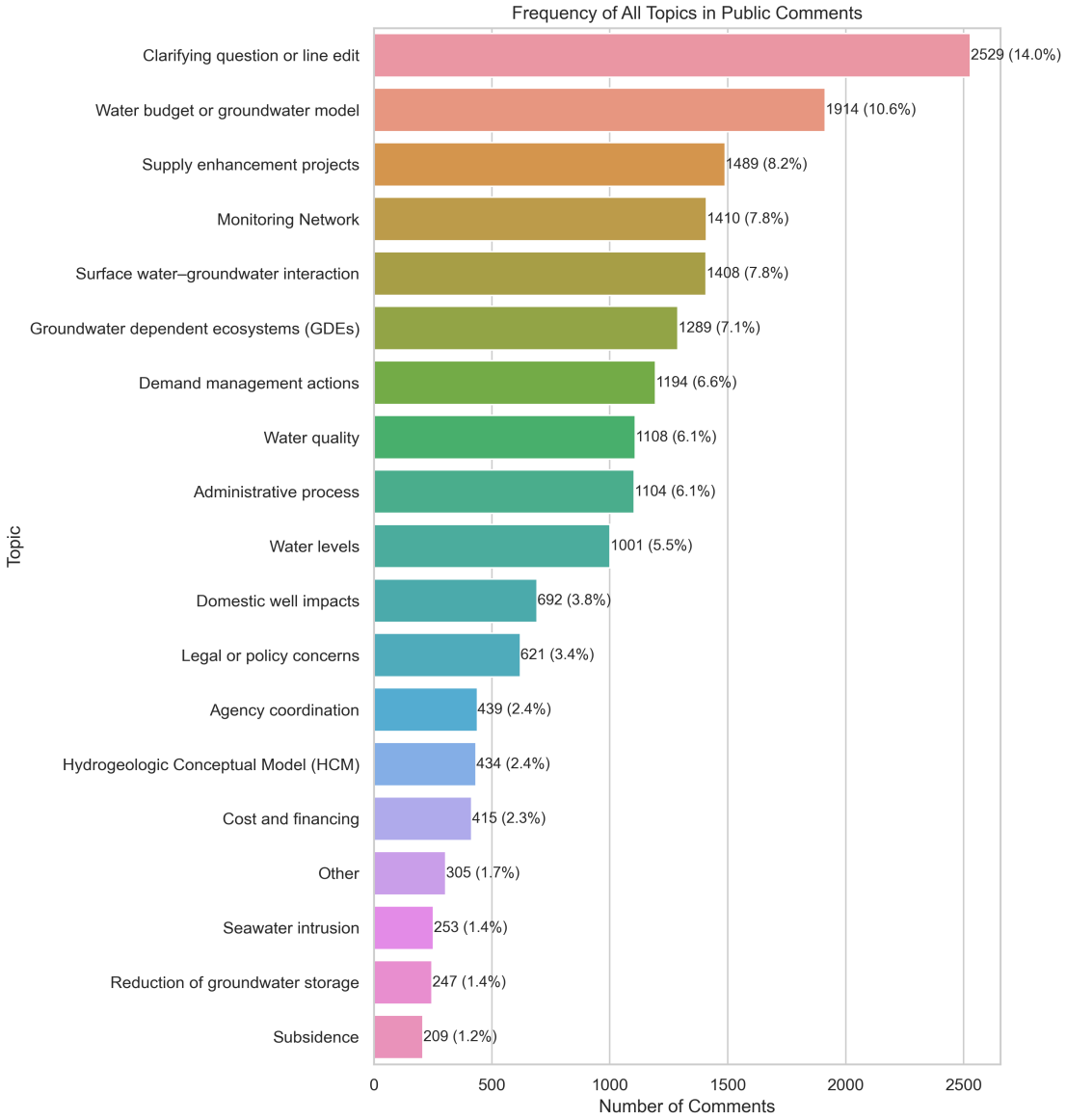



Figure 5.1: This figure shows the number of comments assigned to each of the 19 topics by Claude Sonnet 4. The data labels report both the number of comments and the percent share of total comments.

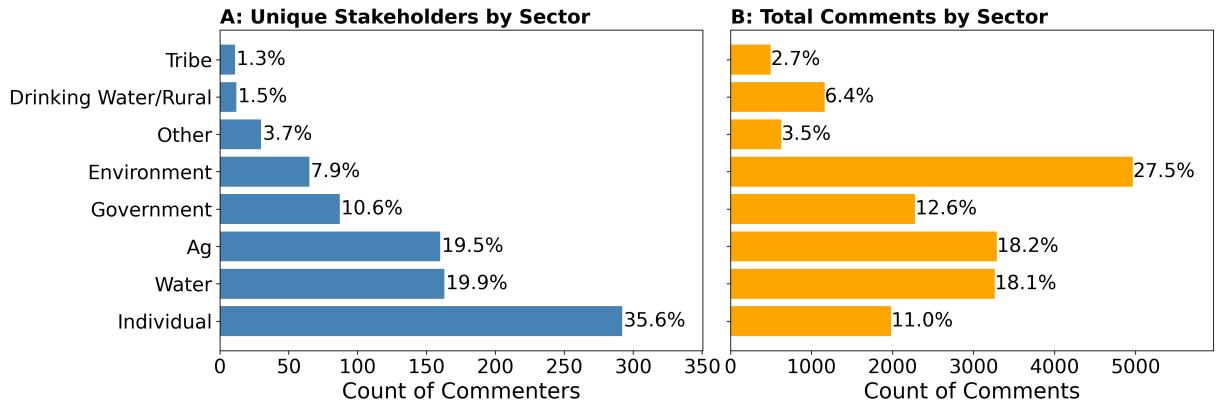



Figure 5.2: Panel A shows the number of unique commenters classified by sector. Panel B displays the total number of comments submitted by sector. Data lables show percentages of total. While non-affiliated individuals represent the largest group of commenters - the environmental sector submitted most comments, followed by the water and ag entities.

Figure A.1 makes clear that commenter sentiment is overwhelmingly negative, but that sentiment differs across sectors and themes. Environmental groups and tribes exhibit the most negative sentiment in their comments, while agriculture and water agencies display the most positive sentiment **??**). The themes with the most negative sentiment are the water budget and groundwater model categories (Figure A.3. The budgets and models are used to determine the scale of the overdraft problem and provide rationale for projects and management actions. For this reason, constituents may have a strong response to these topics, reacting if they perceive that the plan is overstating or understating the problem.

Line edits are the most neutral category, and supply enhancement projects are the topic with most positive sentiment attached. The positive sentiment may occur because commenter view infrastructure investment and increasing water availability favorably. A heat map in Figure A.4 provides a more detailed decomposition of the average sentiment by sector and topic.

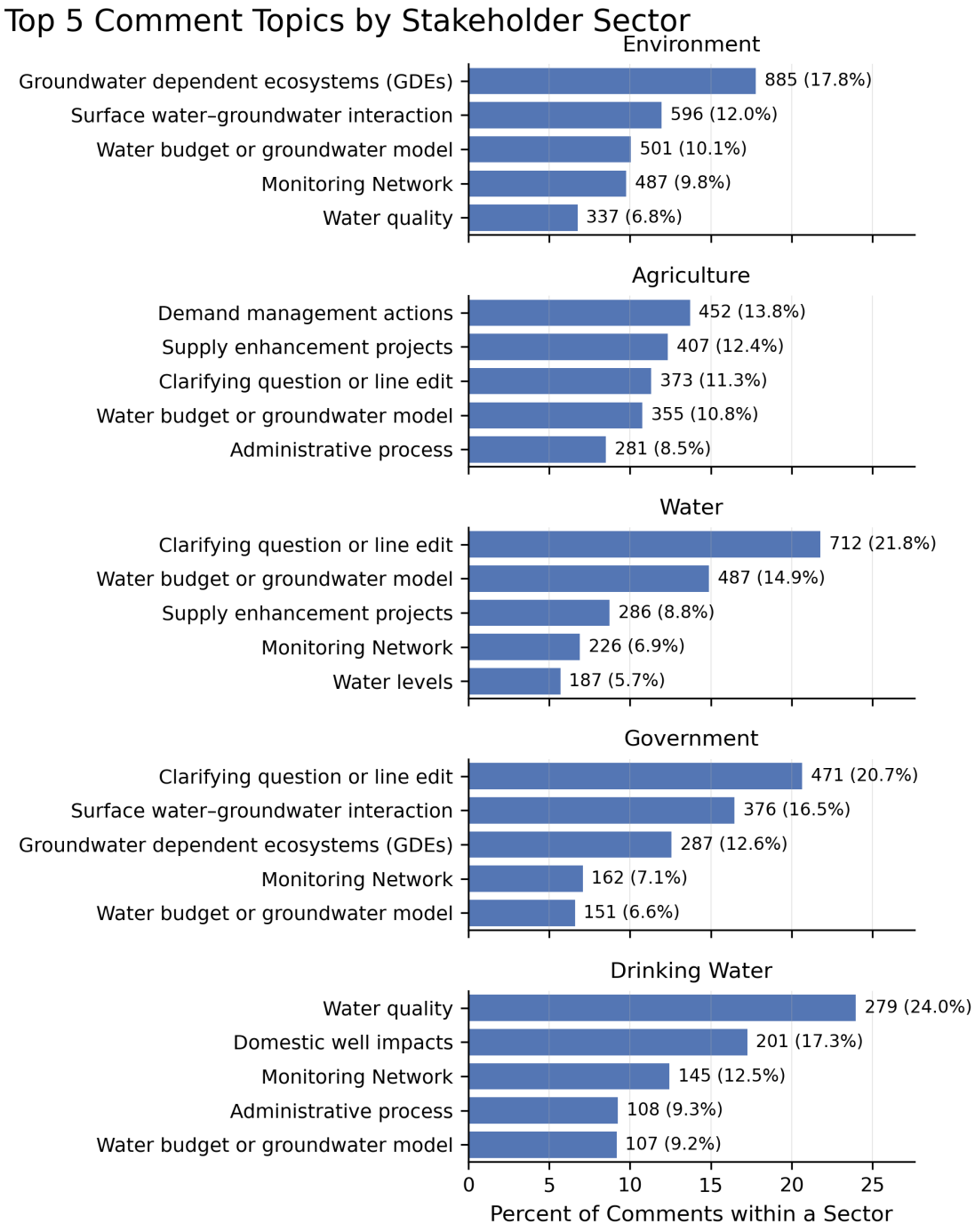



Figure 5.3: Figure displays top five topics mentioned in comments by sector, excluding individuals, tribes and other sector. The data labels show number of comments by topic along with the percent of total within a given sector. Top five topics represent 56.5% of comments made by the environment sector, 56.8% of comments by ag, 58.2% by water, 63.5% by government and 72.3% by drinking water NGOs.

