Scientific Reports

https://doi.org/10.1038/s41598-026-60638-z

Article in Press

Physical water system consolidation alone cannot resolve drinking water failure in California's San Joaquin Valley

Received: 23 March 2026

Accepted: 29 June 2026

Published online: 02 July 2026

Cite this article as: Sangha L. Physical water system consolidation alone cannot resolve drinking water failure in California's San Joaquin Valley. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-60638-z

Laljeet Sangha

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

ARTICLE IN PRESS

Physical water system consolidation alone cannot resolve drinking water failure in California's San Joaquin Valley

Laljeet Sangha, Ph.D.[1]

[1]UC Cooperative Extension Advisor, University of California Agriculture and Natural Resources 2801 2nd St, Davis, CA 95618, email: lsangha@ucanr.edu

Abstract

Small community water systems in California's San Joaquin Valley disproportionately fail to provide safe drinking water, with most serving disadvantaged communities and lacking the technical, managerial, and financial capacity to maintain compliance. The State Water Board designates such systems as failing when they have documented violations of drinking water standards or cannot reliably deliver safe water. Physical consolidation, which involves connecting a smaller failing system to a larger provider through new pipeline infrastructure, is the primary long-term solution promoted by state regulators. Here, we assess consolidation feasibility and capital costs across four counties (Kern, Kings, Tulare, and Fresno) via the state's 2024 screening methodology. The methodology identifies smaller failing and at-risk systems as candidates for joining a larger receiving system with sufficient capacity. Among the 210 eligible systems, 114 (54%) matched a receiving system within three miles, with an estimated total capital need of $413.5 million. Nearly all identified pairs met state funding viability thresholds. The remaining 96 systems, including 49 classified as failing, lacked an eligible receiving partner and will require alternative solutions. Physical consolidation is necessary but insufficient for achieving safe drinking water access across the region.

Keywords

water system consolidation, drinking water, water systems, SAFER, infrastructure costs, California

Introduction

California regulates approximately 7,252 active public water systems, of which 2,815 are community water systems[1]. Of these, 2,407 (85.5%) serve fewer than 3,300 connections yet supply only a small fraction of the state's population. Small systems carry the greatest compliance burden relative to their size. The State Water Resources Control Board's Safe and Affordable Funding for Equity and Resilience (SAFER) risk assessment identified 390

ARTICLE IN PRESS

public water systems as failing and 589 as at-risk, together serving approximately 2.3 million people. Small systems are disproportionately represented: 84% of systems appearing on the 2024 cumulative Failing list average 246 service connections [1]. Nationally, health-based drinking water quality violations affected between 9 and 45 million people annually from 1982--2015, with the incidence of violations being significantly greater in rural and small systems [2]. Small systems face compounding disadvantages of higher per-unit costs, limited operator capacity, and greater dependence on single groundwater sources vulnerable to contamination [3]. In California, 79% of failing systems serve disadvantaged or severely disadvantaged communities [1]. Systems that remain on the failing list disproportionately serve communities of color compared with those that return to compliance[4].

Physical consolidation is the primary policy tool for addressing persistent drinking water failure in California. SB 88 (2015) granted the State Water Board authority to mandate consolidation where a system consistently fails to provide safe drinking water to a disadvantaged community. SB 200 (2019) established the Safe and Affordable Drinking Water Fund, providing up to $130 million annually for drinking water solutions, including consolidation. SB 403 (2021) further expanded consolidation authority to include At-Risk systems, enabling proactive intervention before violations become chronic. A total of 170 public water systems were consolidated statewide between 2019 and 2024, serving approximately 324,101 people [1]. Two hundred and six unique water systems were consolidated statewide between 2015 and 2021, with water quality compliance as the primary motivator in 81% of the cases[5]. Despite this progress, only two of those 206 consolidations were formally mandated. Higher-resource communities complete consolidation at roughly twice the rate of lower-resource ones (defined as 20% above or below the statewide average on at least two of three indicators: median household income, homeownership rate, and proportion of non-Hispanic White residents) [6].

Barriers to consolidation are primarily political rather than technical[7]. They include local sociopolitical dynamics and historical patterns of public investment that have bypassed marginalized communities, as documented in the contested consolidation of Tooleville with Exeter[4] and the consolidation of East Porterville following the 2012-2016 drought[8,9]. Loss of local control is the most commonly cited barrier, reported by 93% of state drinking water administrators surveyed nationally [10]. Systematic assessments of where consolidation is physically feasible and what it would cost at a regional scale remain limited. National proximity analysis revealed

ARTICLE IN PRESS

that 55% of U.S. community water systems are adjacent to another system and that 95% fall within 10 km of their nearest neighbor [11], suggesting that physical distance is often not the binding constraint. However, Euclidean proximity does not account for road-network access, terrain, or infrastructure constraints that determine whether consolidation is physically achievable at a given site.

The State Water Board's 2024 Drinking Water Needs Assessment includes a physical consolidation cost methodology that screens potential consolidation candidates statewide [12]. The methodology classifies each system using two attributes: SAFER risk status and number of service connections (Table 1). Failing systems with 1,000 or fewer connections and at-risk systems with 500 or fewer connections are potential joining systems. Failing systems qualify as receiving systems only if they exceed 1,000 connections, and all other SAFER categories qualify as receivers if they exceed 500 connections. A GIS-based spatial analysis determines proximity through boundary intersection and road-network routing and estimates capital costs via standardized unit cost assumptions for pipeline construction, service lines, connection fees, administration, and environmental review. The methodology is designed as a statewide screening tool to estimate funding needs and does not prescribe specific community solutions or evaluate alternatives to physical consolidation, such as treatment or managerial consolidation. This study applies it across the four counties to produce a system-level map of feasible pairs and to quantify the systems for which physical consolidation is not viable, the point at which policy attention must turn to alternatives such as treatment or new source development.

Kern, Kings, Tulare, and Fresno counties collectively contain 450 public water systems in the SAFER dataset, of which 118 are classified as Failing and 128 as At-Risk, serving a combined population of approximately 594,000 people. Kern, Fresno, and Tulare rank as the three counties with the highest counts of failing water systems in California. Groundwater contamination from nitrate, arsenic, 1,2,3-trichloropropane, and the legacy soil fumigant DBCP is widespread across the region[13], and many affected systems serve disadvantaged communities with limited financial and technical capacity. Consolidation activity in the region is already substantial. As of early 2025, 11 of the 15 mandatory consolidation projects statewide involved systems in Kern and Tulare Counties, including consolidations of small mutual water companies into Lamont PUD, East Niles CSD, and the City of Bakersfield.

