Supporting Information for "Discovering Safe Operating Spaces in Deeply Uncertain Pathways: Evaluating How Implementation Uncertainties Shape Cooperative Multi-City Water Supply Portfolio Management and Investment"

Lillian B. Lau[1], Patrick M. Reed[1], David F. Gold[1]

[1]Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, 14850 USA

Contents of this file

Introduction

This document contains supporting information that specifies, in further detail, the methods of objectives formulation, synthetic streamflow generation, ROF table generation, bootstrapping, clustering, as well as robustness quantification and evaluation. Figures further supporting the results and discussion section of the main text can also be found in this document.

Corresponding author: L.B. Lau (lbl59@cornell.edu)

S1 The Sedento Valley

As a region, the Sedento Valley serves 1.5 million residents. Each city is served by three independent water utilities that supply drinking water to their respective residents. Watertown owns and operates a water treatment plant that draws water from Lake Michael, which in turn is managed by a federal agency. It also has access to supply resources from College Rock Reservoir, where it owns and operates another water treatment facility. Fallsland and Dryville share resources from Autumn Lake. They each access their allocated volumes of water from Autumn Lake via independent water treatment plants. The disproportionate allocation of water supply is apparent: Watertown has direct access to the largest unallocated portion of Lake Michael although it has the smallest population. In contrast, Fallsland and Dryville do not have proportionally sized access to additional allocation from Lake Michael. While all three utilities may request additional allocation from the federal agency, Fallsland and Dryville have to purchase water via treated transfers from Watertown, as the latter operates the only treatment facility on Lake Michael. This has motivated significant regional investments in large transfer interconnections for Fallsland and Dryville to access their respective allocations without capacity constraints.

S2 Objective Functions

Text S2 contains details on the formulation of the objective functions for the Sedento Valley regional test case as first formulated by (Trindade et al., 2020).

1. **Reliability** (f_{REL}): The reliability objective is to be maximized. It is calculated as the fraction of states of world in which the total storage of a utility drops below 20% of its maximum capacity in any given week for each utility:

$$\text{maximize } f_{REL} = \min_i \left[\min_j \left(\frac{1}{N_r} \sum_{i=1}^{N_r} g_{i,j}^y \right) \right] \tag{1}$$

where

$$g_{i,j}^y = \begin{cases} 0 \ \forall w : \dfrac{x_{s,i,j}^{w,y}}{C_j} \geq S_c \\ 1 \ \text{ otherwise} \end{cases} \tag{2}$$

where $g_{i,j}^y = 0$ when there is a week w for a given simulation year y in a particular realization i such that the total storage S of utility j falls below a value of S_c where $S_c = 20\%S$. Otherwise, $g_{i,j}^y = 1$. Here, N_r is the total number of realizations in one function evaluation such that $i \in N_r$.

2. **Restriction frequency** (f_{RF}): The restriction frequency objective is to be minimized. It is calculated as the fraction of years across all realizations in which water use restrictions were triggered in at least one week:

$$\text{minimize } f_{RF} = \max_j \left[\frac{1}{N_r \cdot N_{ys}} \sum_{i=1}^{N_r} \sum_{y=1}^{N_{ys}} h_{i,j}^y \right] \tag{3}$$

where

$$h_{i,j}^y = \begin{cases} 0 \ \forall w : x_{srof,i,j}^{w,y} \leq \theta_{rt,j} \\ 1 \ \text{ otherwise} \end{cases} \tag{4}$$

where $h^y_{i,j} = 0$ if there was a week w in a given year y in a given realization in which water use restrictions were triggered across all years within the planning horizon N_{ys}. $h^y_{i,j} = 1$ otherwise.

3. **Infrastructure net present cost** (f_{INPC}): The infrastructure net present cost objective is to be minimized. It is measured as the average net present cost of all new infrastructure built across all realizations:

$$\text{minimize } f_{INPC} = \frac{1}{N_r} \sum_{i=1}^{N_r} \sum_{y=1}^{BM} \frac{PMT}{(1+d)^y} \tag{5}$$

where BM is the bond term and d is the discount rate of 5% for a given year y of debt service payment PMT since the bond was issued. PMT is calculated as follows and assumes a level debt service bond:

$$PMT = \frac{P[BR(1+BR)^{BM}]}{(1+BR)^{BM} - 1} \tag{6}$$

where P is the principal (i.e., the construction cost), BR is the interest rate to be paid to the debtor (i.e., banks, shareholders) throughout the duration of the bond term BT. All payments are discounted to their present value.

4. **Peak financial cost** (f_{PFC}): The peak financial cost objective is to be minimized. This objective is measured as the expected annual cost of debt including all non-infrastructure assets used to mitigate drought over the planning horizon. These costs are considered revenue losses from water-use restrictions, treated transfer purchases, contingency fund contributions, third-party insurance contract payments, and debt repayment:

$$\text{minimize } f_{PFC} = \max_j \left[\frac{1}{N_r \cdot N_y s} \sum_{i=1}^{N_r} \sum_{y=1}^{N_{ys}} SYPC_{i,j}y \right] \tag{7}$$

where

$$SYPC_{i,j}y = \frac{\sum_{c \in C_j} PMT_{i,j,c} + (\theta_{acfc,j} \cdot ATR^y_{i,j}) + IP^y_{i,j} + ATR_j}{ATR^y_{i,j}} \tag{8}$$

where the $SYPC$ is the single-year peak financial cost, IP is the cost of insurance payments for year y, $PMT_{i,j,c}$ is the debt (bond) payments for infrastructure option $c \in C_j$ where C_j is the set of infrastructure options to be built by utility j in realization i, $\theta_{acfc,j}$ is the annual contingency fund contributions for utility j and ATR is the total annual volumetric revenue. All variables are measured in USD$.