# 6 Why Does the Public Participate in Groundwater Sustainability Plans?

Participation in groundwater management increases when stakeholders anticipate being adversely affected by regulation. In Figure 6.1, we present correlations between average sentiment and the future permitted groundwater decline in **draft plans**, disaggregated by sector. As plans become more stringent, agricultural sentiment towards plans decreases (r= -0.27, p=0.086), while environmental sentiment increases (r = 0.3, p= 0.008). An increase in stringency is also correlated with a higher share of comments submitted by agricultural entities (r=0.22, p=0.045), and a lower share submitted by environmental NGOs (r=-0.20, p=0.076) (Figure 6.2). This pattern suggests that stakeholders are reacting to the plan content, either attempting to steer regulation in a direction aligned with their interests, or laying administrative record groundwork for a future challenge/litigation. For agriculture, a more stringent future permitted groundwater decline metric may require costly reductions in pumping, prompting increased negative participation. In contrast, the environmental NGOs may view higher stringency as being more protective of the ecosystems, reducing their need to weigh in critically on those plans. This finding also aligns with predominant comment themes for each of the groups: supply augmentation and demand management for agriculture, and groundwater dependent ecosystems for the environmental organizations (Figure 5.3.

Stakeholders are not necessarily responding to the magnitude of historical overdraft. While there is a substantial stakeholder mobilization in response to future management, we find no correlations between participation and magnitude of overdraft. In Figure 6.3, we show that agriculture and environment sentiment and participation do not vary systematically with reduction percentage: the amount of groundwater cutback needed in the basin to reach sustainability.

We find that more severe reduction percentages lead to increased scrutiny of the

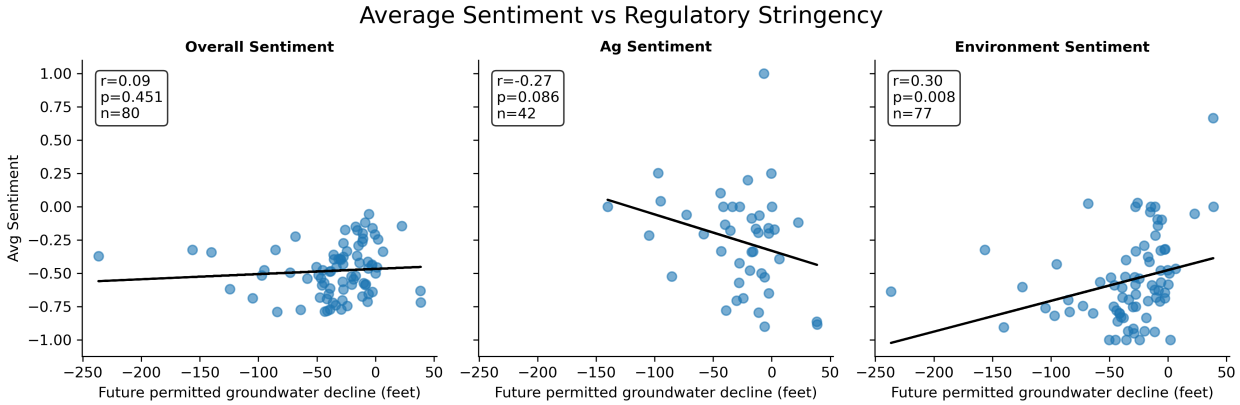



Figure 6.1: Correlation between average sentiment and future permitted groundwater decline. First panel shows overall correlation, while the other two panels show breakdown by sector (ag and environment). While the overall correlation is weak and not significant sector-specific trends diverge. Ag sentiment declines as thresholds become more stringent(r= -0.27, p=0.086), while environmental sentiment increases (r = 0.30, p= 0.008).

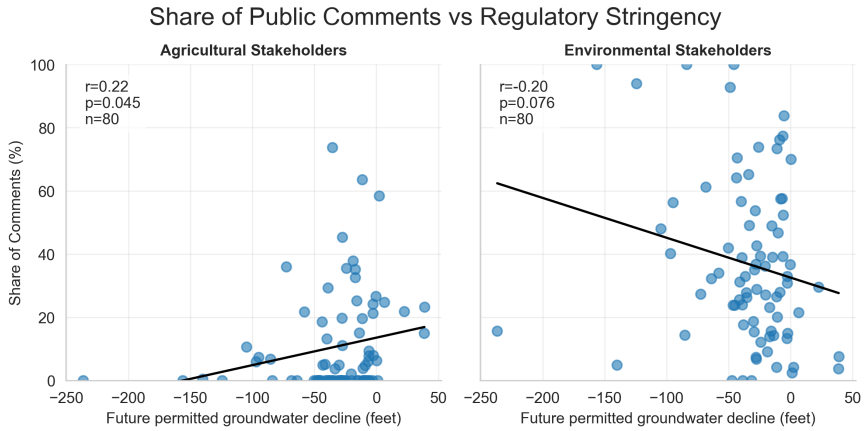



Figure 6.2: Correlation between share of comments and future permitted groundwater decline. Ag share of comments increases as thresholds become more stringent(r=0.22, p=0.045), while environmental share decreases (r=-0.20, p=0.076).

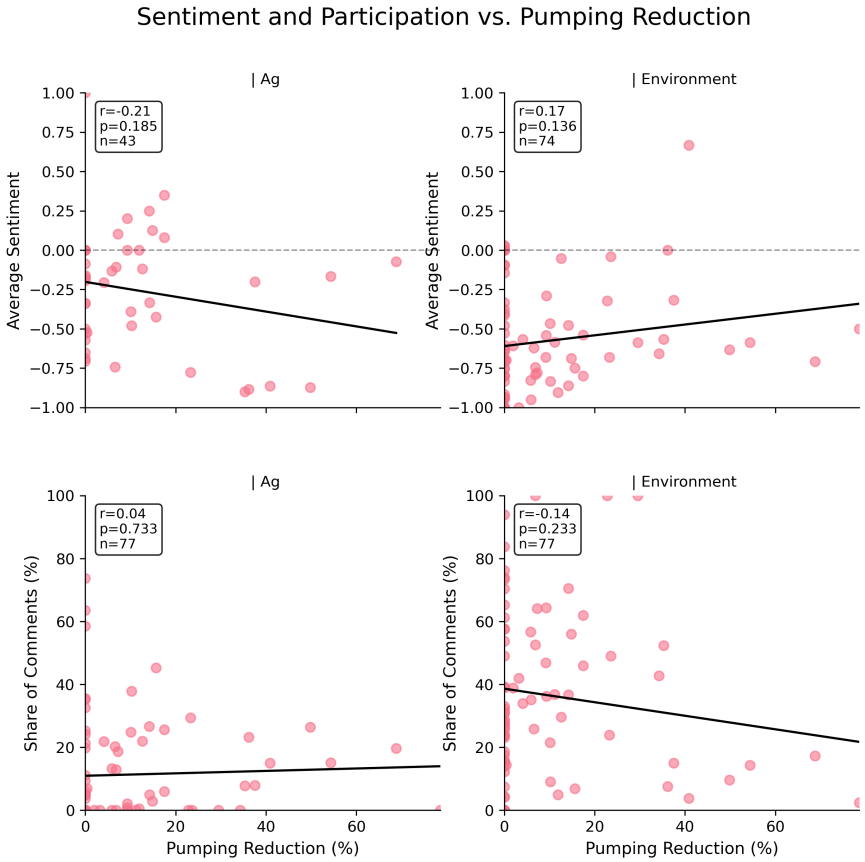



Figure 6.3: Correlation Between Average Sentiment, Share of Comments and Pumping Reduction. We find no evidence that either coalition is responding to the pumping reduction in sentiment or share of comments.

technical foundations of the plans - water budgets and groundwater models. In particular, the overall share of comments about water budgets and groundwater models increases as the reduction percentage rises (r=0.305, p=0.006), and these comments tend to be more negative (r=-0.213, p=0.060). As the scale of required reductions grows, commenters are scrutinizing the credibility of technical analysis underlying the plan.

# 7 Does Participation Affect Regulatory Outcomes?

The regulatory stringency of plans changed minimally between the draft and final plans. We obtained 100 (out of 117) plans for which draft versions were available. The difference in future groundwater decline between drafts and final plans is minimal: 59% of plans did not have any changes in regulatory stringency. At the monitoring well level, this lack of change is even more pronounced: groundwater level thresholds (used to define future groundwater decline) remained the same for 83% of the roughly 2,279 monitoring wells. For the remaining wells, thresholds were either changed (10.5%), wells were added to (2.4%), or removed from (3.6%) the monitoring network.