ARTICLE IN PRESS

This study applies the State Water Board's 2024 physical consolidation methodology across Kern, Kings, Tulare, and Fresno Counties, extending it from a per-system screening procedure to a regional assessment that maps feasible consolidation pairs and identifies the systems without a viable partner[12]. Proximity is measured as road-network distance rather than straight-line distance, since connecting pipeline follows existing road corridors, and pairs within 3 miles of road distance are treated as feasible. Among the 210 eligible joining systems identified, 114 (54%) matched a receiving system through boundary intersection or road-network routing within three miles, with an estimated total capital need of $413.5 million. Nearly all pairs met state funding viability thresholds, and model results aligned with the state's active mandatory consolidation pipeline. The remaining 96 joining systems lacked an eligible receiver, highlighting that physical consolidation alone cannot resolve the region's drinking water challenges.

Results

System Classification

The analysis evaluated all public water systems in the four-county boundary layer against the State Water Board's 2024 classification criteria. Systems were assigned as potential joining systems, potential receiving systems, or neither, on the basis of SAFER status and service connection count (Table 1).

ARTICLE IN PRESS

Across the four study counties, 794 public water systems were evaluated for consolidation eligibility (450 with a SAFER risk classification and 344 without; Table 2). The classification identified 210 potential joining systems and 74 potential receiving systems. Kern County contained the largest number of joining systems (93), followed by Fresno (58) and Tulare (56). Kings County, with only 19 total systems, yielded 3 joining and 6 receiving candidates. Not all joining systems result in consolidation pairs; a system qualifies as a joiner on the basis of its own attributes but may lack an eligible receiver within the spatial proximity thresholds defined in the analysis. Systems without a SAFER classification could not qualify as joining systems and none met the receiving-system size threshold, so their inclusion did not affect the results.

Proximity Analysis

The proximity analysis identified 114 consolidation pairs across the four counties (Table 2). Among these, 45 pairs (39%) were identified through

ARTICLE IN PRESS

boundary intersections, where the service area of the joining system overlaps with that of the receiving system. The remaining 69 pairs (61%) were identified through road-network routing at distances of three miles or less. Across the four counties, 26 joining systems had more than one eligible receiver within the threshold (1 in Kings, 3 in Tulare, 4 in Fresno, and 18 in Kern) and were matched to the larger receiver. All remaining matched systems had a single eligible receiver.

Kern County produced the most consolidation candidates (64 pairs), reflecting both its large number of joining systems and the availability of nearby receivers. Tulare County yielded 22 pairs, predominantly through route analysis (17 of 22). Fresno County produced 26 pairs with similar distributions of intersect and route types. Kings County produced only 2 pairs, one of each type, which is consistent with its small system inventory and limited number of joining candidates.

The ratio of joining systems to consolidation pairs varied across counties. In Kern, 64 of 93 joining systems (69%) were matched to a receiving system. In Tulare, 22 of 56 (39%) were matched. In Fresno, 26 of 58 (45%) were matched. In Kings, 2 of 3 (67%) were matched. Systems that were not matched lacked an eligible receiver within the intersection or three-mile route threshold or had no routable road connection to a candidate receiver.

Figure 1 presents the spatial distribution of the consolidation pairs for each county, showing the joining systems, receiving systems, and road-network connection routes. Interactive versions of these maps with detailed system-level popups are available in the project repository (Data Availability Statement).

ARTICLE IN PRESS

Cost Estimation

The total estimated capital costs across all 114 pairs ranged from $449,000 to $14.9 million per project. Intersect consolidations were consistently less expensive than route consolidations because of the shorter assumed pipeline distance. Across all four counties, the median total capital cost was $1,081,740 for intersection pairs (IQR $709,506 to $1,807,598) and $5,432,863 for route pairs (IQR $2,718,522 to $6,417,305) (Table 3). The median cost per connection was $30,048 for intersection pairs (IQR $24,101 to $44,344) and $138,549 for route pairs (IQR $59,367 to $401,082). County-level cost breakdowns are presented in (Table 3).

The cost per connection exhibited substantial variation driven primarily by system size. For the intersection pairs, the median cost per connection

ARTICLE IN PRESS

ranged from $19,459 (Kings) to $31,478 (Kern) (Table 3). For route pairs, the median cost per connection ranged from $97,443 (Kern) to $1,930,513 (Kings). The extreme per-connection cost in Kings reflects a single route pair serving only 3 connections, where fixed project costs are distributed across a very small customer base.

This pattern of elevated per-connection costs in small systems was consistent across counties. Systems with fewer than 20 connections routinely exceeded $100,000 per connection regardless of route distance, whereas systems with more than 75 connections generally remained below $96,000 per connection (Figure 2). The relationship between system size and unit cost reflects the fixed-cost structure of consolidation projects: pipeline construction, environmental review, administration, and engineering costs are largely independent of the number of connections served. Intersection pairs therefore represent the lowest-cost connections in the analysis.

Funding Viability

Of the 114 consolidation pairs identified, 113 (99%) met the State Water Board's funding viability thresholds (Table 2). The single pair that exceeded the threshold was the Athal mutual water system in Kern County, a 55-connection at-risk system paired with the East Niles CSD through route analysis. Its estimated capital cost of $7.26 million marginally exceeded the $7.2 million total cost cap applied to systems with fewer than 75 connections by a few thousand dollars. This system is currently in a mandatory consolidation process with East Niles CSD, indicating that the state has determined the project to be viable through a more detailed engineering and funding evaluation than the screening model provides.

ARTICLE IN PRESS

The total estimated capital need for all fundable consolidation pairs across the four counties was approximately $413.5 million, with Kern County accounting for the largest share ($222.0 million), followed by Tulare ($97.3 million), Fresno ($79.0 million), and Kings ($15.2 million). These estimates represent the modeled cost of physically connecting all identified joining systems to their matched receivers and do not include additional costs for treatment upgrades, backup source development, or other essential infrastructure that may be required in individual systems (Table 2).

Discussion

Consolidation Potential in the San Joaquin Valley

ARTICLE IN PRESS

Nearly half of the joining systems identified across the four counties lack a physical consolidation path under the current screening criteria. Among the 210 classified joining systems, 96 (46%) lacked a consolidation path, including 49 failing systems and 47 at-risk systems. The unmatched systems concentrate in Tulare County (34 of 56, 61%), and Fresno County (32 of 58, 55%) and with a lower proportion in Kern (29 of 93, 31%). These systems face compounding disadvantages of geographic isolation and the limited institutional capacity that halves consolidation completion rates in lower-resource communities[6].

Geography is the most frequently cited barrier to consolidation nationally, reported by state drinking water administrators in more than 85% of the states surveyed[10]. The 39% boundary-intersection rate observed here falls well below the 55% national adjacency rate, reflecting the dispersed settlement of San Joaquin Valley agricultural Valley agricultural communities[11]. Similar geographic constraints limit regionalization in Central Appalachia, where rugged terrain and sparse settlement increase unit infrastructure costs and reduce the pool of viable receiving systems[14]. A similar pattern appears in central and west Texas, where regional water planning measured small-system densities ranging from one system per 26 square miles in clustered areas to one per 108 to 188 square miles in dispersed rangeland[15]. In the most dispersed subregions, connecting systems would require transmission lines of 18 to 30 miles, distances the analysis identified as too far to convey water economically. Geographic isolation, rather than rurality alone, set the limit on feasible regionalization.