5. **Worst-case first-percentile cost** (f_{WCC}): The worst-case first-percentile cost objective is to be minimized. This objective represents the top 1% single-year drought mitigation costs observed across all SOWs over the planning horizon. It is measured as follows:

$$\text{minimize } f_{WCC} = \max_{j}$$

where

$$SYWC^{y}_{i,j} = \frac{max(RL^{y}_{i,j} + TC^{y}_{i,j} - (\theta_{acfc,j} \cdot ATR^{y}_{i,j} - YIPO^{y}_{i,j}, 0)}{ATR^{y}_{i,j}} \tag{10}$$

where $SYWC$ is the single-year worst-case first percentile cost, IP is the cost of insurance payments for year y, RL is the revenue losses from water-use restrictions, TC are the costs of treated transfer purchases, $YIPO$ is the total cost of insurance payouts over year y, CF is the total available contingency funds, ATR is the total annual volumetric revenue and $\theta_{acfc,j}$ is the annual contingency fund contributions for utility j. All variables are measured in USD$.

S3 Decision Value Ranges

Table S1 to Table S3 show the decision variable ranges used to generate potential stakeholder risk and action values.

S4 Calculating Risk of Failure

This study uses two types of ROF action triggers: short-term ROF triggers ($sROF$) that trigger short-term drought mitigation actions (Caldwell & Characklis, 2014), and long-term ROF triggers ($lROF$) that trigger the new candidate infrastructure investments (Zeff et al., 2016) in a utility's infrastructure pathways. At any given week, a utility's $sROF$ represents the probability that its reservoir storage will drop below 20% of its total capacity at any point during a moving window of the subsequent 52 weeks through the full 20 simulated years, across 500 hydro-climatic realizations. Each drought mitigation instrument (i.e., restrictions, transfers, insurance payments) is assigned an associated $sROF$. The instrument is triggered if the $sROF$ exceeds a risk threshold found to be Pareto-approximately optimal during DU Optimization.

The $lROF$ captures a utility's capacity-to-demand ratio and is calculated on an annual basis. Moreover, it measures the probability that it's reservoir storage will drop below 20% of its total capacity at any point during a moving window of the next 78 weeks, or a year and a half, over a period of 20 years across 500 hydro-climatic realizations. This study assumes that all reservoirs begin at 100% capacity. Each utility has an assigned $lROF$ trigger, and an associated ranking of infrastructure options included in the ICO. The ICO is a decision variable vector of infrastructure construction prioritized ordering for each utility, and $\boldsymbol{I_{ME}}$, the infrastructure options within the set of already-built or potential infrastructure, $\boldsymbol{I_B}$. For elements within $\boldsymbol{I_{ME}}$, it is enforced that two infrastructure options cannot be built in the same realization.

The calculation for the ROF metric is derived from (Trindade et al., 2019) and is computed as follows:

$$x^{w}_{rof,j} = \frac{1}{N_{rof}} \sum_{y'=0}^{w} f^{w}_{y',j}(\boldsymbol{NI^{y'}}, \boldsymbol{E^{y'}}) \tag{11}$$

where

$$f^{w}_{y',j} = \begin{cases} 0 \ \forall w' \in (y', w), ..., (y', w + T_{rof}) \ \dfrac{x^{y,w'}_{s',j}}{C_j} \geq S_c \\ 1 \ \text{ otherwise} \end{cases} \tag{12}$$

and

$$x^{y,w'}_{s',j} = f\left(C_j, \boldsymbol{UD^w_j}, \boldsymbol{NI^{y',w'}_j}, \boldsymbol{E^{y',w'}_j}, \boldsymbol{W^{y',w'}_j} | \Psi_s\right) \tag{13}$$

In Equations 11 to 13, w' and y' indicate a week and year simulated using 52 weeks' of past hydro-climatic data. Next, $x^{w}_{rof,j}$ is the ROF for utility j in week w, otherwise known as the probability of failure remaining under a utility-specified failure threshold. Next, the variable $f^{w}_{y',j}$ is a binary variable where 0 indicates that the ROF threshold has been crossed (failure) and 1 indicates otherwise. It is conditional upon the volume of the combined storage $x^{y,w'}_{s',j}$ of utility j for any year y' for a given realization, divided by the total storage capacity of utility j being at least as great as critical storage S_c.

In Equation 13, hydrologic data variables $\boldsymbol{NI^{y',w'}_j}$, $\boldsymbol{E^{y',w'}_j}$ and $\boldsymbol{W^{y',w'}_j}$ denote total natural inflows to all reservoirs, evaporation rates and reservoir spillage of utility j in year y' prior to current week w used in one N_{rof} simulation. The hydrologic data variables within each N_{rof} simulation is a combination of 50 years of historical data and 20 years of synthetically-generated data.

Variable T_{rof} can carry two values:$T_{rof} = 52$ weeks is used to capture short-term ROFs due to single-year drought, and $T_{rof} \geq 78$ (1.5 years) to capture long-term ROFs due to droughts that are at least 2 years long. The variable $x^{y,w'}_{s',j}$ is the vector of storage states calculated in one year-long ROF simulation using hydrologic data from the past year y'. In turn, $x^{y,w'}_{s',j}$ is a function of unrestricted demand in week w, $\boldsymbol{UD^w}$ where demand is met without triggering any water-use restrictions or purchasing treated transfers, and Ψ_s, which is a metric of sampled DU factors obtained through DU Optimization. On leap

years (when w' may equal 53 weeks), it is assumed that the first week of the historical year will be used.

Overall, to generate a ROF table, a matrix of dimension $L \times W$ (where L is the total number of discretized initial total storage levels (i.e., 20%, 25%, 30%,...100%) and W is the total number of weeks in the time series) is first initialized. For each initial storage level, hydrologic and demand data from week w to $w + T_{rof}$ is selected, and C_j for each week is evaluated while monitoring for failure. Failure triggers an increment of 1 toward the total number of failures. Otherwise, the total number of failures remains unchanged. This is repeated N_{rof} times, resulting in $\sum_{y'=0}^{w} f_{y',j}^{w}(\boldsymbol{NI^{y'}}), \boldsymbol{E^{y'}}$. The final value is then divided by N_{rof} to result in the $x_{rof,j}$ for a given week j. The process is repeated until the ROF for all $w \in W$ is evaluated.