We find no evidence that stakeholder sentiment or participation influenced changes in regulatory stringency. In Figure (figure 7.1), we show that plans which adjusted their groundwater metrics from draft to final do not differ in stakeholder participation and sentiment. Further, we find no correlation between the magnitude of the change in regulatory stringency and comment sentiment (agricultural: r=0.19, p=0.20; environmental: r=0.07, p=0.50) or participation (ag: r=0.10, p=0.36; environment: r=0.06, p=0.58) (figure 7.2). Together, these results indicate that despite the substantial scale of public participation in the comment process, it is correlated with no change in regulatory stringency.

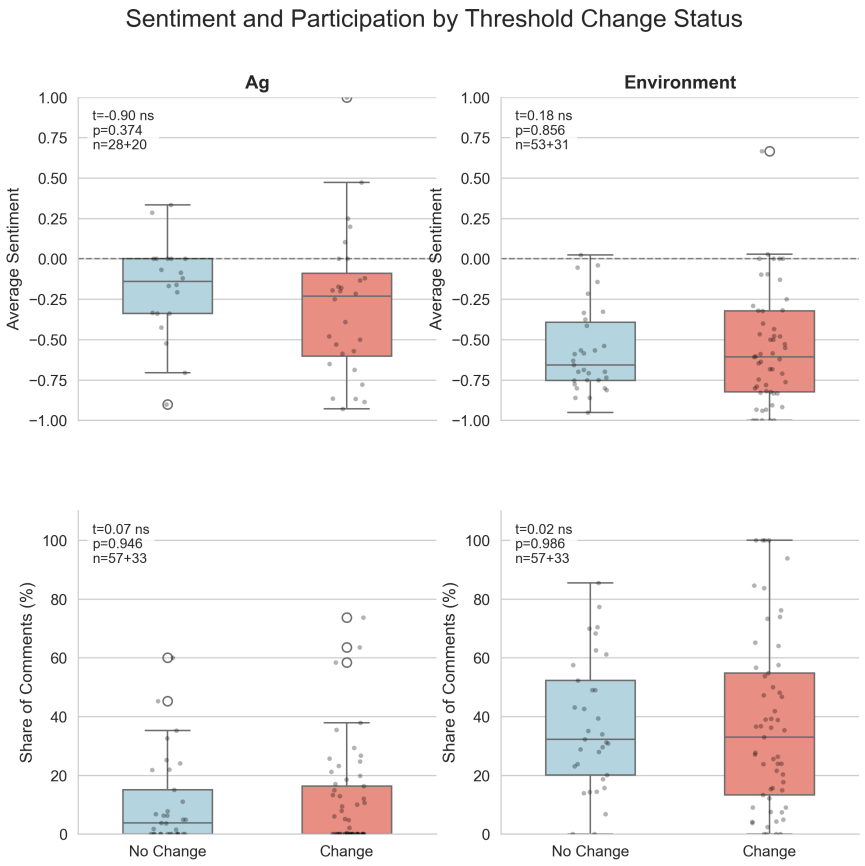



Figure 7.1: Difference in Means Between Plans with Changes and No Changes in Stringency and Stakeholder Sentiment and Participation. We find no difference in stakeholder participation in plans with and without changes to regulatory stringency.

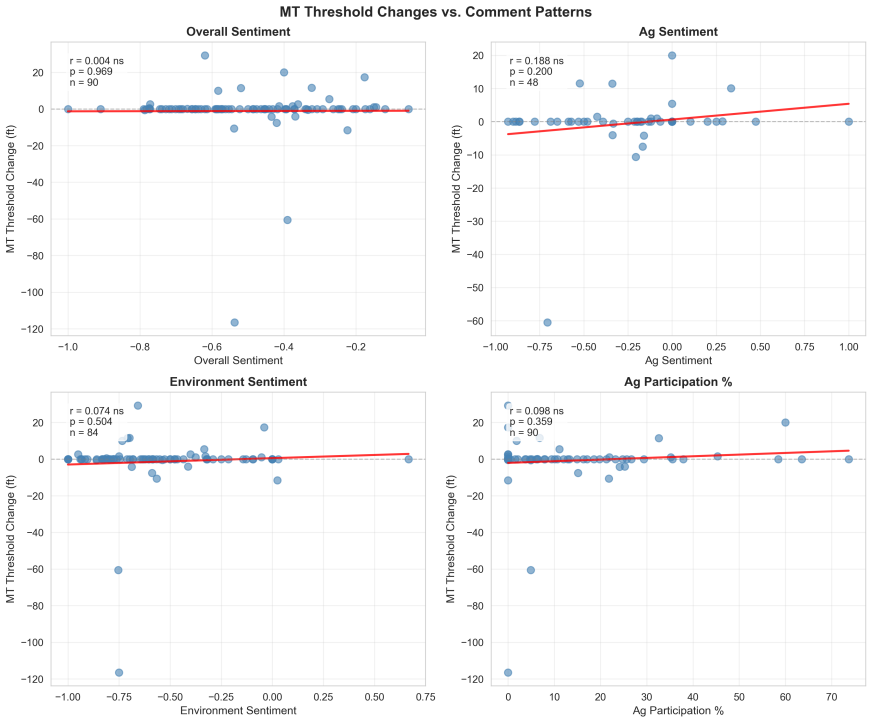



Figure 7.2: Correlation between the Magnitude of Change in Future Permitted Groundwater Level and Stakeholder Sentiment and Participation. We find no systematic relationship between stakeholder participation and magnitude of change in future permitted groundwater level between draft and final plans.

# 8 Data and Methods

In total, we obtain 18,061 individual comments for 106 GSPs in 83 basins. Six basins did not receive any comments, or received comments but did not provide commenter names or full letters, so we exclude them from analysis.

In most plans, agencies followed a standardized format when organizing and responding to public comments. They constructed public comment/response tables where letters were split into multiple pieces according to topic, since a single letter may address several issues. We manually extract the comment matrices from these plans. A subset of plans provided comment letters in full, without standardized formatting. To maintain consistency, we mimicked the agency formatting by parsing the letters into thematic segments, while excluding introductions, attachments and tables.

We combine all comments into a single database. We use python pdf libraries `pdfplumber` and `tabula` to extract the table structure, including the commenter name, affiliation (where available), comment text and the GSA response. Where automated extraction was not possible due to formatting issues, we extract the text manually.

We manually attribute a sector to each commenter. Using identities and affiliations, we construct a sector variable that classifies each comment as originating from agriculture, environment, government, tribes, water agencies, drinking water NGOs or non-affiliated individuals. This variable enables us to identify if themes and concerns differ across stakeholder groups.

We classify comment topics and sentiment using Claude 4 Sonnet [3]. We choose 19 predetermined topics that align with the sections in the GSPs to use in the classification prompt (see Appendix C for the full prompt). For each comment, Claude assigns one of 19 topics and provides a confidence score (0–1); scores below 0.7 are considered low-confidence and are flagged for human review. Claude also assigns a sentiment (positive (1), negative (-1), neutral(0)), identifies whether the comment refers to the GSP specifically, and codes stance

[3]May 2025 release, claude-sonnet-4-20250514

(supports (1), opposes (-1), neutral (0)). Finally, Claude flags comments which may contain multiple topics, though we require it to choose a predominant theme for each comment.

Since out primary unit of analysis is a groundwater basin, and sentiment and topic are computed at comment level, we aggregate these metrics by computing mean sentiment scores, sectoral share of comments by basin and count of topics. This allows us to align stakeholder participation to sustainability metrics, which are already defined at the basin level.

Sustainability metrics are constructed independently of comment analysis using plan reported hydrologic data, described in Appendix D. Finally, we estimate Pearson correlations between basin-level participation metrics and regulatory variables, weighing each basin equally. We report correlation coefficients, p-values and sample size for each correlation.

# 9 Discussion

We find that stakeholder coalitions respond to future regulatory stringency, but not to the amount of current groundwater overdraft. Ag stakeholders submit more negative comments on more stringent plans, and the pattern is reversed for environmental NGOs. Despite substantial strategic participation, the minimum thresholds largely do not change between draft and final plans. Where thresholds did change, there was no association between the change and the volume or sentiment of participation.

Why did regulatory stringency remain largely unchanged, despite extensive stakeholder participation? The most likely explanation is procedural. Public comments arrived late in the process, only months before the final plans were due for submission to the state. Revisions to thresholds would be difficult and costly to do on a compressed timeline, since many GSAs relied on consultants for groundwater modeling and other technical analyses. Another explanation is that regulatory stringency was already shaped by stakeholders earlier in the process through other forms of participation. If thresholds in draft plans were already

a result of negotiation and compromise in board and committee meetings and workshops, then public comment process would be less likely to affect changes. For example, (**?**) finds that stakeholder integration in the plan development leads to more protective plans. Finally, it is possible that instead of changing the plans, stakeholders commented to establish administrative record: in anticipation of future litigation, or to signal their positions to the Department of Water Resources which ultimately made decisions about adequacy of plans.

We bring new empirical evidence to two central questions in notice and comment rulemaking literature: what motivates stakeholder participation, and whether public input influences rules. The implementation of the Sustainable Groundwater Management Act serves as a natura policy experiment to examine this question: close to a hundred basins developed groundwater management rules, under a common framework and timeline. Provisions in the law required stakeholder participation, creating a detailed record of engagement and allowing systematic comparisons. We exploit this setting by constructing consistent metrics of regulatory stringency and groundwater overdraft, and using a large language model to analyze over 18,000 comments submitted on the groundwater sustainability plans

Our findings ultimately imply that comment process is not where the negotiation over regulatory stringency happens. This raises important questions for policymakers involved in environmental regulation. If meaningful influence on rulemaking occurs earlier, through integration of stakeholders, then policymakers should focus on ensuring diverse stakeholder representation in governance. While public comment serves an important procedural justice role, the costs of the process (the delay in regulation, staff time to respond to comment etc.) should be weighed against the benefits. The comment-related delay was doubled in the context of SGMA, where plans submitted to DWR immediately went through another round of comment submitted to the state.