For these systems, alternative long-term solutions such as centralized treatment, decentralized treatment, or managerial consolidation need to be evaluated. For systems where physical consolidation is not feasible, shared-service arrangements such as joint operators, regional procurement, and managerial partnerships offer alternative pathways to improve technical and financial capacity[16]. Among California consolidations completed between 2015 and 2021, water quality compliance was the primary driver in 81% of cases[5], underscoring the urgency of identifying alternative solutions for the 49 Failing systems that lack a consolidation partner. System fragmentation compounds the risks for unmatched systems, as small uncoordinated providers lack the supply redundancy and interconnections that support drought resilience [17].

For the remaining 114 joining systems, physical consolidation is spatially feasible, with an estimated total capital need of $413.5 million, of which 113

ARTICLE IN PRESS

met the State Water Board's funding viability thresholds. Kern County, with the largest inventory of joining systems (93) and the most receiving systems (34), produced 64 pairs accounting for more than half of the total estimated capital need ($222.0 million). Tulare County produced 22 pairs ($97.3 million) despite having 56 joining systems, reflecting fewer eligible receivers and greater distances between systems. Fresno County produced 26 pairs ($79.0 million), and Kings produced only 2 pairs ($15.2 million). This variation reflects the spatial configuration of potential receivers relative to joining systems, not simply the number of failing systems in each county.

Validation Against Active Consolidation Projects

The model's results align closely with the State Water Board's current mandatory consolidation pipeline. In Kern County, all 11 systems currently in mandatory consolidation proceedings appeared in the analysis and were matched to the same receiving systems selected by the state: Lamont PUD, East Niles CSD, City of Bakersfield, and CWS-Bakersfield. Ten of these 11 pairs were classified as fundable. One exception, the Athal Mutual Water System ($7.26 million, 55 connections), marginally exceeded the $7.2 million total cost cap for systems with fewer than 75 connections[1]. The state has nonetheless proceeded with this consolidation through its mandatory authority, illustrating that the viability thresholds function as screening parameters rather than absolute funding limits.

In Tulare County, the model identified East Orosi CSD paired with Orosi PUD ($4.5 million) and Sultana CSD paired with the City of Dinuba ($9.3 million), both of which correspond to active or planned consolidation efforts[1]. The model did not identify the Cutler PUD as a joining system because its service connections exceed the 500-connection threshold for at-risk systems. This is consistent with the classification criteria rather than a model failure; the state's mandatory authority is not limited by the screening thresholds used in the cost assessment. In Fresno County, Las Deltas Mutual Water Company was identified and paired with the City of Firebaugh ($1.8 million), matching a system with an active state-funded administrator appointment. This convergence between the model output and the state's active consolidation pipeline provides confidence that the screening methodology produces results that are consistent with consolidation decisions being implemented on the ground.

Cost Drivers and System Size

ARTICLE IN PRESS

The cost analysis reveals a consistent pattern across all four counties: system size is the dominant driver of per-connection costs. Small systems with fewer than 20 connections routinely exceeded $100,000 per connection, whereas systems with more than 75 connections generally remained below $96,000. This pattern reflects the fixed-cost structure of consolidation projects. Pipeline construction, environmental review (CEQA), engineering, and administration represent substantial base costs that are largely independent of the number of connections served. The three-mile pipeline costs are the same whether it connects 3 households or 300. National cost function analysis confirms significant economies of scale for community water systems, with unit costs declining as system size increases[18]. Fiscally weaker small districts also face higher borrowing costs and more constrained capital programs, amplifying the per-connection cost burden[19]. Similarly, performance assessments of North Carolina community water systems revealed that small systems with limited technical, managerial, and financial capacity face systematic disadvantages in maintaining infrastructure[20], which is consistent with the elevated per-connection costs observed here. Infrastructure investment disparities are well documented nationally, with small systems and communities of color receiving disproportionately less capital investment per connection[2].

The split threshold described earlier addresses this directly: evaluating small systems on total project cost rather than per-connection cost prevents otherwise feasible projects from being screened out by high per-connection metrics. In this analysis, 113 of 114 pairs passed the applicable threshold, and the only pair that did not pass exceeded the total cost cap by less than 1%.

Route consolidations were consistently more expensive than intersect consolidations, with median costs approximately five times higher ($5.3 million versus $1.1 million). This difference is driven almost entirely by pipeline distance: intersect pairs assume 1,000 feet of pipeline, whereas route pairs require road–network distances that can extend to three miles (approximately 15,840 feet) plus a 1,000-foot buffer. Intersect opportunities therefore likely represent the lowest-cost consolidation pathway among the costs modeled here. Cost is one consideration in prioritization, but not the only one. Funding decisions also weigh the severity of need, including active health-based violations, the presence of disadvantaged communities, and the duration systems have relied on interim supplies. A low-cost connection for a system without recent violations may warrant lower priority than a costlier connection for a low-income system facing ongoing exposure.

ARTICLE IN PRESS

Limitations

The cost estimates produced by this analysis are screening-level approximations based on standardized unit cost assumptions. They do not account for site-specific conditions, including topography, soil conditions, existing utility conflicts, right-of-way acquisition, or hydraulic capacity constraints. Actual project costs may differ substantially from modeled estimates. The State Water Board's methodology documentation notes this limitation and emphasizes that the cost assessment is designed to estimate statewide funding needs rather than to prescribe specific community solutions [12]. A unit cost analysis of U.S. water infrastructure projects revealed significant variability in delivered costs depending on the project delivery method and scale[21], suggesting that standardized unit cost assumptions may under- or overestimate costs for individual projects.

The analysis evaluates physical consolidation only. Systems that do not meet the distance or cost viability criteria may still be candidates for managerial consolidation, centralized treatment, decentralized treatment, or other long-term solutions modeled in subsequent stages of the state's cost assessment. The total capital need estimated here ($413.5 million) represents only the physical consolidation component of the broader investment required to achieve safe drinking water access in the study region. Hybrid configurations that combine decentralized infrastructure with centralized governance and shared operations may offer practical solutions for isolated systems[22].

The classification criteria exclude some systems that the state may choose to consolidate through its mandatory authority. As observed with Cutler PUD in Tulare County, systems exceeding the connection thresholds for joining eligibility (by a few thousand in our analysis) can still be mandated to consolidate on the basis of site-specific evaluation. The screening thresholds should be understood as defining the scope of the modeled analysis rather than the full universe of possible consolidations.