S5 Formulation of the Social Planner and Pragmatist Compromise Portfolios

The Social Planner pathway policy's least-squares formulation is shown in Equation 14:

$$LS = min \sum_{j=1}^{j} [w_i(S_j^* - S_{i,j}]^2 \tag{14}$$

where S_j^* is the maximum robustness achieved for utility j in the Pareto-approximate set, $S_{i,j}$ is the robustness for actor j resulting from solution i, m is the total number of negotiating actors and w_j is a weighting applied to the actor j where here $w_j = 1$ for all actors (everyone is treated equally). It selects a solution based on the minimum total squared distance from all utilities' ideal (maximally robust) solutions. Notably, it does not differentiate the magnitude of the potential loss of a utility's performance. Therefore, it may result in a policy that favors a utility with a better ability to withstand challenging scenarios, or with more resources, as it is 'blind' to such disparities. The Social Planner formulation may therefore be inappropriate when regional members' access to resources is disproportionately scaled to their demand requirements.

Next, the Pragmatist stakeholder is described using the power-index (aka Shapley-Shubik Power Index) formulation described in Equation 15 to Equation 18:

$$PW = \min_i CV \tag{15}$$

$$CV = \frac{\sigma_i}{\bar{\alpha_i}} \tag{16}$$

$$\alpha_{i,j} = \frac{w_j(S_j^* - S_{i,j})}{\sum_{j=0}^{m}(S_j^* - S_{i,j})} \tag{17}$$

$$\sum_{j=0}^{m} \alpha_j = 1 \tag{18}$$

where $\bar{\alpha}_i$ and σ_i are the mean and standard deviations of power index values $\alpha_{i,j}$ across all actors j for solution i, S_j^* is the maximum robustness achieved for utility j in the Pareto-approximate set, $S_{i,j}$ is the robustness for utility j resulting from solution i, and m is the total number of negotiating utilities. This distribution is weighted by the ranking of the influence (perceived power) that each stakeholder has on the system. Here, the weights w_j of all utilities are set equal to 1. This allows our analysis to be consistent with the Sedento Valley analysis in (Gold et al., 2022).

The Pragmatist compromise strategy minimizes the coefficient of variation (CV) of the power index (PW) across all stakeholders' robustness as demonstrated in Equation 16. This policy pathway is deemed 'cooperatively stable' assuming that a utility will view the others as having their fair share of gains and losses. However, the value of PW is also a measure of a utility's potential to improve their allocation. The higher a utility's loss-to-gain ratio, the more likely they are not to cooperate, unintentionally or otherwise. Although it may be seen as a more practical way of selecting a regional compromise compared to the Social Planner compromise, the Pragmatist approach can also conceal its region-wide performance and robustness implications, as well as impacts on cooperative members' performance.

S6 Bootstrapped Hydro-climatic Scenarios Diagnostics

Bootstrapping is utilized in the analysis of the Sedento Valley test case to reduce the number of hydro-climatic realizations needed to re-evaluate the Social Planner and Pragmatist policy pathways and reduce the computational demand associated with maintaining an adequate sampling of hydro-climatic extremes. The bootstrapping method employed here is derived from Zatarain Salazar et al. (2017); Trindade et al. (2019), where bootstrapping large ensembles of synthetic streamflow timeseries were found to satisfactorily approximate the original (large-ensemble) streamflow probability distribution and preserve streamflow extremes across the ensemble. The bootstrapped samples are then used to generate new ROF tables required for simulation in WaterPaths, which is more computationally efficient than generating new ROF tables for the full ensemble of NI realizations.

The number of natural inflow NI realizations (and their associated demand and infrastructure construction multipliers) is reduced by random-uniformly selecting 500 NI realizations conditioned upon the DM. Bootstrapping halves the computational time required to evaluate all portfolio perturbations across all realizations while maintaining the hydro-climatic extremes that each portfolio perturbation might encounter.

The sufficient number of $n = 100, 200, ..., 1000$ hydro-climatic realizations (and their associated demand and infrastructure construction multipliers) to obtain stable performance objective values that represent the full 1,000 realizations was evaluated by selecting a sample of n-realizations from the full set of hydro-climatic realizations. Next, this sample of n-realizations was each paired with 1 DU SOW. One set of decision variables was evaluated across this realization-SOW combination. This step is repeated for all values of n. The standard deviation in each objective for each n-value against n was plotted (Figure S1). The elbow point was located, and its corresponding value of n was used as the number of bootstrapped hydro-climatic realizations. To further verify that $n = 500$

hydro-climatic realizations are sufficient, the probability distributions of the original performance objectives and the bootstrapped performance objectives were compared: To bootstrap the original 1,000 natural inflow (NI) realizations and their associated demand (D) and evaporation rates (E), all realizations were ordered according to their maximum drought magnitudes (DM) as calculated in Supporting Information 2. This study uses the SSI_6 method as described in Herman et al. (2016) and Kirsch et al. (2013) to calculate the drought magnitudes (DM) of all droughts throughout one inflow realization's simulation period.

The Standardized Streamflow Indicator (SSI_6) as used in (Herman et al., 2016) was used to evaluate the drought magnitude (DM) for all natural inflow realizations $NI_i \; \forall i \in N$ where N is the total number of realizations. To calculate the SSI_6, each NI_i, discretized at a weekly timescale, is assumed to follow a lognormal distribution, such that

$$Y_t = lnQ_t \tag{19}$$

which is then standardized:

$$Z_t = \frac{Y_t - \hat{\mu_Y}}{\hat{\sigma_y}} \tag{20}$$

The SSI_6 is evaluated as the 6-month moving average of Z_t. Drought events are defined as when $SSI_6 < 0$ continuously for at least three months, and $SSI_6 < -1$ at least once during this three-month interval.

The sum of SSI_6 over the total duration T of the planning horizon is defined as the drought magnitude or DM:

$$DM = \sum_{t=1}^{T} SSI_{6,t} \tag{21}$$

Next, each realization was split into $\frac{1000}{NI}$ bins where $n = 500$ is the total number of bootstrapped realizations. From each bin, one hydro-climatic realization was sampled to be placed into the overall set of bootstrapped natural inflow (NI_B), demand (D_B),

and evaporation rate (E_B) realizations. Three of these combined variables form one bootstrapped hydro-climatic realization.