One limitation of our work is that our analysis relies on Pearson correlations, which are not causal estimates. Thus, we cannot rule out unobserved variables that may jointly determine both regulatory stringency setting and public comment participation. Next, we

measure how comments affect changes in permitted groundwater decline - but comments could also result in changes to projects and management actions - or other aspects of the plans. Finally, we choose 2015 as the baseline for calculating our permitted groundwater decline, and use a median across the wells to create a basin level metric. While we show that using 2023 as a baseline yields similar results (see D), there could be alternative ways to aggregate sustainability metrics.

# References


Bostic, D., L. Mendez-Barrientos, R. Pauloo, K. Dobbin, and V. MacClements. 2023. "Thousands of domestic and public supply wells face failure despite groundwater sustainability reform in California's Central Valley." *Scientific Reports* 13, Open access, Published: 08 September 2023.

Bruno, E.M., and K. Jessoe. 2024. "Designing water markets for climate change adaptation." *Nature Climate Change* 14:331–339.

—. 2021. "Missing markets: Evidence on agricultural groundwater demand from volumetric pricing." *Journal of Public Economics* forthcoming.

California Department of Water Resources. 2019. "Basin Prioritization." Accessed: February 6, 2025.

—. 2025. "GSP Monitoring Data." `https://data.cnra.ca.gov/dataset/gspmd`, California Natural Resources Agency Open Data Portal. Dataset contains monitoring sites and groundwater level, subsidence, and streamflow measurements submitted by GSAs via SGMA Portal's Monitoring Network Module. Accessed: YYYY-MM-DD.

California State Legislature. 2014. "California Water Code §10721, Sustainable Groundwater Management Act." Division 6, Part 2.74, Chapter 2, Added by Stats. 2014, Ch. 346, Sec. 3.

Carleton, T., L. Crews, and I. Nath. 2024. "Is the world running out of fresh water?" In *AEA Papers and Proceedings*. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 114, pp. 31–35.

Coase, R.H. 2013. "The problem of social cost." *The journal of Law and Economics* 56:837–877.

Costa, M., B.A. Desmarais, and J.A. Hird. 2019. "Public comments' influence on science use in US rulemaking: The case of EPA's national emission standards." *The American Review of Public Administration* 49:36–50.

Crow, D.A., E.A. Albright, and E. Koebele. 2016. "Environmental Rulemaking Across States: Process, Procedural Access, and Regulatory Influence." *Environment and Planning C: Government and Policy* 34:1222–1240.

—. 2020. "Evaluating Stakeholder Participation and Influence on State-Level Rulemaking." *Policy Studies Journal* 48:953–981.

Department of Water Resources. 2017. "Sustainable Management Criteria Best Management Practice." Technical report, Department of Water Resources, November.

Escriva-Bou, A., H. McCann, E. Blanco, B. Gray, E. Hanak, J. Lund, B. Magnuson-Skeels, and A. Tweet. 2016. "Accounting for California's Water: Technical Appendix." Technical appendix, Public Policy Institute of California.

Gordon, H.S. 1954. "The economic theory of a common-property resource: the fishery." In *Classic Papers in Natural Resource Economics*. Springer, pp. 178–203.

Haeder, S., and S.W. Yackee. 2015. "Influence and the Administrative Process: Lobbying the U.S. President's Office of Management and Budget." *American Political Science Review* 109:507–522.

Hartman, R. 2018. "The harvesting decision when a standing forest has value." In *Economics of Forestry*. Routledge, pp. 295–302.

Hotelling, H. 1931. "The Economics of Exhaustible Resources." *Journal of Political Economy* 39:137–175.

Hui, I., N. Ulibarri, and B. Cain. 2018. "Patterns of Participation and Representation in a Regional Water Collaboration." *Policy Studies Journal*, pp. , First published: 04 November 2018.

Jasechko, S., H. Seybold, D. Perrone, Y. Fan, M. Shamsudduha, R.G. Taylor, O. Fallatah, and J.W. Kirchner. 2024. "Rapid Groundwater Decline and Some Cases of Recovery in Aquifers Globally." *Nature* 625:715–721.

Kerwin, C.M., and S.R. Furlong. 2018. *Rulemaking: How government agencies write law and make policy*. Cq Press.

Legislative Analyst's Office. 2021. "What Can We Learn From How the State Responded to the Last Major Drought?" Policy report, Legislative Analyst's Office, Sacramento, California.

Lund, J., J. Medellin-Azuara, J. Durand, and K. Stone. 2018. "Lessons from California's 2012–2016 Drought." *Journal of Water Resources Planning and Management* 144, Accessed: February 6, 2025.

Naughton, K., C. Schmid, S.W. Yackee, and X. Zhan. 2009. "Understanding Commenter Influence during Agency Rule Development." *Journal of Policy Analysis and Management* 28:258–277.

Pigou, A. 2017. *The economics of welfare*. Routledge.

Richey, A.S., B.F. Thomas, M.H. Lo, J.T. Reager, J.S. Famiglietti, K. Voss, S. Swenson, and M. Rodell. 2015. "Quantifying renewable groundwater stress with GRACE." *Water resources research* 51:5217–5238.

Rosenblum-Sellers, H. 2025. "Good Cause for Goodness' Sake: A New Approach to Notice-and-Comment Rulemaking." *Michigan Law Review*, October, pp. , MLR Online.

Scott, T., N. Ulibarri, and R. Scott. 2020. "Stakeholder involvement in collaborative regulatory processes: Using automated coding to track attendance and actions." *Regulation & Governance* 14:219–237.

Scott, T.A., N. Ulibarri, and R.P. Scott. 2018. "Stakeholder involvement in collaborative regulatory processes: Using automated coding to track attendance and actions." *Regulation & Governance* 14:219–237.

Siebert, S., J. Burke, J.M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F.T. Portmann. 2010. "Groundwater use for irrigation – a global inventory." *Hydrology and Earth System Sciences* 14:1863–1880.

Smith, K.B., and M.J. Licari. 2006. *Public Administration: Power and Politics in the Fourth Branch of Government*. [Publisher name needed].

Smith, R.R. 2021. "SGMA in the Field: Early Efforts at Defining Sustainability in California's Critically Overdrafted Basins." *University of the Pacific Law Review* 52, Rebecca R.A. Smith is a Partner at Downey Brand LLP.

Stavins, R.N. 2011. "The problem of the commons: Still unsettled after 100 years." *American Economic Review* 101:81–108.

Ulibarri, N., T.A. Scott, and O. Perez-Figueroa. 2019. "How does stakeholder involvement affect environmental impact assessment?" *Environmental Impact Assessment Review* 79:106309.

UNESCO World Water Assessment Programme. 2022. "United Nations World Water Development Report 2022: Groundwater: Making the Invisible Visible." Accessed online.

Yackee, J.W., and S.W. Yackee. 2006. "A bias towards business? Assessing interest group influence on the US bureaucracy." *The Journal of Politics* 68:128–139.

Yackee, S.W. 2019. "The politics of rulemaking in the United States." *Annual Review of Political Science* 22:37–55.

# A Appendix A. Public Comment Results

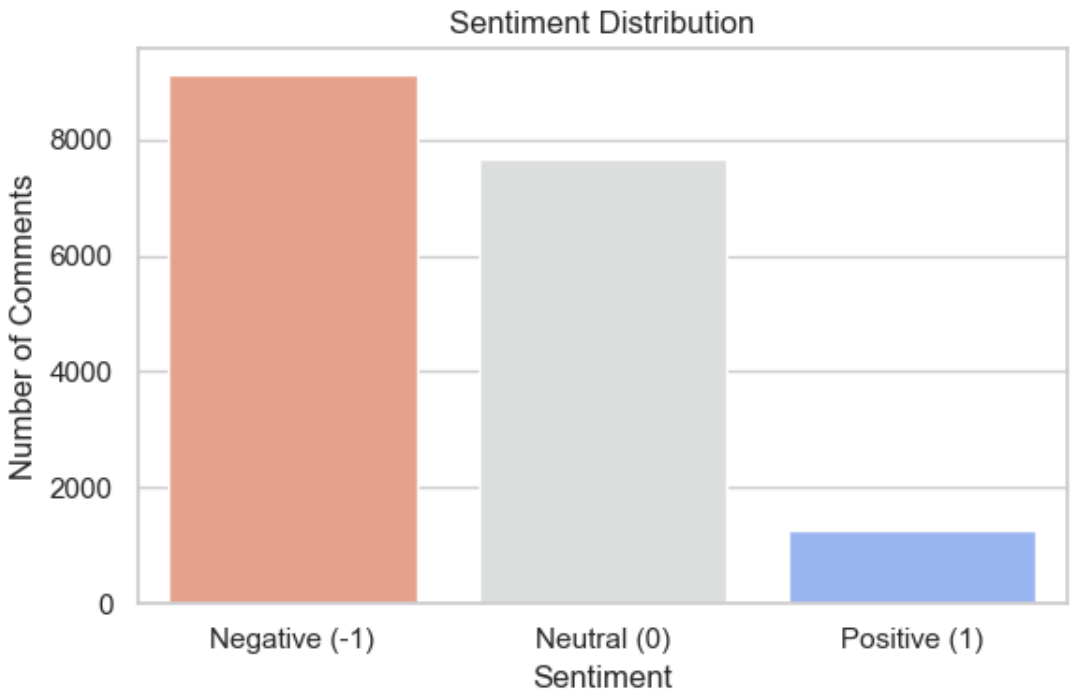



Figure A.1: Distribution of comment sentiment, showing number of positive (1,280), neutral (7,674) and negative comments (9,107).