The intersect classification identifies any spatial overlap between the joining and receiving system service area boundaries, regardless of the size or location of the overlapping area. In one case in Fresno County (Las Deltas Mutual Water Company), the boundary overlap involved a small, fragmented portion of the service area distant from the main population center, where a pipeline connection would still be required in practice. The fixed 1,000-foot pipeline assumption for intersecting pairs partially accounts for such cases, but the actual infrastructure need may be closer to route

ARTICLE IN PRESS

consolidation for fragmented systems. This was the only instance observed across the 45 intersection pairs in this analysis. Therefore, intersection classifications for systems with fragmented or noncontiguous service area boundaries should be reviewed on a case-by-case basis.

The results highlight both the opportunity and the limits of physical consolidation as a long-term solution. Physical consolidation is a necessary but insufficient component of a comprehensive strategy for achieving safe drinking water access in the San Joaquin Valley. Future work should extend this analysis statewide to quantify the physical consolidation potential across California's 58 counties. A complementary equity analysis examining which disadvantaged communities have viable consolidation options but have not yet been consolidated would further inform funding prioritization. Evaluating the sensitivity of the results to alternative distance thresholds, cost assumptions, and receiver selection criteria would strengthen the evidence base for policy decisions surrounding water system consolidation.

Methods

Study Area and Data Sources

This study examines physical consolidation opportunities among public water systems in four counties in California's southern San Joaquin Valley: Kern, Kings, Tulare, and Fresno (Figure 3). These counties share interconnected groundwater basins, overlap in agricultural land use and contamination sources, and collectively represent the region with the highest concentration of failing water systems in the state. The four counties contain water systems that range from large municipal providers serving tens of thousands of connections to small community systems serving mobile home parks, labor camps, and unincorporated communities. The landscape includes intensively irrigated agricultural land on the valley floor, arid foothills along the Sierra Nevada and Coast Range margins, and mountain communities along the Kern River, presenting a range of consolidation scenarios from dense urban-adjacent pairings to isolated rural systems separated by miles of farmland.

The analysis integrates three primary data sources (Table 4). System-level risk classifications and attributes were obtained from SAFER risk assessment dataset. This dataset provides each public water system's SAFER status (Failing, At-Risk, Potentially At-Risk, Not Assessed, or Not At-Risk), service connections, population served, economic status designation, and geographic coordinates. The service area boundaries were obtained

ARTICLE IN PRESS

from the California Drinking Water System Area Boundaries layer maintained by the State Water Board. Eleven systems present in the SAFER dataset but absent from the boundary layer (6 in Kern, 4 in Tulare, 1 in Fresno) were assigned estimated service areas by buffering their SAFER-reported coordinates by 0.5 miles. Of these 11 systems, 6 entered consolidation pairs as joining systems and none served as receiving systems. All 6 were matched through boundary intersection rather than road-network routing, and each pair was financially viable with a cost per connection within the range observed for systems with mapped boundaries.

Road network data were obtained from OpenStreetMap (OSM) via the Overpass API and processed locally via the dodgr package in R [23,24]. All spatial data were projected to the California Albers equal-area coordinate system (EPSG:3310) for distance calculations.

System Classification

Each water system was classified as a potential joining system, a potential receiving system, or neither, following the criteria established in the State Water Board's 2024 Physical Consolidation Methodology (Table 1). Joining systems are smaller systems that may benefit from consolidation into a larger provider. Receiving systems are larger systems with the capacity to absorb joining systems and serve their customers.

Classification depends on two attributes: the system's SAFER status and its number of service connections. The joining-system thresholds are 1,000 or fewer connections for failing systems and 500 or fewer for at-risk systems. For receiving systems, the connection threshold varies by SAFER status: failing systems qualify only above 1,000 connections, while all other categories, including At-Risk, Potentially At-Risk, Not At-Risk, Not Assessed, and systems with no SAFER data, qualify above 500 connections (Table 1). The classification permits systems without a SAFER status to serve as receivers where sufficiently large, though none did so in this study region.

Classification did not restrict systems by federal classification type. Any water system with a Failing or At-Risk SAFER status meeting the connection thresholds qualified as a potential joining system, including non-community systems such as schools, since these face the same exposure to unsafe drinking water that consolidation is intended to address. Systems that did not meet the criteria for either category were excluded from the consolidation analysis. A system cannot be classified as both a joining

ARTICLE IN PRESS

system and a receiving system simultaneously, as the connection thresholds are mutually exclusive within each SAFER status category.

This classification step produces a pool of eligible candidates. The actual pairing of joining systems to specific receiving systems occurs in proximity analysis, where each joining system is matched to the largest eligible receiver within the proximity threshold, established by boundary intersection or road-network distance.

Proximity Analysis

The proximity analysis identifies geographically plausible consolidation pairs through two stages: boundary intersection for systems with overlapping service areas and road network routing for systems that require new connecting infrastructure.

Boundary Intersection

The joining systems whose service area boundaries spatially overlap with the boundary of at least one eligible receiving system represent the most straightforward consolidation candidates. The boundary intersection was evaluated via the full polygon geometry of each system's service area. For each joining system with one or more intersecting receivers, the receiver with the largest number of service connections was selected as the consolidation partner. Joining systems identified through this stage were excluded from the subsequent route analysis to avoid duplicate pairing.

Road Network Routing

Combining systems that do not share a boundary with any eligible receiver requires new connecting infrastructure, making the road network distance the relevant measure of proximity. A routable road network was constructed from OpenStreetMap (osm) data by excluding segments classified as construction, proposed, footway, path, cycleway, bridleway, steps, or pedestrian, retaining only segments accessible to motor vehicles. The filtered network was converted to a weighted directed graph via the dodgr package with a motorcar weighting profile [24].

This substitutes for the State Water Board methodology, which uses ESRI StreetMap and ArcGIS Network Analyst for routing. The two road networks differ in coverage, topology, and classification, which may produce different route distances for some pairs. OpenStreetMap data are community-contributed and may have gaps in rural areas, though coverage in California's Central Valley is comprehensive. A comparison of road-network

ARTICLE IN PRESS

datasets in New Mexico found that OpenStreetMap-derived travel times closely approximate those from ESRI StreetMap, with larger discrepancies concentrated in less urbanized areas [25]. A national assessment found higher OpenStreetMap completeness and positional accuracy in metropolitan counties and lower coverage in nonmetropolitan areas[26], although the San Joaquin Valley's extensive agricultural road network is well represented.

The true mathematical centroid of each joining system's service area polygon serves as the routing origin. In cases where an irregularly shaped polygon causes the centroid to fall outside the service area boundary, the routing engine snaps the point to the nearest node on the road network. The proportion of centroids falling outside their service area was recorded as a quality assurance measure.

The route destination for each joining-receiving pair was determined by intersecting the road network with the outer boundary of the receiving system's service area. This intersection produces a set of road entry points representing locations where physical roads cross into the receiving system. For each candidate pair, the road entry point with the shortest road network distance to the joining centroid was selected as the destination. This ensures that route distances reflect actual points of infrastructure access rather than arbitrary perimeter coordinates.