To ensure that NI_B captured the full range of extreme inflow values across the planning horizon, a z-test and comparison of the flow duration curves of the original and bootstrapped NI_B realizations are performed. The flow duration curves (FDCs) for each inflow source were plotted (Figure S3. The FDCs of the low flows (Figure S4) were also generated to verify that the distribution extreme low inflows of the bootstrapped realizations were not significantly different from that of the full realization set.

S7 K-Means Clustering

K-Means clustering is an unsupervised machine learning algorithm that attempts to identify k similar groups within a given input set, where k is the number of clusters to be identified by the algorithm. K-Means clustering partitions input data into k-clusters where each cluster contains similar features as defined by their Euclidean distances. Formally, K-Means clustering is explained as follows.

Given k-subsets of the input dataset $\boldsymbol{x}$, now split into $C_1, C_2, ..., C_n$ groups. Each data point x_i in $\boldsymbol{x}$ is assigned to a cluster $C_l \forall l \in [1, k]$ such that each cluster does not overlap, defined in Equation 22.

$$C_l \cap C_{l'} = \varnothing \text{ for } l \neq l' \tag{22}$$

Next, the centroid for the cluster Ci is computed (Equation 23):

$$\mu_l = \frac{1}{||C_l||} \sum_{i \in C_l} x_i \text{ for } l = 1, 2, ..., k \tag{23}$$

Each cluster C_l is then updated by assigning each x_i to the cluster whose μ_l it is closest to. This process is repeated until convergence, which is when the cluster assignments of each x_i do not change.

In this experiment, K-Means clustering was conducted to identify distinctive clusters of the frequency of triggered infrastructure investments across all hydro-climatic realizations to form coherent infrastructure investment pathways across the planning horizon of each utility. The `sklearn.cluster.KMeans` function within the `scikit-learn` Python library (Pedregosa et al., 2011) was implemented using $n_{clusters} = 3$ for Watertown and Dryville, and $n_{clusters} = 2$ for Fallsland. This denotes high, moderate, and light infrastructure investment and construction frequencies for Watertown and Dryville, and high and moderate frequencies for Fallsland. The lower number of clusters for Fallsland was chosen as

there was no significant third cluster that was identified using the K-Means algorithm. The function was initialized using n_{init}= 10 centroid seeds.

S8 Regional Implications of Implementation Uncertainty

Figure S5 shows the cumulative distributions of attained performance for the objectives used to define the robustness satisficing criteria. Figures S5a and c illustrate regional performance tradeoffs similar to that of Figure 6. The cumulative distribution function (CDF) plots in Figures S5b and d show the likelihood that the three key satisficing criteria (indicated by the dashed black line on each CDF subplot) for the SP and PR compromise will be met.

From Figure S5a, moderate policy perturbations primarily drive changes in the SP compromise pathway's regional reliability, restriction frequency and worst-case cost, with little to no change in the remaining two performance objectives. Figure S5b reveals that the drivers of these changes are Watertown's restriction frequency, Dryville's worst-case cost, and Fallsland's reliability. Implementation uncertainties strongly impact Watertown where its frequency of achieving its restriction frequency target drops from 99% in the original unperturbed policy to approximately 40% for many of the perturbed instances. A similar degree of effects occurs for Dryville's frequency achieving the worst-case cost goal, falling from almost a 100% to a 20% for perturbed instances. Implementation uncertainties most strongly influence a decrease in Fallsland's frequency of achieving the reliability performance goal across the DU Re-Evaluation SOWs, dropping from approximately 75% to 50%.

Overall, the satisficing criteria that changed most significantly due to policy perturbations vary across the utilities, demonstrating the asymmetric effects that implementation uncertainty has on the different Sedento Valley utilities' ability to meet their individual robustness satisficing criteria. It also reveals that moderate deviations in the implementation of all three utilities underlying pathway policies' decision variables have the potential to significantly degrade the region's ability to meet all three satisficing criteria. Without

this information, utilities could be cooperating under the assumption that all parties hold equal sway over all satisficing criteria of the Sedento Valley, and may not emphasize the monitoring needed for a better understanding of how a specific utility's pathway policy implementation is shaping regional robustness.

Next, a similar analysis of the PR regional performance tradeoffs (Figure S5c) and its distributions of attained performance across its perturbed instances (Figure S5d) shows that moderate policy perturbations induce a higher regional restriction frequency and worst-case cost but improve regional peak financial cost. The perturbed instances of the PR compromise pathway are consistent with the original solution's soft-path approach, showing no increases in infrastructure investment debt across the Sedento Valley's utilities. Figure S5c shows that the envelope of performance tradeoffs for the perturbed instances of the PR pathway show that several of them yield a decrease in regional peak financial cost, which is the sum of a utility's drought mitigation costs and financial obligations. Relative to the original PR solution, these perturbed instances reduce short-term drought mitigation costs by more actively restricting water use versus purchasing more costly treated transfers. However, this also introduces the potential for increases in financial variability arising from revenue losses, increasing the region's worst-case cost (Figure S5c).

The performance attainment distributions for the perturbed instances of the PR compromise pathway shown in Figure S5d highlight that regional changes in robustness are driven primarily by Dryville. Dryville is frequently the worst-performing utility in the region within the minimax formulation used in the DU Optimization and when evaluating regional robustness. More specifically, Dryville's moderate policy perturbations cause its frequency of achieving restriction frequency and worst-case cost performance goals in the DU Re-Evaluation sampled SOWs to decrease from 60% to as low as 30%, and

from 95% to as low as 20%, respectively. Watertown and Fallsland also show degraded performance with implementation uncertainty for their restriction frequency and worst-case cost satisficing criteria, but neither are as significant as those faced by Dryville. Dryville's ability to successfully meet all of its robustness criteria is heavily influenced by the assumption that the Sedento Valley regional utilities precisely implement the PR compromise pathway's actions. The PR compromise pathway policy disproportionately impacts Dryville's vulnerability to implementation uncertainty. It is worth noting that the original PR compromise pathway policy resulted in Dryville achieving significant robustness gains. However, policy perturbations would likely result in Dryville facing the largest loss in robustness across the three utilities.