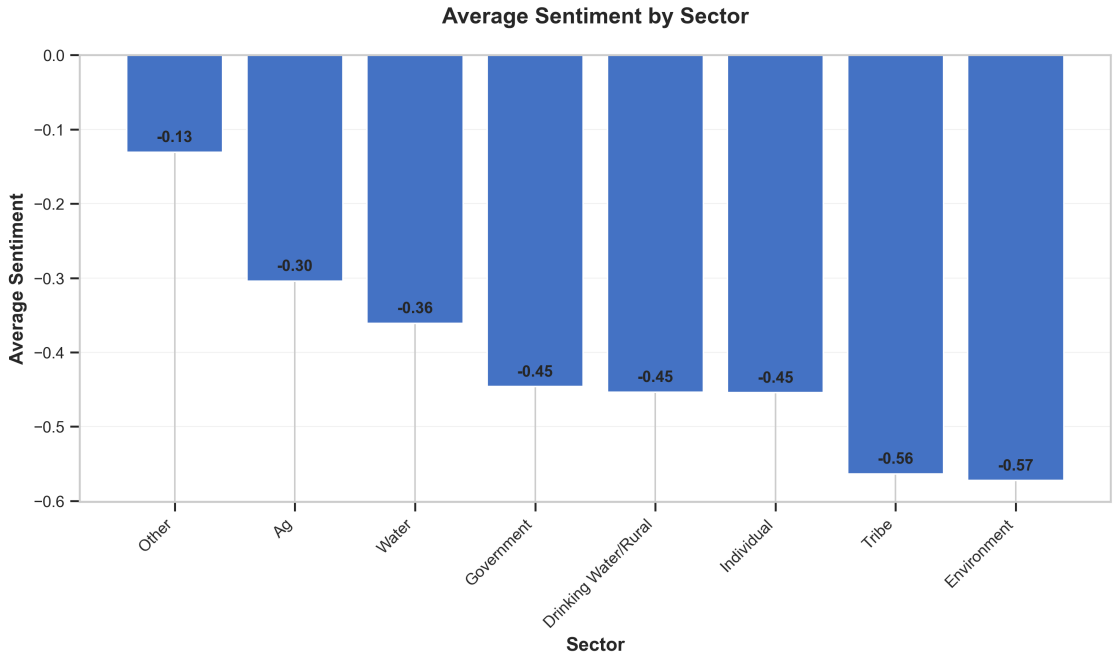



Figure A.2: Average sentiment by sector. Tribes and environmental sector exhibit the lowest average sentiment, while ag and water agencies show the highest average sentiment.

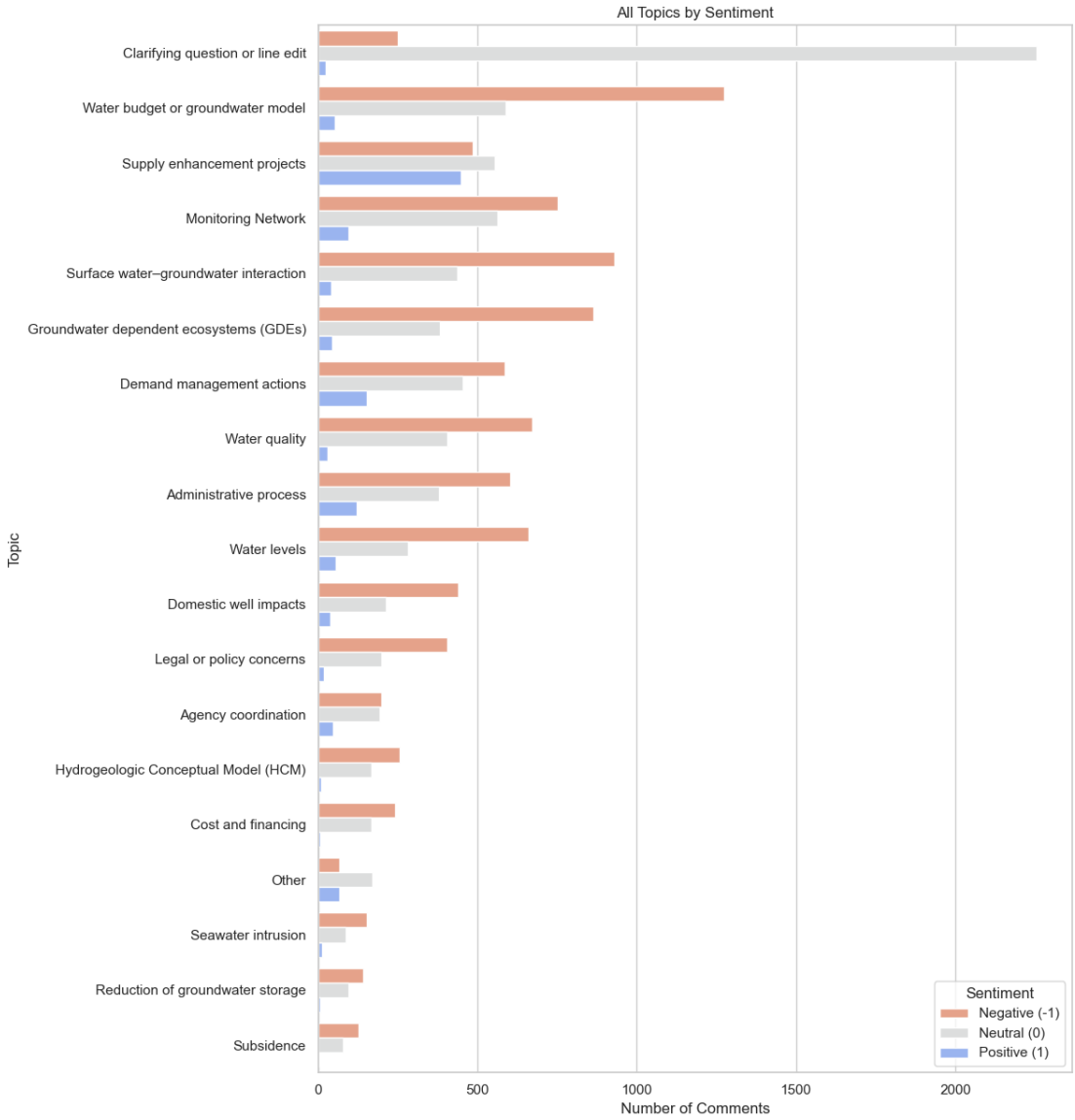



Figure A.3: Figure displays number of comments by topic and sentiment. Line edits account for the largest neutral sentiment category, water budget or groundwater model largest negative category and supply enhancement projects largest positive category.

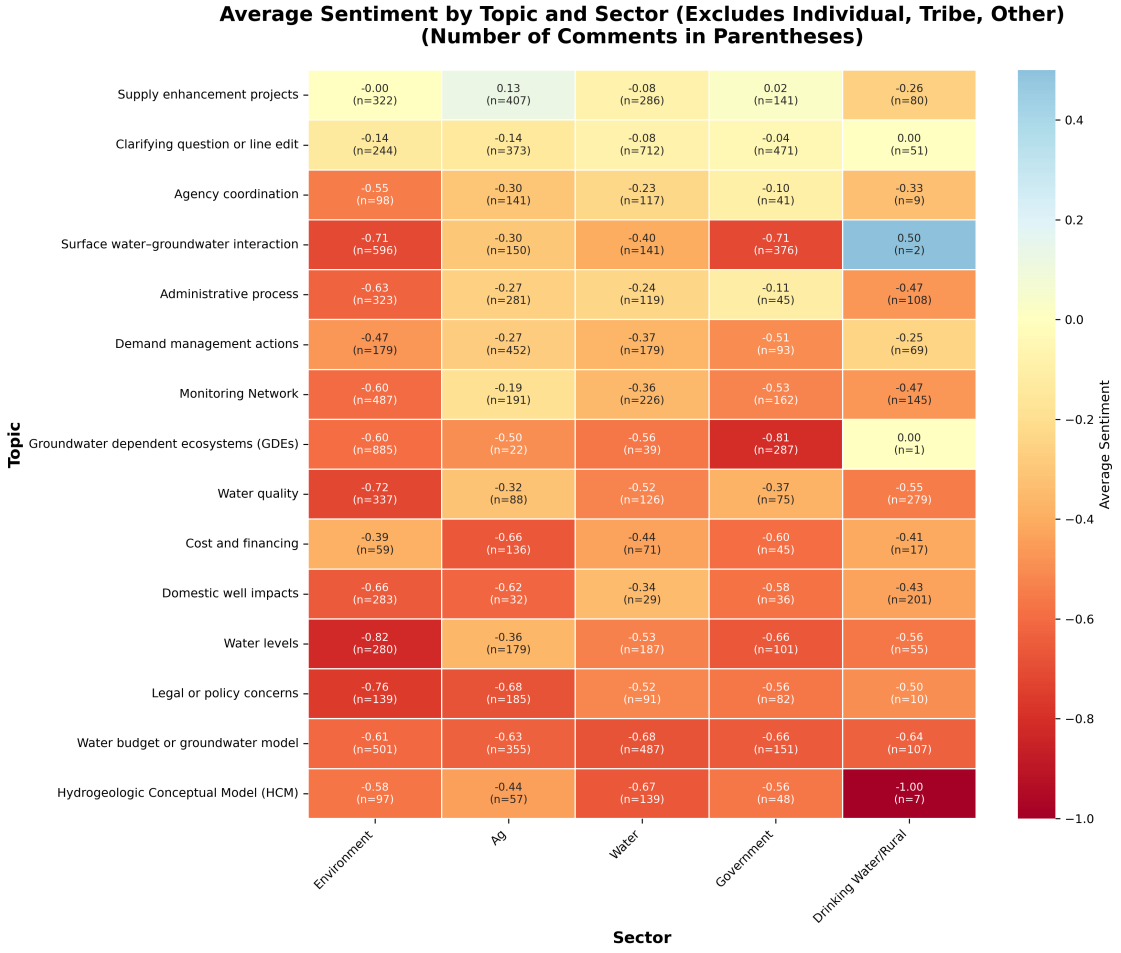



Figure A.4: Average sentiment by sector and topic. Number of comments in the parentheses. Exlcudes the breakdown for individuals, tribes and other sector.