A preliminary screen identified all joining-receiving pairs within a 10-mile Euclidean distance to reduce the computational burden. For each candidate pair, the shortest road network distance was calculated from the joining centroid to the selected road entry point. Pairs with a road distance of three miles or less were retained. Among the retained candidates, each joining system was matched to the receiving system with the largest number of service connections, prioritizing the receiver with the greatest capacity to absorb and sustainably serve the joining system rather than the one requiring the shortest connection.

Cost Estimation

Capital costs for each consolidation pair were estimated following the cost methodology published with the 2024 Drinking Water Needs Assessment. The estimate covers the infrastructure and administrative components required to physically connect a joining system to a receiving system. Table 5 summarizes all unit cost assumptions.

The largest cost driver is pipeline construction, which is calculated as the product of the pipeline length and the per-linear-foot unit cost. For

ARTICLE IN PRESS

intersecting pairs, a fixed pipeline distance of 1,000 feet was assumed. For route pairs, the pipeline distance equals the road network distance converted to feet plus a 1,000-foot buffer. This buffer follows the State Water Board cost methodology, which adds a fixed 1,000 feet to the GIS-calculated route distance to account for limitations in the GIS distance and route analysis. Pipeline and service line costs are subject to a regional adjustment factor that reflects geographic variation in construction costs across California. The State Water Board methodology applies this factor by county type, increasing costs in suburban (+30%) and urban (+32%) counties relative to a rural baseline. All four study counties are classified as rural and therefore receive a 0% adjustment, so the base unit costs apply without regional inflation.

For systems with missing or zero connection counts, a surrogate value was derived by dividing the reported population by 3.3 persons per connection and rounding it up. This surrogate was used in place of reported connections for all connection-driven cost components. This surrogate applied to 344 systems across the four counties (83 in Kern, 1 in Kings, 83 in Tulare, and 177 in Fresno). None of these systems entered a consolidation pair, as joining or receiving systems, so the surrogate did not affect any reported cost estimate.

The subtotal of all the components was multiplied by an additive adjustment factor of 1.481, comprising contingency, planning, engineering, and inflation allowances, as detailed in Table 5. A cost-per-connection metric was calculated for each pair by dividing the total capital cost by the effective number of service connections.

ARTICLE IN PRESS

Funding Viability Screening

Each consolidation pair was evaluated against the funding viability thresholds established in the State Water Board's 2024 Cost Assessment Model. These thresholds determine whether a consolidation project falls within the range of costs that existing funding programs can support.

The screening applies different criteria depending on the system size. For joining systems with fewer than 75 service connections, the total capital cost must not exceed $7.2 million. For joining systems with 75 or more connections, the cost per connection must not exceed $96,000. Table 6 presents these thresholds.

All spatial operations were performed in R version 4.4.2 [27] via the sf package for geometry operations, dodgr for road network construction and

ARTICLE IN PRESS

shortest-path routing, and osmdata for OpenStreetMap data retrieval [23]. Data manipulation and analysis were performed via the tidyverse suite of packages [28] in R, including dplyr for data wrangling and ggplot2 for static figures [29,30]. Interactive maps were produced via leaflet, and static publication figures were produced via tmap and ggspatial [31–33].

Data availability statement

The following data, models, or code generated or used during the study are available in a repository online. All analysis code is available at [Zenodo DOI]. The SAFER Risk Assessment dataset and the 2025 Drinking Water Needs Assessment report are archived in the repository at https://github.com/laljeet/CA-Physical-Water-System-Consolidation/tree/main/safer_archive. Service area boundary data are available from the California Drinking Water System Area Boundaries dataset https://github.com/laljeet/CA-Physical-Water-System-Consolidation/blob/main/data_in/water_systems_ca.gpkg . OpenStreetMap road network data are available at https://www.openstreetmap.org. County boundaries were obtained from the U.S. Census Bureau TIGER/Line program via the R tigris package. Interactive maps of the consolidation results for each study county are available at https://laljeet.github.io/CA-Physical-Water-System-Consolidation/.

Code availability

All analysis code is available at https://github.com/laljeet/CA-Physical-Water-System-Consolidation and archived at Zenodo (https://doi.org/10.5281/zenodo.19156569). The analysis was performed in R version 4.4.2.

Acknowledgments

No external funding supported this work.

Author contributions

L.S. designed the study, performed the analysis, and wrote the manuscript.

Competing interests

The author declares no competing interests.

References

ARTICLE IN PRESS

1. SWRCB. *2025 Drinking Water Needs Assessment*. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/needs.html (2025).
2. Allaire, M., Wu, H. & Lall, U. National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences* **115**, 2078–2083 (2018).
3. Pierce, G. & González, S. R. Publicly available data reveal systemic vulnerability among California mobile home park water systems. *Environmental Justice* **10**, 168–173 (2017).
4. Balazs, C. L. & Ray, I. The Drinking Water Disparities Framework: On the Origins and Persistence of Inequities in Exposure. *American Journal of Public Health* **104**, 603–611 (2014).
5. Dobbin, K. B., McBride, J. & Pierce, G. Panacea or Placebo? The Diverse Pathways and Implications of Drinking Water System Consolidation. *Water Resources Research* **59**, e2023WR035179 (2023).
6. Dobbin, K. B., McBride, J. & Pierce, G. Recent Trends in Water System Consolidation: Lessons from California. *Journal AWWA* **117**, 10–17 (2025).
7. Dobbin, K. B. & Fencl, A. L. Institutional diversity and safe drinking water provision in the United States. *Utilities Policy* **73**, 101306 (2021).
8. Egge, M. & Ajibade, I. A community of fear: emotion and the hydro-social cycle in East Porterville, California. *Journal of Political Ecology* **28**, (2021).

ARTICLE IN PRESS

ARTICLE IN PRESS

9. Egge, M. & Ajibade, I. Water struggles and contested use: A capabilities assessment of household water security in marginalized communities. *Journal of Environmental Management* **341**, 118047 (2023).

10. Norriss, J., Cunningham, M., DeRosa, A. R. & Vedachalam, S. Too Small to Succeed: State-Level Consolidation of Water Systems. *Journal AWWA* **113**, 8–15 (2021).

11. Sowby, R. B. & Price, R. B. Proximity of U.S. community water systems: implications for regional collaboration. *Utilities Policy* **97**, 102079 (2025).

12. SWRCB. *Supplemental Appendix: Physical Consolidation Cost Estimate Methodology*. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/documents/needs/2024/2024costassessment-physical-consolidation.pdf (2024).

13. Burow, K. R., Floyd, W. D. & Landon, M. K. Factors affecting 1,2,3-trichloropropane contamination in groundwater in California. *Science of the Total Environment* **672**, 324–334 (2019).

14. Ormsbee, L., Byrne, G., McNeil, D., Yost, S. & Garner, E. The Challenges of Water Infrastructure in the Central Appalachia Region of the United States. *Journal of Water Resources Planning and Management* **152**, 05025017 (2026).