S9 Robustness Across All Portfolio Perturbed Instances

Figures S6 and S7 show the ranked full range of robustness values attained by all perturbed instances of the SP and PR compromise portfolios respectively.

The set of decision variables that correspond to the original compromise portfolio and its least-robust instance can be observed in Figure S8.

Figure S9 is a three-dimensional scatter plot showing how perturbations in long-term investment ROF triggers in the Social Planner compromise pathway policy can affect the utilities' abilities to achieve their short-term reliability and long-term infrastructure net present cost performance goals. In the three illustrated clusters, a higher ROF trigger value indicates a higher tolerance for risk. The triangle denotes the original Social Planner compromise, and the cross denotes its least-robust perturbed instance. Across all utilities, attaining higher reliability is sensitive to relatively modest changes in the long-term infrastructure ROF trigger. This trend is most clear for Dryville (Figure S9b), where the utility may fail to meet its reliability satisficing criterion (reliability $\geq 98\%$) by tolerating more long-term infrastructure risk while also reducing debt burden through the elimination or delay of several supply infrastructure options (Figure S9). Figure S9c also illustrates the large financial impact caused by modest changes in a Fallsland's infrastructure trigger, where a slight alteration can result in up to a $80 million change in its infrastructure net present cost. This sunk cost for Fallsland represents a significant financial risk. These findings are characteristic of a 'hard-path' policy that relies on the development of new supply infrastructure to ensure a reliable supply of water.

Given the complex, two-way interaction between short-term management and long-term planning actions, the implementation of cooperative pathway policies must therefore be carefully executed, lest the benefits of cooperation are lost. Thus, utilities should be informed of the level of precision required during implementation for regional cooperative

water supply investment and management pathways to perform as expected. Identifying these operational tolerances requires delineating individual safe operating spaces (SOS) which are decision variable operational tolerance ranges in which each utility's robustness remains the same or improves from its original robustness value.

Table S1. Watertown water supply portfolio decision variables.

Decision variable	**Abbreviation**	**Lower bound (%)**	**Upper bound (%)**
Restriction ROF trigger	RT_W	0	100
Lake Michael allocation	LMA_W	33.4	90
Annual reserve fund contribution (%of annual revenue)	RF_W	0	10
Insurance ROF trigger	IT_W	0	100
Insurance payment (%of annual revenue)	IP_W	0	2
Infrastructure construction long-term ROF trigger	INF_W	0	100
New River Reservoir ranking	NRR	1st	8th
College Rock Expansion (Low) ranking	CRL	1st	8th
College Rock Expansion (High) ranking	CRH	1st	8th
Water Reuse I ranking	$WR1$	1st	8th
Water Reuse II ranking	$WR2$	1st	8th

Table S2. Dryville water supply portfolio decision variables.

Decision variable	**Abbreviation**	**Lower bound (%)**	**Upper bound (%)**
Restriction ROF trigger	RT_D	0	100
Transfer ROF trigger	TT_D	0	100
Lake Michael allocation	LMA_D	5	33.4
Annual reserve fund contribution (%of annual revenue)	RF_D	0	10
Insurance ROF trigger	IT_D	0	100
Insurance payment (%of annual revenue)	IP_D	0	2
Infrastructure construction long-term ROF trigger	INF_D	0	100
Sugar Creek Reservoir ranking	SCR	1st	8th
Water Reuse ranking	WR	1st	8th

Table S3. Fallsland water supply portfolio decision variables.

Decision variable	**Abbreviation**	**Lower bound (%)**	**Upper bound (%)**
Restriction ROF trigger	RT_F	0	100
Transfer ROF trigger	TT_F	0	100
Lake Michael allocation	LMA_F	5	33.4
Annual reserve fund contribution (%of annual revenue)	RF_F	0	10
Insurance ROF trigger	IT_F	0	100
Insurance payment (%of annual revenue)	IP_F	0	2
Infrastructure construction long-term ROF trigger	INF_F	0	100
New River Reservoir ranking	NRR	1st	8th
Water Reuse ranking	WR	1st	8th