# B Appendix B. Sentiment Correlation and Validation

We examine the relationship between stance and sentiment using conditional probabilities. We find that opposing stance was strongly associated with negative sentiment (99.6%), and support with positive (79.7%) and neutral sentiment (20.2%). Likewise, neutral stance is associated with neutral sentiment (99.9%). The strong link between opposition and negative sentiment indicates that opposition to the plans uses a more emotionally charged negative language. Supporters, in contrast, show more neutral sentiment. In general, the close relationship between stance and sentiment seems to suggest that sentiment analysis classifier is performing as expected.

Table 3: Distribution of sentiment scores by stance category

| **Stance** | **Negative (-1)** | **Neutral (0)** | **Positive (1)** |
|---|---|---|---|
| Support | 0.001252 | 0.202128 | 0.796621 |
| Neutral | 0.000954 | 0.998774 | 0.000273 |
| Oppose | 0.996047 | 0.003624 | 0.000329 |
| Unknown | 0.000000 | 1.000000 | 0.000000 |

We find that the vast majority of the comments refer to the plans B.1, and only a small subset targets other topics or general processes.

Finally, we present several validation checks below. We prompt Claude to assign only one topic to each comment. As a validation, we ask it to flag comments which it considers to be multi-topic. In figure B.2 we show that Claude found most comments to be single topic (73.8%), but a sizable chunk of comments is classified as multi-topic (26.2%). This does present an opportunity for further review. A t-test for difference in means ($t = 33.700$, $p < 0.001$) finds that comments flagged for multiple topics tend to have a longer word count on average. This suggest that Claude is classifying longer comments as multi-topic - indicating that there is a real complexity in the text, as opposed to classification error. Further work may include human review, and allowing Claude to provide multiple themes for those comments as opposed to forcing it to choose a primary theme.

As an additional validation check, we asked Claude to provide a topic classification "confidence" score. As a rule of thumb, a high confidence score (above 0.7) indicates high confidence in the topic, while a low score (below 0.7) indicates low confidence in the selected topic. In the histogram B.3, we find that vast majority of the classified comments have a score of 0.7 and above. In fact only 88 comments of the 18k total are below the threshold. We do find a small, but statistically significant difference ($t = 14.794$, $p < 0.0001$) in confidence score means in multi-topic (0.856) and single topic (0.869) comments. Still, for both groups, Claude confidence scores are above 0.7, on average.

Looking at the confidence scores by topic, we find that most low confidence comments (76 out of 88) were placed in the "Other" category. This makes intuitive sense - other is a category where Claude classified comments that don't fit into other categories, so naturally the difficult to classify comments would belong here. In figure B.4, we show a boxplot chart of Claude confidence scores by topic. It seems that topics with more distinct language/terminology show higher confidence scores. For example, subsidence, line edits, groundwater

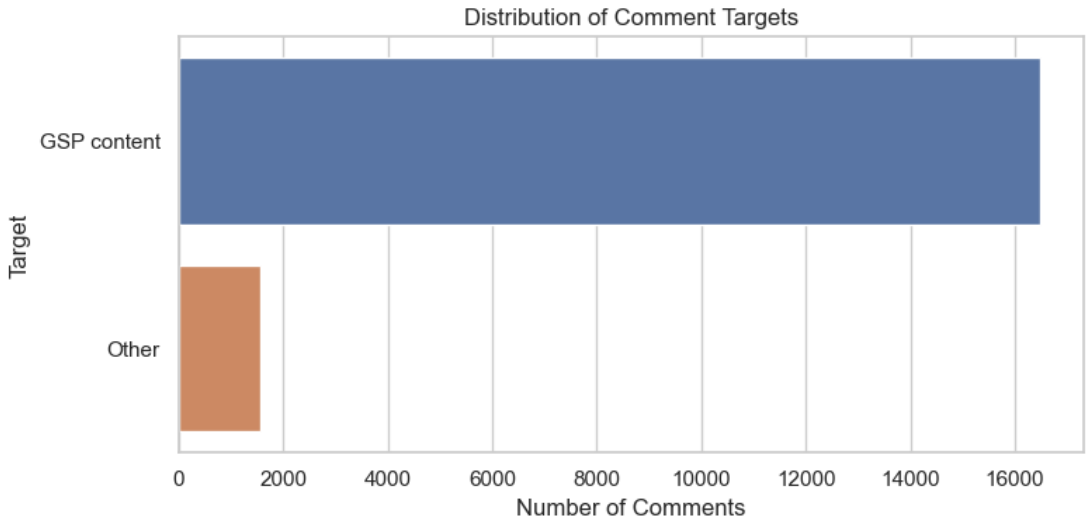



Figure B.1: Number of comments about the GSP content and about other concerns.

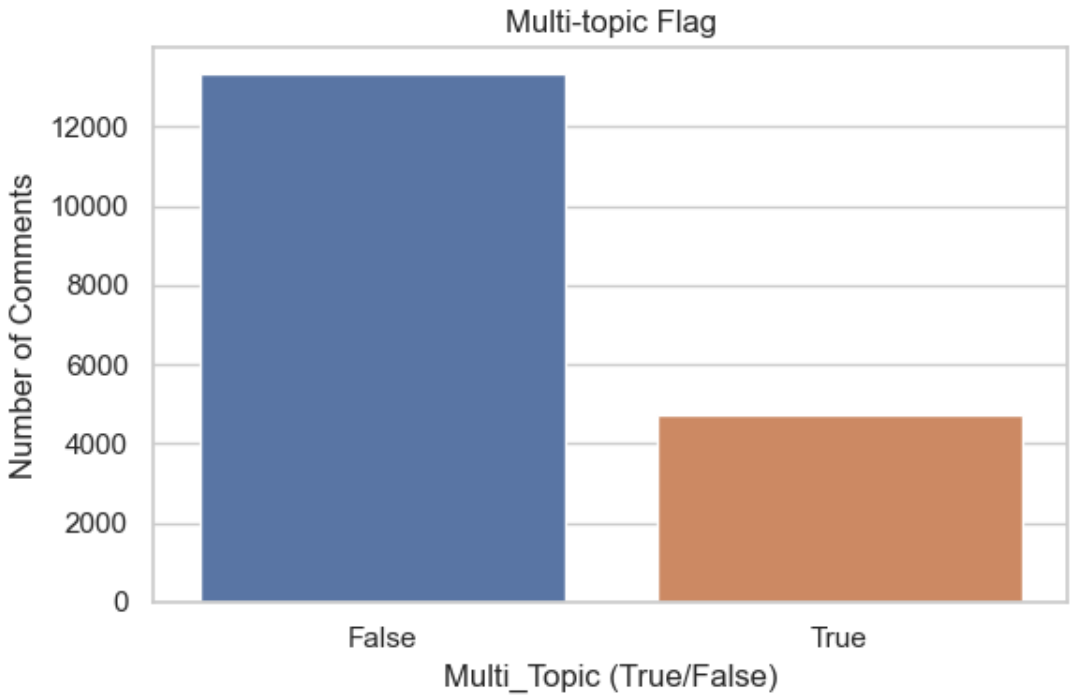



Figure B.2: Number of comments classified as single topic, and those flagged as multi-topic.

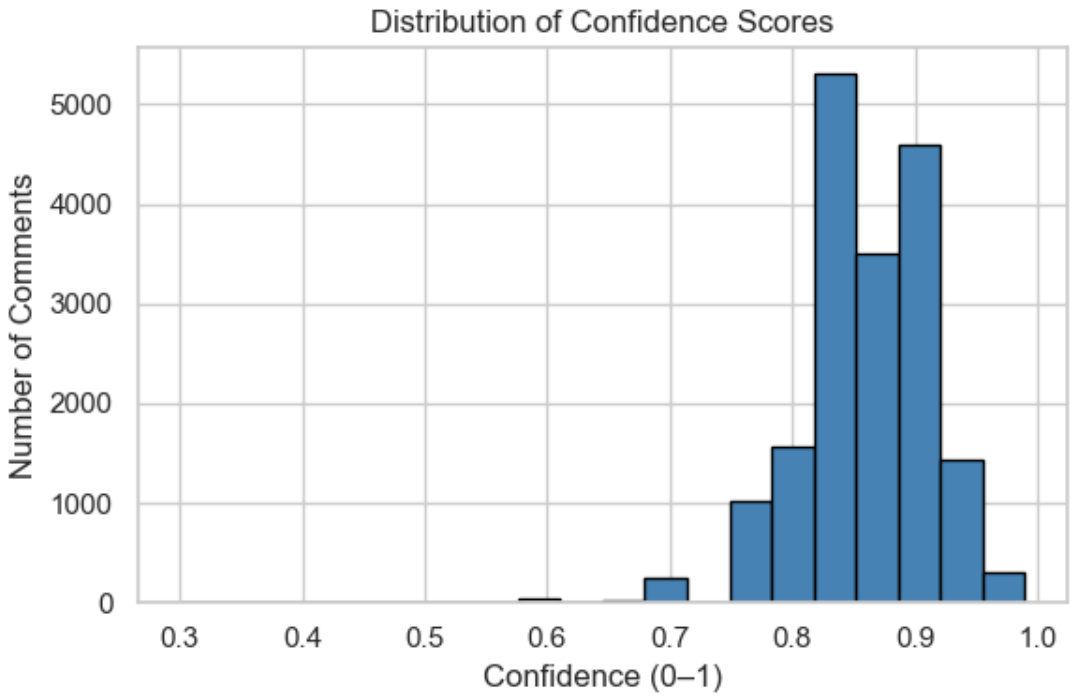



Figure B.3: Distribution of topic-classification confidence scores. Scores above 0.7 indicate high confidence in the topic. 88 comments are classified below 0.7

dependent ecosystems, water quality, don't necessarily overlap with other topics. Whereas, groundwater storage reduction, water levels, water budgets and hydrologic conceptual model are more related and could exhibit some level of overlap, reducing the confidence scores on average.