15. HDR. *Regionalization Strategies to Assist Small Water Systems in Meeting New SDWA Requirements (Study 3)*.

ARTICLE IN PRESS

ARTICLE IN PRESS

https://www.twdb.texas.gov/publications/reports/contracted_reports/doc/0704830692_RegionG/Brazos%20G%20Study%203%20Report.pdf (2009).

16. Milman, A., Ashjian James, O. & Macuch, C. Motivating the formation of partnerships by small water systems. *Utilities Policy* **85**, 101680 (2023).
17. Mullin, M., Vedachalam, S., Kerins, R. & Teodoro, M. P. The effects of drinking water service fragmentation on drought-related water security. *Science* **368**, 274–277 (2020).
18. Shih, J.-S., Harrington, W., Pizer, W. A. & Gillingham, K. Economies of scale in community water systems. *Journal AWWA* **98**, 100–108 (2006).
19. Scott, T. A., Moldogaziev, T. & Greer, R. A. Drink what you can pay for: Financing infrastructure in a fragmented water system. *Urban Studies* **55**, 2821–2837 (2018).
20. Bell, E. V., Hansen, K. & Mullin, M. Assessing performance and capacity of US drinking water systems. *Journal of Water Resources Planning and Management* **149**, 04022073 (2023).
21. Feghaly, J., El Asmar, M. & Ariaratnam, S. T. Exploring Unit Cost Performance for Water Infrastructure Delivery. *Practice Periodical on Structural Design and Construction* **28**, 04022060 (2023).
22. Stoler, J., Thomson, B. M. & Adams, E. A. (De)centralized water futures: Key dimensions of infrastructure, governance, and operations and maintenance. *WIREs Water* **12**, e70035 (2025).

ARTICLE IN PRESS

ARTICLE IN PRESS

23. Fellows, I. & Stotz, J.-P. OpenStreetMap: Access to Open Street Map Raster Images. (2025).

24. Padgham, M. *et al.* dodgr: Distances on Directed Graphs. (2025).

25. Lin, Y., Lippitt, C., Beene, D. & Hoover, J. Impact of travel time uncertainties on modeling of spatial accessibility: a comparison of street data sources. *Cartography and Geographic Information Science* **48**, 471–490 (2021).

26. Li, D. *et al.* Mapping the completeness and positional accuracy of OpenStreetMap road data at the county level in the contiguous United States. *Transactions in GIS* **29**, e70077 (2025).

27. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2024).

28. Wickham, H. & RStudio. tidyverse: Easily Install and Load the 'Tidyverse'. (2023).

29. Wickham, H. *et al.* ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. (2026).

30. Wickham, H. *et al.* dplyr: A Grammar of Data Manipulation. (2026).

31. Cheng, J. *et al.* leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. (2025).

32. Dunnington, D., Thorne, B. & Hernangómez, D. ggspatial: Spatial Data Framework for ggplot2. (2025).

33. Tennekes, M. *et al.* tmap: Thematic Maps. (2025).

Tables

ARTICLE IN PRESS

Table 1 Classification criteria for joining and receiving systems, based on SWRCB 2024 Physical Consolidation Methodology.

SAFER Status	Joining System	Receiving System
Failing	≤ 1,000 connections	> 1,000 connections
At-Risk	≤ 500 connections	> 500 connections
Potentially At-Risk	Excluded	> 500 connections
Not At-Risk	Excluded	> 500 connections
Not Assessed	Excluded	> 500 connections
No SAFER data	Excluded	> 500 connections

Table 2 System classification, consolidation pairs, and funding viability results by county Unmatched joining systems include Failing and At-Risk systems with no eligible receiver within the intersect or 3-mile route threshold.

	Kern	Kings	Tulare	Fresno	Total
System Classification					
Total water systems	266	19	200	309	**794**
Joining systems	93	3	56	58	**210**
Receiving systems	34	6	14	20	**74**
Proximity Analysis					
Intersect pairs	29	1	5	10	**45**
Route pairs	35	1	17	16	**69**
Total consolidation pairs	64	2	22	26	**114**
Unmatched joining systems	29	1	34	32	**96**
Funding Viability					
Fundable pairs	63	2	22	26	**113**
Not fundable pairs	1	0	0	0	**1**
Total capital need ($ millions)	$222.0	$15.2	$97.3	$79.0	**$413.5**

ARTICLE IN PRESS

Table 3 Median estimated capital costs by county and consolidation type. Cost per connection reflects total capital cost divided by effective service connections in the joining system.

County	Type	n	Median Total Cost	Total Cost IQR	Median Cost per Connection	Cost/Connection IQR
Kern	Intersect	29	$1,007,293	$709,506–$1,435,363	$31,478	$26,098–$44,344
	Route	35	$5,515,449	$3,552,160–$6,768,512	$97,443	$47,203–$249,487
Kings	Intersect	1	$9,457,017	_	$19,459	_
	Route	1	$5,791,539	_	$1,930,513	_
Tulare	Intersect	5	$1,342,304	$1,081,740–$2,235,667	$26,846	$22,813–$30,048
	Route	17	$5,819,145	$2,593,262–$6,093,134	$109,409	$49,370–$541,255
Fresno	Intersect	10	$1,053,823	$676,935–$1,835,515	$30,985	$23,994–$48,083

ARTICLE IN PRESS

	Route	16	$4,321,678	$2,582,650–$5,501,567	$339,466	$149,588–$1,754,449

Table 4 Summary of Data sources and record counts used in this analysis

Data Source	**Provider**	**Records**
SAFER Risk Assessment	SWRCB	3,206 systems statewide
Service Area Boundaries (SABL)	SWRCB	5,050 polygons
OpenStreetMap highway network	OSM Contributors	498,157 road segments (four study counties)
County boundaries	U.S. Census Bureau (TIGER/Line)	58 counties

Table 5 Unit cost assumptions for physical consolidation capital cost estimation (2024 dollars).