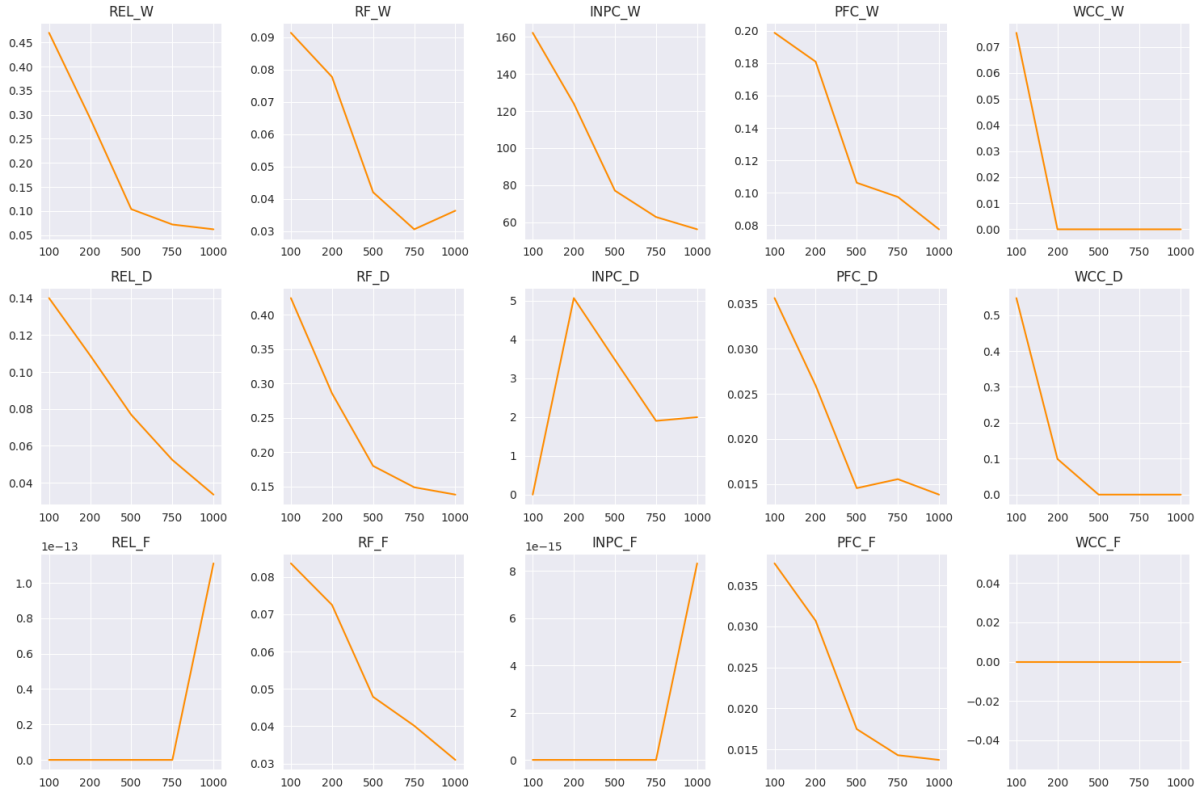

Figure S1. The standard deviation between the original performance objective value and the bootstrapped performance objective value (y-axis) across n-realizations (x-axis). Each column represents a performance objective and each row represents each utility (Watertown, Dryville, and Fallsland)

.

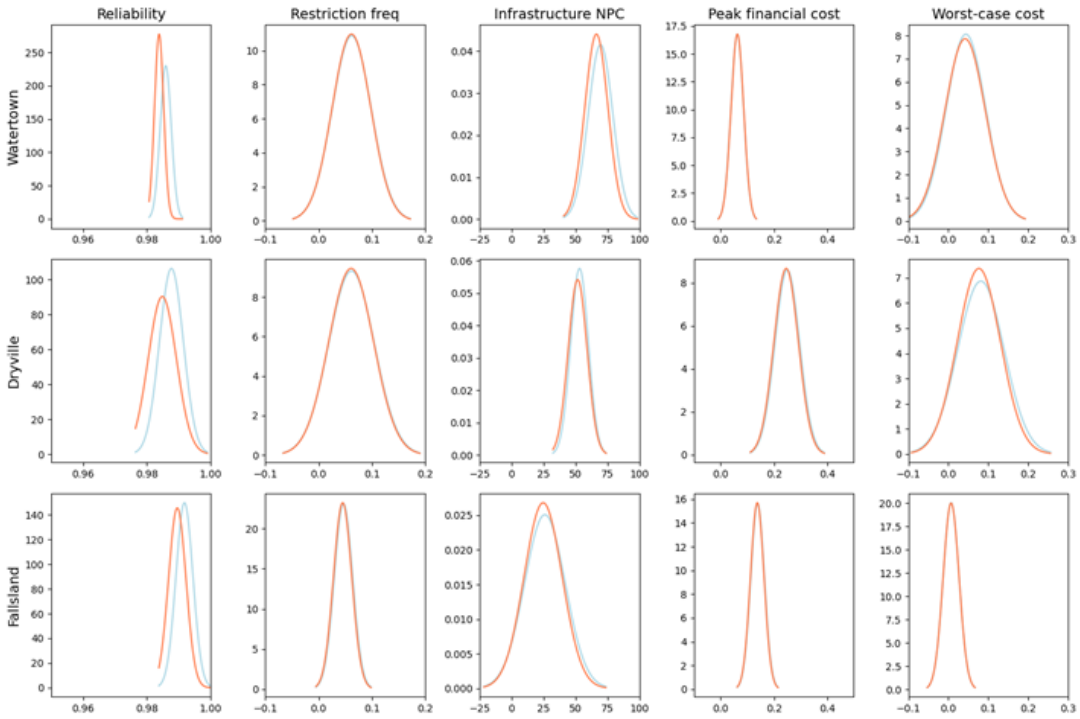

Figure S2. The probability distribution of the original performance objective value (blue) and the bootstrapped performance objective values (orange) evaluated across n-realizations, where $n = 500, 1000$. Each column represents a performance objective and each row represents each utility (Watertown, Dryville, and Fallsland).