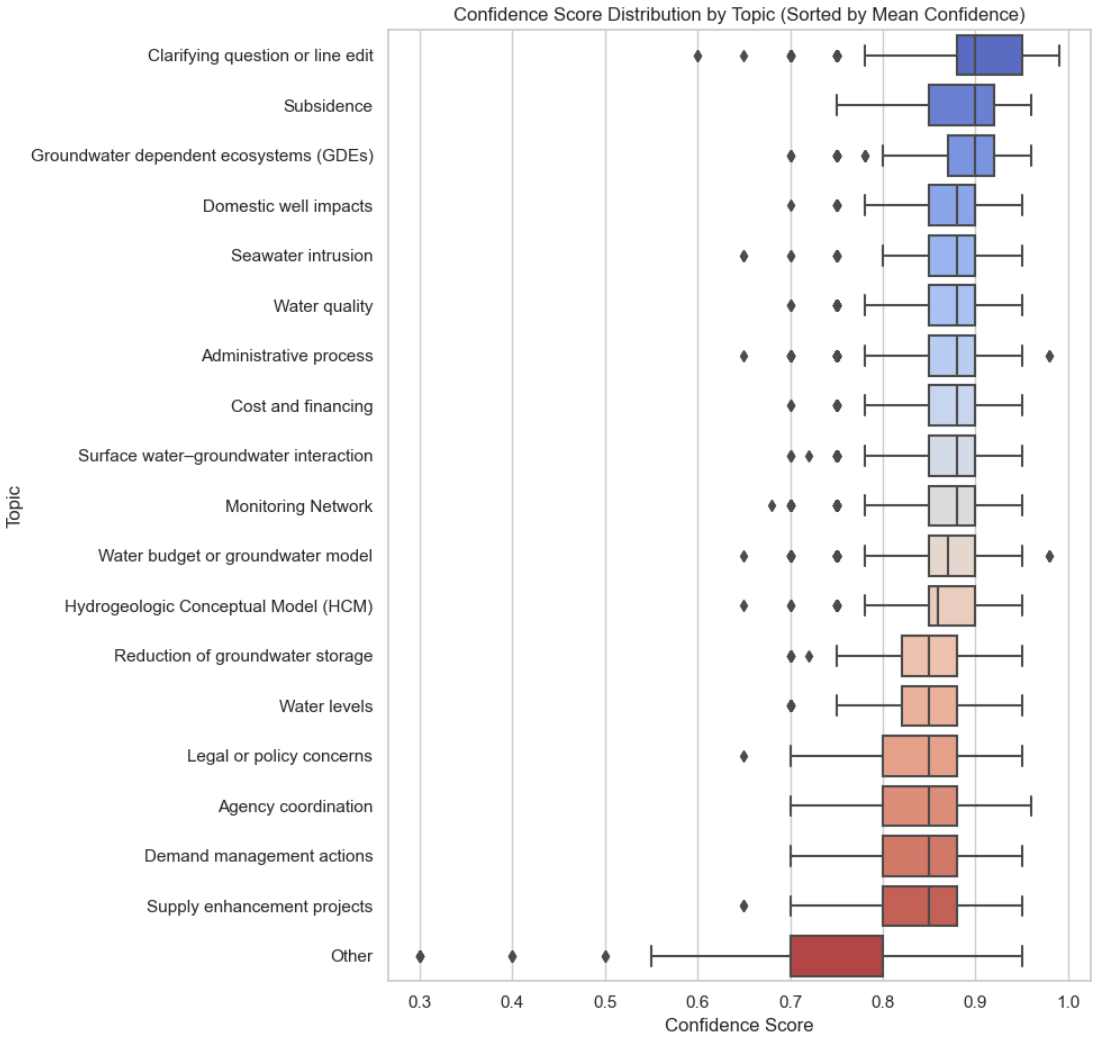



Figure B.4: Boxplot of confidence scores by topic. Topics with more distinct language/terminology show higher confidence scores.

# C Appendix C. Public Comment Analysis: Full Prompt for Claude

```
You are analyzing public comments submitted on groundwater sustainability plans (GSPs) in
     California. Each comment may express feedback, concerns, or requests about technical
     content, policy choices, or planning procedures.

Your task is to do four things for each comment:

1. Assign **ONE topic** from a predefined list that best captures the main focus of the
    comment.
2. Assign a **sentiment score**:
   - -1 = negative tone (complaints, opposition, concern)
   - 0 = neutral or factual
   - 1 = positive tone (support, agreement, encouragement)
3. Assign a **target**:
   - "GSP content" the comment addresses specific technical or substantive components of
       the plan (e.g., thresholds, models, objectives, water budgets, projects, sections,
        appendices, line edits on the plan text)
   - "Other" the comment addresses planning procedures, public engagement, transparency,
       general opinions, or any content not directly tied to the GSP's technical design
4. Assign the **stance** the commenter takes toward the target:
   - "support" the comment expresses agreement or approval
   - "oppose" the comment expresses disagreement, concern, or objection
   - "neutral" the comment is descriptive, factual, or does not clearly indicate a
       position

Also include:
- confidence a float from 0 to 1 representing how confident you are in the topic
    classification
- "multi_topic": true or false true if the comment clearly references more than one topic
    , even though only one was assigned

Please provide a structured JSON list with exactly N items, one for each numbered comment
    , in the same order. Do not omit any.

---

### Allowed topics:
1. **Water levels** Declining groundwater levels, historic or current trends,
    sustainability, or the adequacy of minimum thresholds and measurable objectives.
2. **Water quality** Contamination, pollutants, drinking water safety, or adequacy of
    water quality minimum thresholds and measurable objectives.
3. **Subsidence** Land sinking, damage to infrastructure, or adequacy of minimum
    thresholds and measurable objectives for subsidence.
4. **Monitoring Network** Requests for more data collection, well monitoring, network
    design, or transparency; data gaps and representative monitoring sites (RMS).
5. **Surface watergroundwater interaction** Impacts on rivers, streams, wetlands, or
    springs, including adequacy of associated minimum thresholds and measurable
    objectives.
6. **Administrative process** Comments on public engagement, transparency, meeting
```

```
    procedures, timelines, or community inclusion in the planning process.
7. **Domestic well impacts** Dry wells, domestic water access issues, or protection of
    small well users.
8. **Supply enhancement projects** Recharge, reservoirs, stormwater capture, or managed
    aquifer recharge.
9. **Demand management actions** Pumping allocations, land fallowing, water markets, and
    other mechanisms to reduce water use.
10. **Legal or policy concerns** SGMA compliance, legal vulnerability, CEQA concerns,
    jurisdictional issues.
11. **Clarifying question or line edit** Minor factual clarifications, copyedits, or
    questions lacking substantive content.
12. **Water budget or groundwater model** Concerns about assumptions, uncertainty,
    calibration, or transparency of models and water budgets.
13. **Hydrogeologic Conceptual Model (HCM)** Stratigraphy, aquifer structure, hydraulic
    properties, or basin boundaries.
14. **Agency coordination** Comments about inter-basin agreements, overlapping GSAs, or
    inconsistent policies between agencies like county general plans, the Irrigated Lands
     Regulatory Program (ILRP), or other regional frameworks.
15. **Reduction of groundwater storage** Long-term storage loss, overdraft, or adequacy
    of minimum thresholds and measurable objectives related to groundwater storage.
16. **Groundwater dependent ecosystems (GDEs)** Impacts to springs, phreatophytes,
    wetlands, or sensitive species, including adequacy of related thresholds and
    objectives.
17. **Cost and financing** Implementation costs, fairness of fee structures, economic
    burden on stakeholders.
18. **Seawater intrusion** Concerns about saltwater contaminating coastal aquifers,
    including monitoring, modeling, and adequacy of protective thresholds and projects.
19. **Other** Comments that do not clearly fit into any category above.

---

### Topic Overlap Guidance

Some topic categories have overlapping language. Use the following guidance to resolve
    edge cases:

- Water levels vs. Reduction of groundwater storage
  - Use Water levels if the comment focuses on observed trends, well measurements, or
      concern about declining levels.
  - Use Reduction of groundwater storage if the comment refers to long-term loss of total
       storage volume, overdraft, or storage-related thresholds and sustainability
      indicators.

- Water levels vs. Domestic well impacts
  - Use Domestic well impacts if the comment highlights effects on individual or small
      wells.
  - Use Water levels if the comment is general (e.g., levels have dropped 20 feet basin-
      wide).

- Monitoring vs. Threshold adequacy topics
  - Use Monitoring for comments requesting more wells, transparency, better data sharing,
       or system design.
  - Use other topics (e.g., Water levels, Water quality, GDEs) if the comment focuses on
      whether thresholds or measurable objectives are sufficient.
```

```
- Surface watergroundwater interaction vs. GDEs
  - Use Surface watergroundwater interaction if the focus is on river or stream depletion
      , flows, or hydrologic connections.
  - Use GDEs if the focus is on ecological impacts plants, wildlife, sensitive species.

- Water quality vs. Seawater intrusion
  - Use Seawater intrusion when the concern is specifically about saltwater encroachment
      into coastal aquifers.
  - Use Water quality for all other contamination concerns (e.g., nitrates, arsenic,
      industrial pollution).

When in doubt, choose the topic that best matches the commenters intent, even if another
    technical category could also apply. Use "Other" with low confidence if the comment
    is too vague.

---

### Multi-Topic Comments

If a comment clearly spans more than one predefined topic:
- Still assign ONE primary topic (the most emphasized, detailed, or first-mentioned one).
- Set: "multi_topic": true

This allows downstream reviewers to track comments that reference multiple themes, even
    if only one was formally assigned.

---

### Output Format:
Return a JSON array, one object per comment, with the following keys:
- "topic": one of the 19 topics above (use the full name)
- "sentiment": -1, 0, or 1
- "target": "GSP content" or "Other"
- "stance": "support", "oppose", or "neutral"
- "confidence": a float from 0 to 1 representing your confidence in the topic
    classification
- "multi_topic": true or false

---

### Example Comments:
1. "The groundwater thresholds dont account for domestic wells in the southern part of
    the basin."
2. "Theres a typo in Table 2-3. It says storwater instead of stormwater."
3. "Glad the agencies are finally considering recharge basins its overdue."
4. "This plan doesnt adequately protect GDEs. You havent identified any thresholds."
5. "The water budget assumptions about crop use seem unrealistic."
6. "There wasn't enough community engagement during the planning process."

---

### Output:
[
```

```
  {
    "topic": "Domestic well impacts",
    "sentiment": -1,
    "target": "GSP content",
    "stance": "oppose",
    "confidence": 0.95,
    "multi_topic": false
  },
  {
    "topic": "Clarifying question or line edit",
    "sentiment": 0,
    "target": "GSP content",
    "stance": "neutral",
    "confidence": 0.98,
    "multi_topic": false
  },
  {
    "topic": "Supply enhancement projects",
    "sentiment": 1,
    "target": "GSP content",
    "stance": "support",
    "confidence": 0.92,
    "multi_topic": false
  },
  {
    "topic": "Groundwater dependent ecosystems (GDEs)",
    "sentiment": -1,
    "target": "GSP content",
    "stance": "oppose",
    "confidence": 0.88,
    "multi_topic": false
  },
  {
    "topic": "Water budget or groundwater model",
    "sentiment": -1,
    "target": "GSP content",
    "stance": "oppose",
    "confidence": 0.87,
    "multi_topic": false
  },
  {
    "topic": "Administrative process",
    "sentiment": -1,
    "target": "Other",
    "stance": "oppose",
    "confidence": 0.80,
    "multi_topic": false
  }
]

]
```