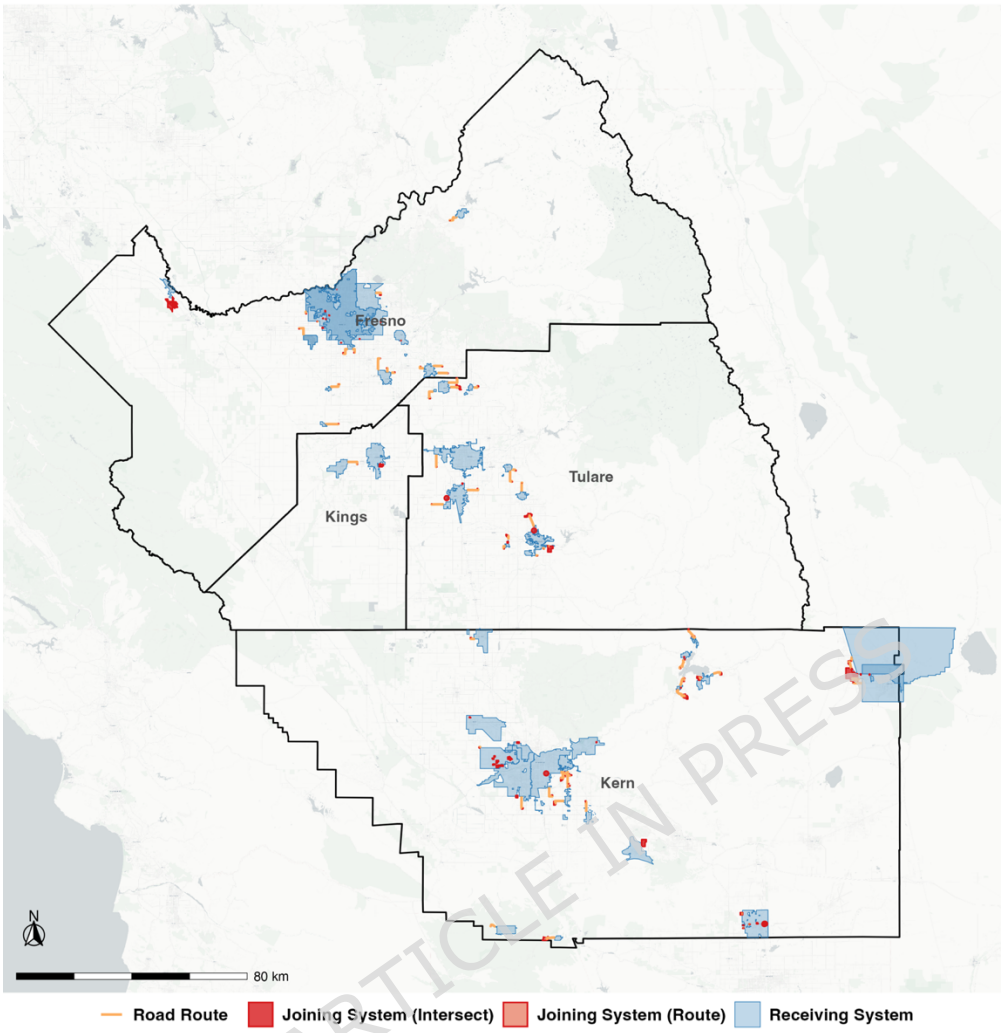
Component	**Value**	**Unit**
Pipeline construction	$220	per linear foot
Service line	$6,200	per connection
Connection fee	$5,437	per connection
Administration	15%	of pipeline + service line cost
CEQA (intersect)	$25,000	per project
CEQA (route)	$100,000	per project
Contingency	20%	of subtotal
Planning/construction	10%	of subtotal
Engineering	15%	of subtotal
Inflation	3.1%	of subtotal

ARTICLE IN PRESS

Table 6 Funding viability thresholds for Cost Assessment Model.

System Size	**Viability Metric**	**Threshold**
< 75 service connections	Total capital cost	≤ $7,200,000
≥ 75 service connections	Cost per connection	≤ $96,000

ARTICLE IN PRESS



ARTICLE IN PRESS

Figure 1 Spatial distribution of consolidation pairs across the four study counties. Maps created by the authors in R version 4.4.2 (R Core Team, 2024; https://www.R-project.org/) using the sf, tmap, and ggspatial packages. Road network data available under the Open Database License (https://www.openstreetmap.org/copyright). Additional spatial data sources are listed in the Data Availability statement.

ARTICLE IN PRESS

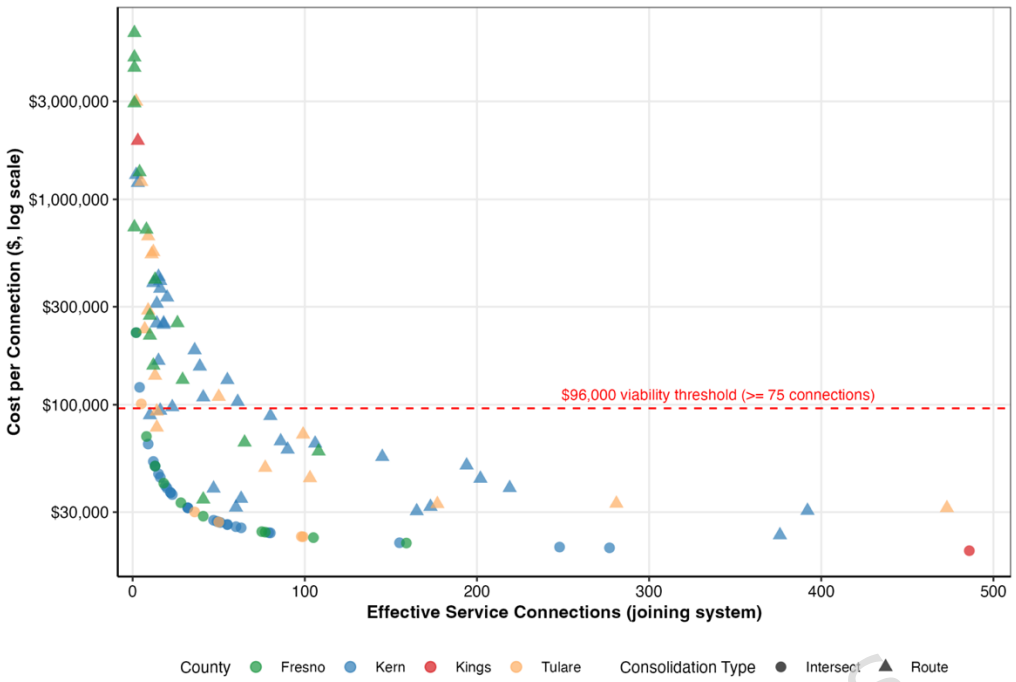


Figure 2 Cost per connection as a function of joining system size (effective service connections) for all consolidation pairs across four counties