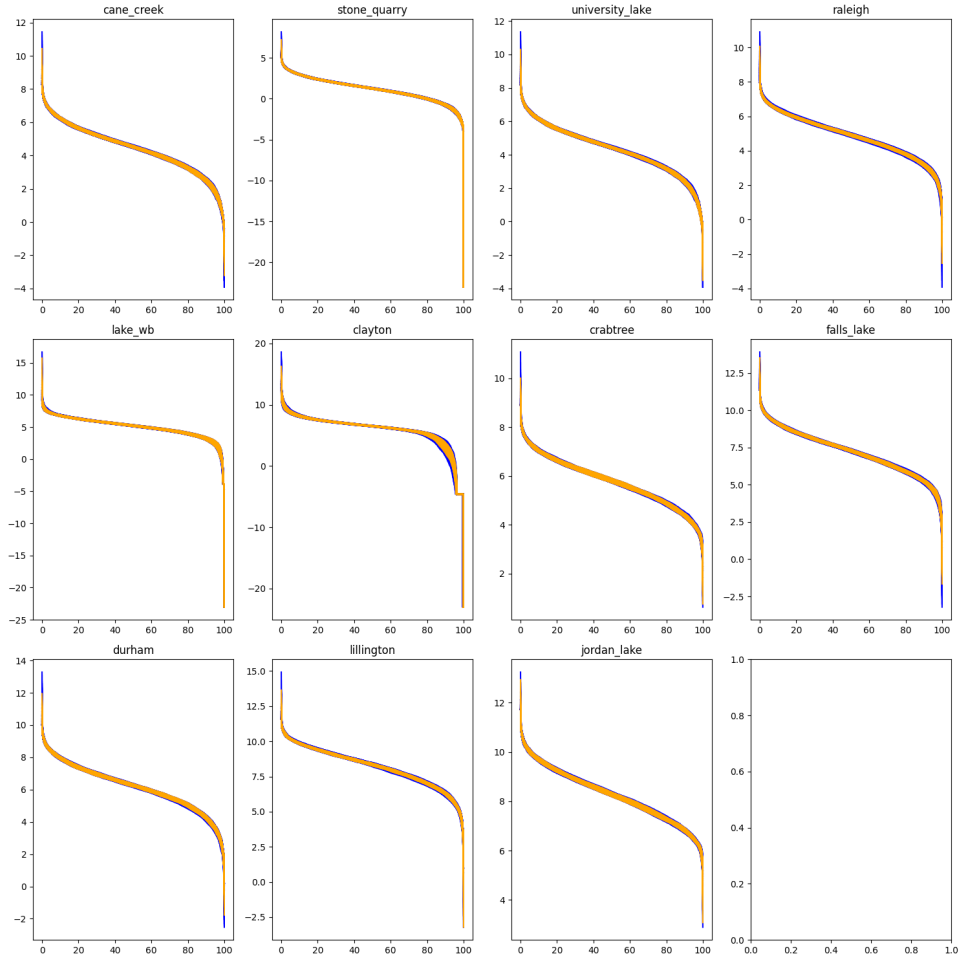

Figure S3. The flow duration curves for all inflow sources across the entire Sedento Valley. The x-axis denotes cumulative probability from 0% to 100%, and the y-awis denotes the total inflow in million gallons (MG).

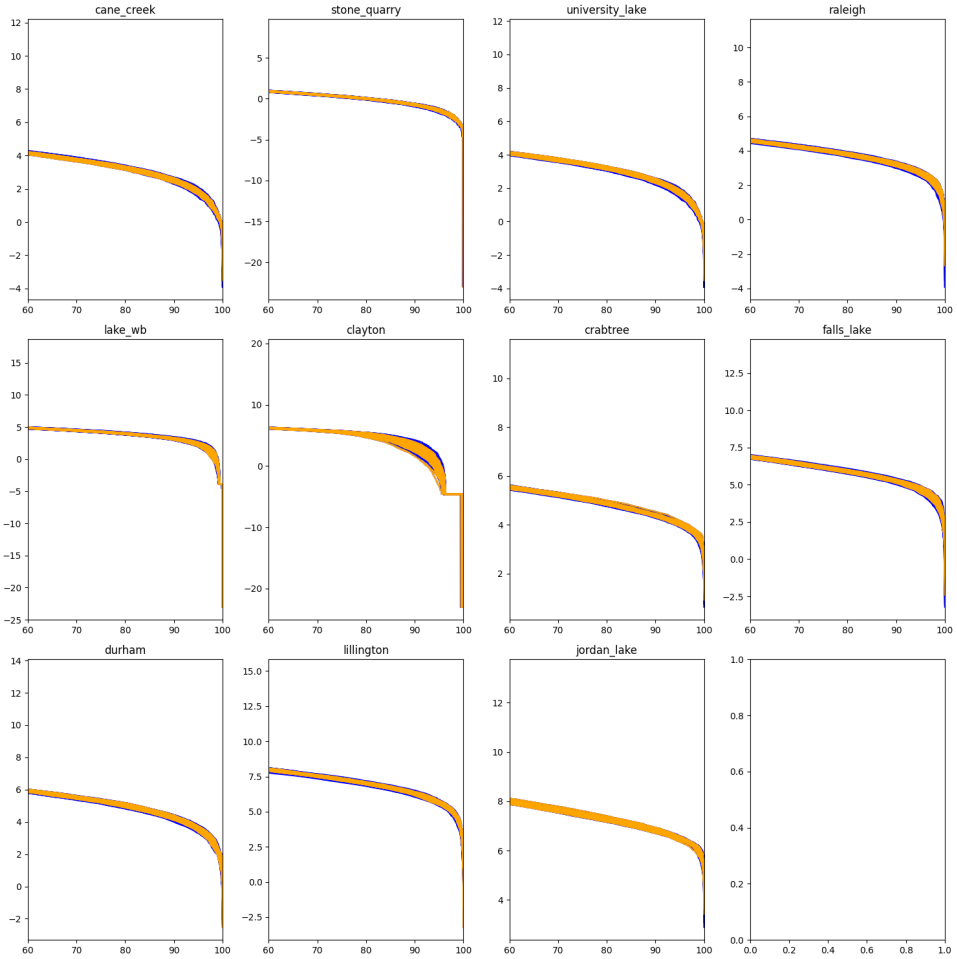

Figure S4. The flow duration curves for only low inflows of up to 20MGD for all inflow sources across the entire Sedento Valley. The x-axis denotes cumulative probability from 0% to 100%, and the y-awis denotes the total inflow in million gallons (MG).