# D Appendix D. Overdraft and Sustainability Metrics

To calculate our future permitted groundwater decline metric, we first collect minimum thresholds (MT) for chronic decline of groundwater levels in draft and final groundwater sustainability plans. The minimum thresholds are defined at well level, and they represent the point at which undesirable results would start to occur under the law. A GSA can choose any number of monitoring wells at which to define thresholds.

The minimum thresholds are given in units of groundwater elevation, or in units of groundwater depth. We link our database of thresholds to monitoring site data provided by the Department of Water Resources to obtain ground surface elevation California Department of Water Resources (2025). For MTs in elevation units, we calculate "depth to water" by subtracting the groundwater elevation from the well's ground surface elevation. To compare the MT depth to water to recent conditions, we also calculate the depth to water at each representative site in 2015. Once the 2015 differences are calculated, we take a median per basin to represent the future permitted groundwater decline.

We chose 2015 as a baseline year because under SGMA, agencies are only required to address undesirable results occurring after January 1, 2015. Since a well can have multiple measurements in a given year, we used the measurement in June, or the first available measurement after June.

Of the 2,279 wells in our database that had a minimum threshold for water levels defined, 70.2% had 2015 readings available, and the remaining 30% did not. This could be due to lack of measurement or lack of publicly available data. If a well does not have a reading in 2015, it is by definition excluded from our definition of the "future permitted groundwater decline".

To assess whether excluding wells without 2015 data affects our results, we repeat the analysis using 2023 as a baseline year - since 82.4% of wells in the sample had measurements in that year. We find that addition of wells and a change in baseline year generally did not meaningfully affect our findings.

We find little evidence of a systematic relationship between pumping reduction and future permitted groundwater decline, two of our regulatory stringency metrics. Figure D.1 plots the future permitted groundwater decline and pumping reduction for each groundwater sustainability plan, and illustrates the fitted line using the data (in blue) as well as the fitted line if these metrics were perfectly correlated (in red). The lack of a correlation may occur because one of these metrics incorporates a regulatory threshold, while the other metric captures a water accounting exercise. The law requires plans to set minimum thresholds, one portion of the future permitted groundwater decline metric, that if exceeded can trigger regulatory intervention and legal action. In contrast, pumping reductions simply capture the reduction in historical water use that would need to occur to meet sustainable yields.

This figure makes clear that most groundwater plans explicitly allow for declining water tables, including in overdrafted basins. With few exceptions, almost all the reported future permitted groundwater decline metrics are negative, indicating that groundwater basins set their sustainability thresholds below the 2015 depth to the water table. The absence of a systematic positive correlation between the two metrics suggests that more stringent well level thresholds are not necessarily set in locations where larger pumping reductions are needed to meet sustainability targets.

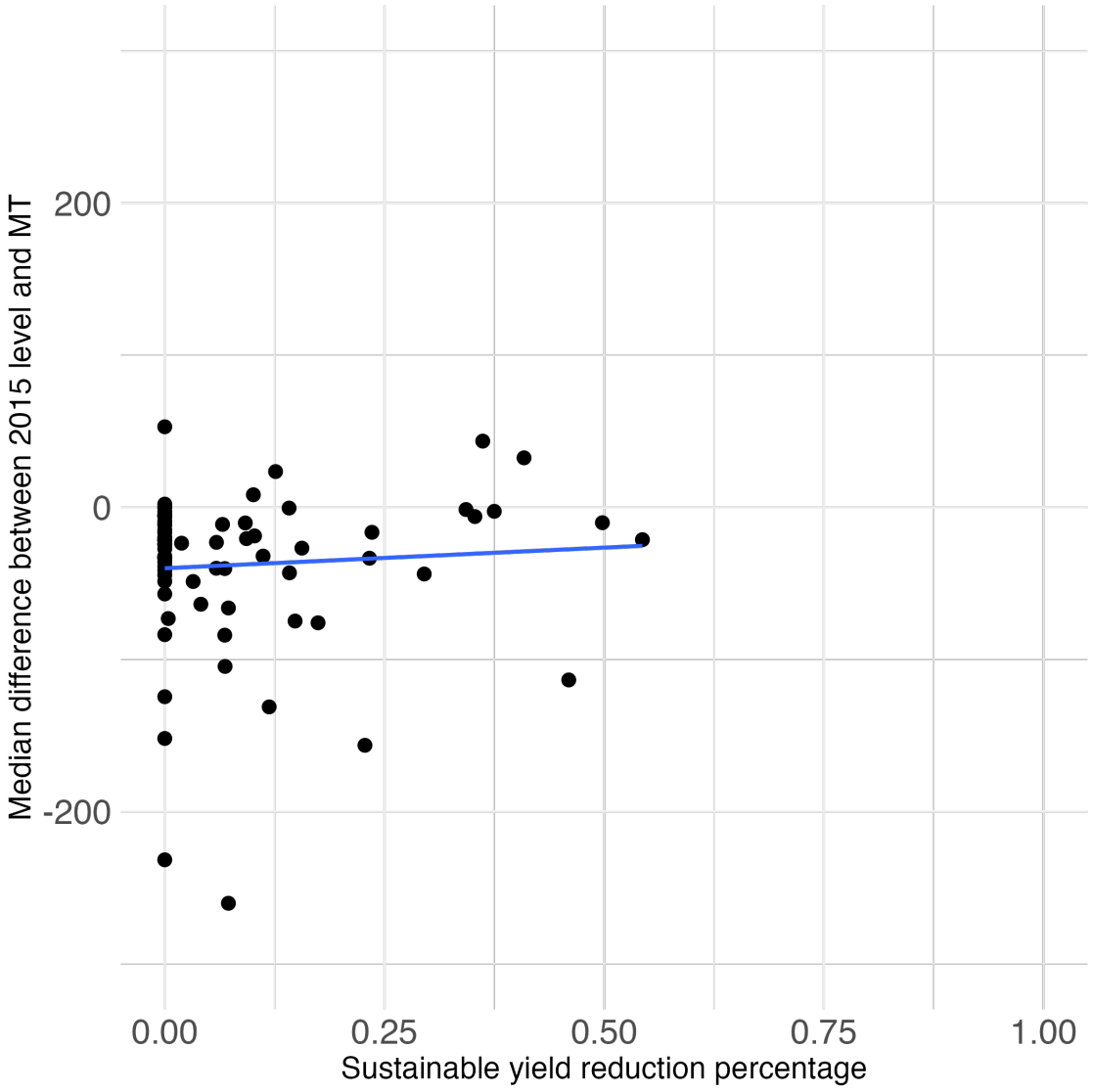

Median difference between 2015 level and MT
200
0
-200
0.00
0.25
0.50
0.75
1.00
Sustainable yield reduction percentage

Figure D.1

To better understand what drives threshold definition we examine how the future permitted groundwater decline correlates with the historically lowest groundwater levels recorded in each basin from 2000-15. We find that basins with historically deeper depths to the water table tend to set less protective thresholds relative to 2015 conditions ($r=-0.226$, $p=0.043$). This suggests that water managers are setting thresholds that are based on historic lows to allow more management flexibility, rather than establishing protective targets.