ARTICLE IN PRESS

ARTICLE IN PRESS

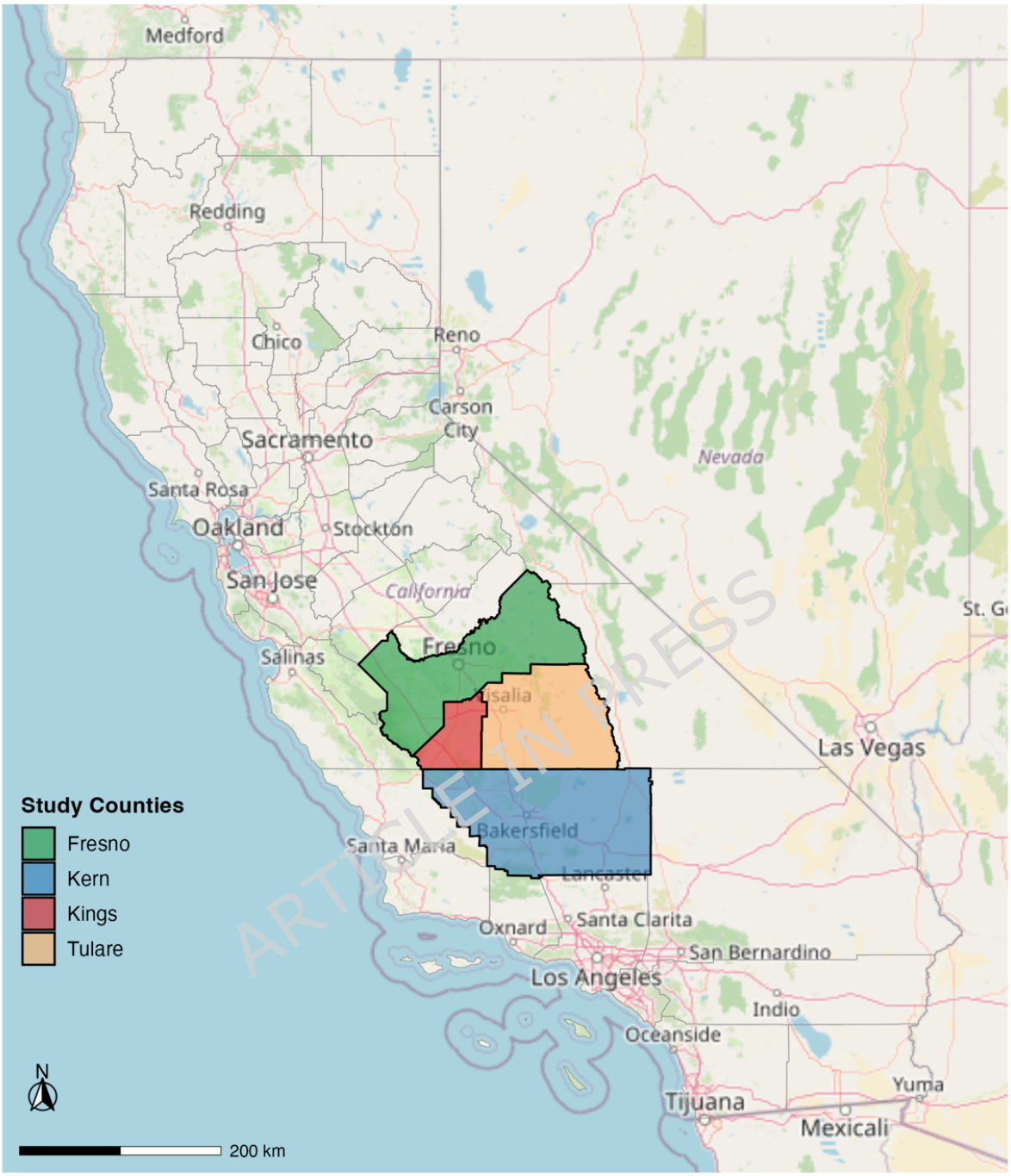


Figure 3 Study area comprising four San Joaquin Valley counties (Kern, Kings, Tulare, and Fresno) within California. Maps created by the authors in R version 4.4.2 (R Core Team, 2024; https://www.R-project.org/) using the sf, tmap, and ggspatial packages. Spatial data sources are listed in the Data Availability statement.

ARTICLE IN PRESS

Tables

Table 1 Classification criteria for joining and receiving systems, based on SWRCB 2024 Physical Consolidation Methodology.

SAFER Status	Joining System	Receiving System
Failing	≤ 1,000 connections	> 1,000 connections
At-Risk	≤ 500 connections	> 500 connections
Potentially At-Risk	Excluded	> 500 connections
Not At-Risk	Excluded	> 500 connections
Not Assessed	Excluded	> 500 connections
No SAFER data	Excluded	> 500 connections

Table 2 System classification, consolidation pairs, and funding viability results by county Unmatched joining systems include Failing and At-Risk systems with no eligible receiver within the intersect or 3-mile route threshold.

	Kern	Kings	Tulare	Fresno	Total
System Classification					
Total water systems	266	19	200	309	**794**
Joining systems	93	3	56	58	**210**
Receiving systems	34	6	14	20	**74**
Proximity Analysis					
Intersect pairs	29	1	5	10	**45**
Route pairs	35	1	17	16	**69**
Total consolidation pairs	64	2	22	26	**114**
Unmatched joining systems	29	1	34	32	**96**
Funding Viability					
Fundable pairs	63	2	22	26	**113**
Not fundable pairs	1	0	0	0	**1**
Total capital need ($ millions)	$222.0	$15.2	$97.3	$79.0	**$413.5**

Table 3 Median estimated capital costs by county and consolidation type. Cost per connection reflects total capital cost divided by effective service connections in the joining system.

County	Type	n	Median Total Cost	Total Cost IQR	Median Cost per Connection	Cost/Connection IQR
Kern	Intersect	29	$1,007,293	$709,506–$1,435,363	$31,478	$26,098–$44,344
	Route	35	$5,515,449	$3,552,160–$6,768,512	$97,443	$47,203–$249,487
Kings	Intersect	1	$9,457,017	_	$19,459	_
	Route	1	$5,791,539	_	$1,930,513	_
Tulare	Intersect	5	$1,342,304	$1,081,740–$2,235,667	$26,846	$22,813–$30,048
	Route	17	$5,819,145	$2,593,262–$6,093,134	$109,409	$49,370–$541,255

ARTICLE IN PRESS

Fresno	Intersect	10	$1,053,823	$676,935–$1,835,515	$30,985	$23,994–$48,083
	Route	16	$4,321,678	$2,582,650–$5,501,567	$339,466	$149,588–$1,754,449

Table 4 Summary of Data sources and record counts used in this analysis

Data Source	**Provider**	**Records**
SAFER Risk Assessment	SWRCB	3,206 systems statewide
Service Area Boundaries (SABL)	SWRCB	5,050 polygons
OpenStreetMap highway network	OSM Contributors	498,157 road segments (four study counties)
County boundaries	U.S. Census Bureau (TIGER/Line)	58 counties

Table 5 Unit cost assumptions for physical consolidation capital cost estimation (2024 dollars).

Component	**Value**	**Unit**
Pipeline construction	$220	per linear foot
Service line	$6,200	per connection
Connection fee	$5,437	per connection
Administration	15%	of pipeline + service line cost
CEQA (intersect)	$25,000	per project
CEQA (route)	$100,000	per project
Contingency	20%	of subtotal
Planning/construction	10%	of subtotal
Engineering	15%	of subtotal
Inflation	3.1%	of subtotal

ARTICLE IN PRESS

Table 6 Funding viability thresholds for Cost Assessment Model.

System Size	**Viability Metric**	**Threshold**
< 75 service connections	Total capital cost	≤ $7,200,000
≥ 75 service connections	Cost per connection	≤ $96,000