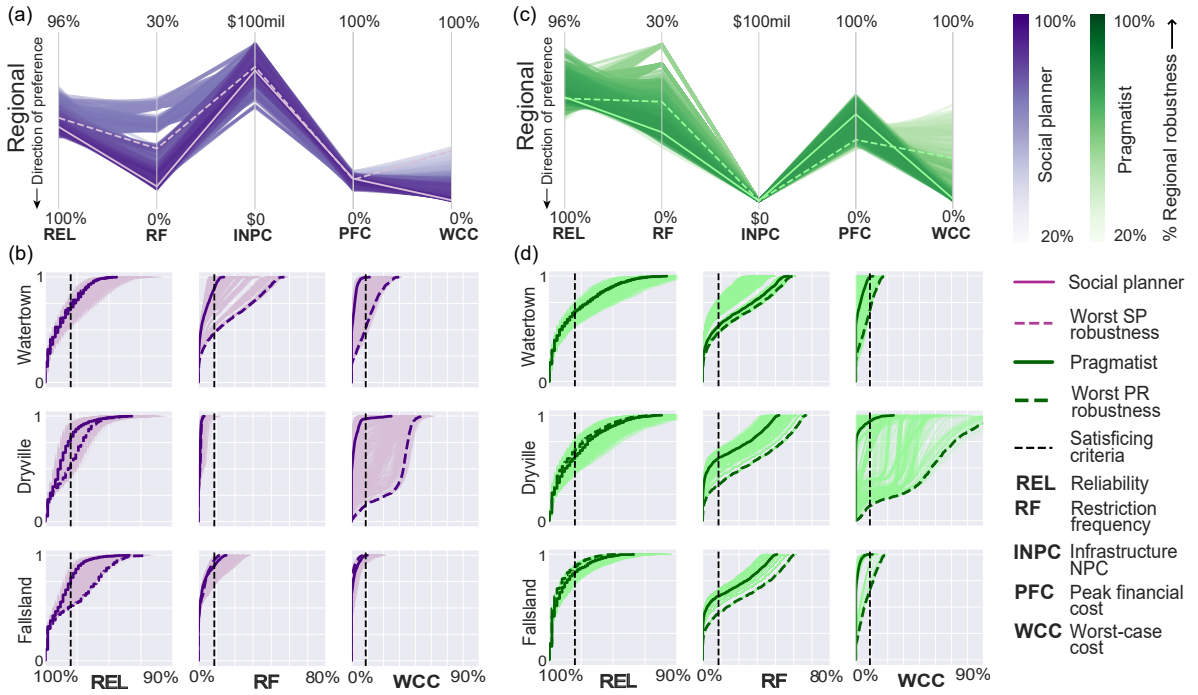

Figure S5. The perturbed envelope of regional performance tradeoffs and performance attainment for the three objectives used to assess robustness. The purple subplots denote the SP compromise; the green subplots denote the PR compromise. Panels (a) and (c) show the parallel axis plots of the regional performance objective tradeoffs. The color and gradient of the lines indicating the value of regional robustness, where a darker color indicates a higher regional robustness, and vice versa. Panels (b) and (d) show cumulative distribution performance attainment plots for the three performance objectives used to assess robustness. Each row shows the distribution of performance attainment for each utility. Each column represents one of the three objectives used to measure robustness. The robustness performance goals are indicated by the black dashed lines. Each colored line in the attainment distribution plots denotes the cumulative distribution of robustness for a given perturbed instance of either the original SP or PR compromise pathway policies.

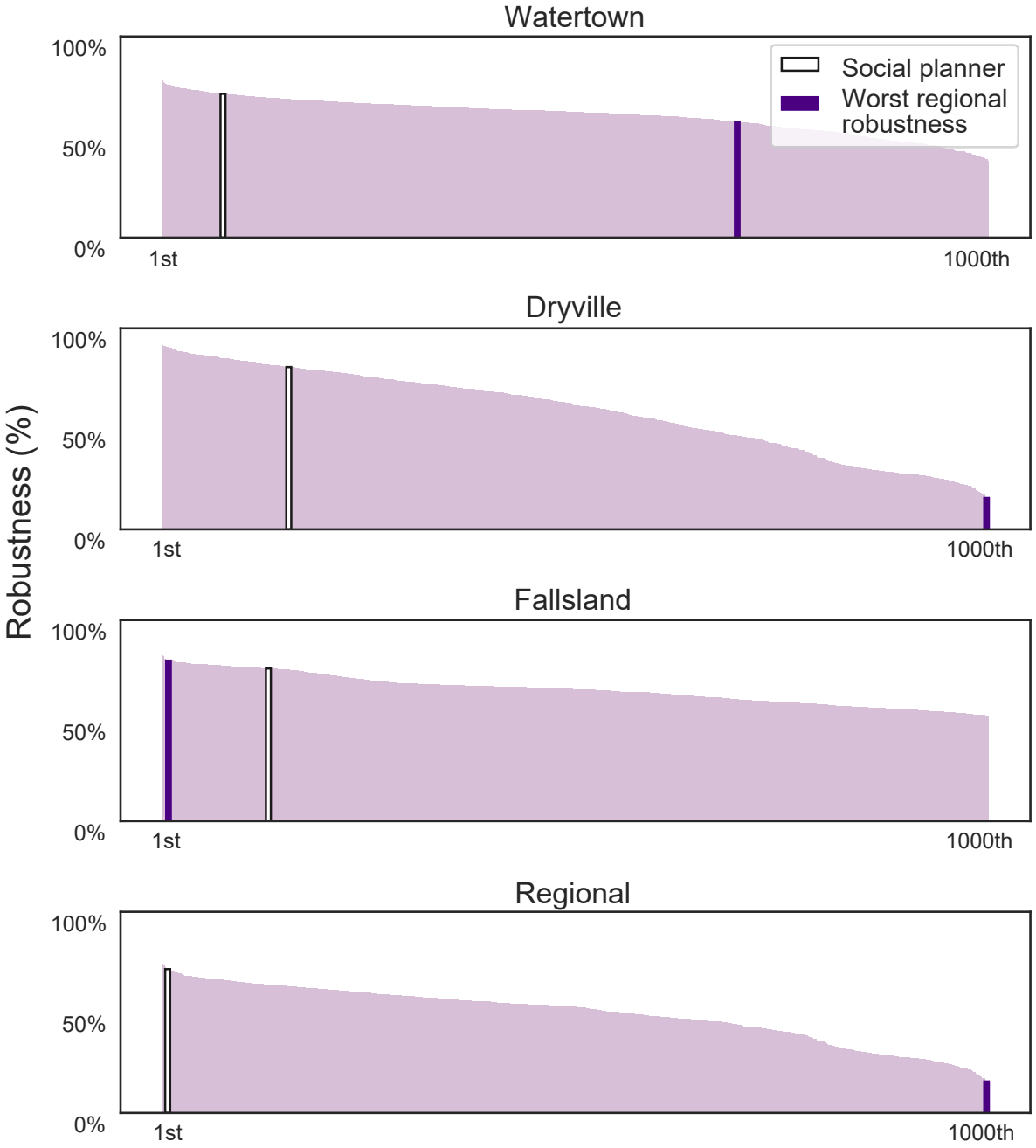

Figure S6. Ranked robustness of the original Social Planner compromise portfolio and that of all its perturbed instances. The white bar indicates the original compromise portfolio, and the purple solid bar denotes the least-robust perturbed instance across the entire region. Each row denotes all utilities and the region.

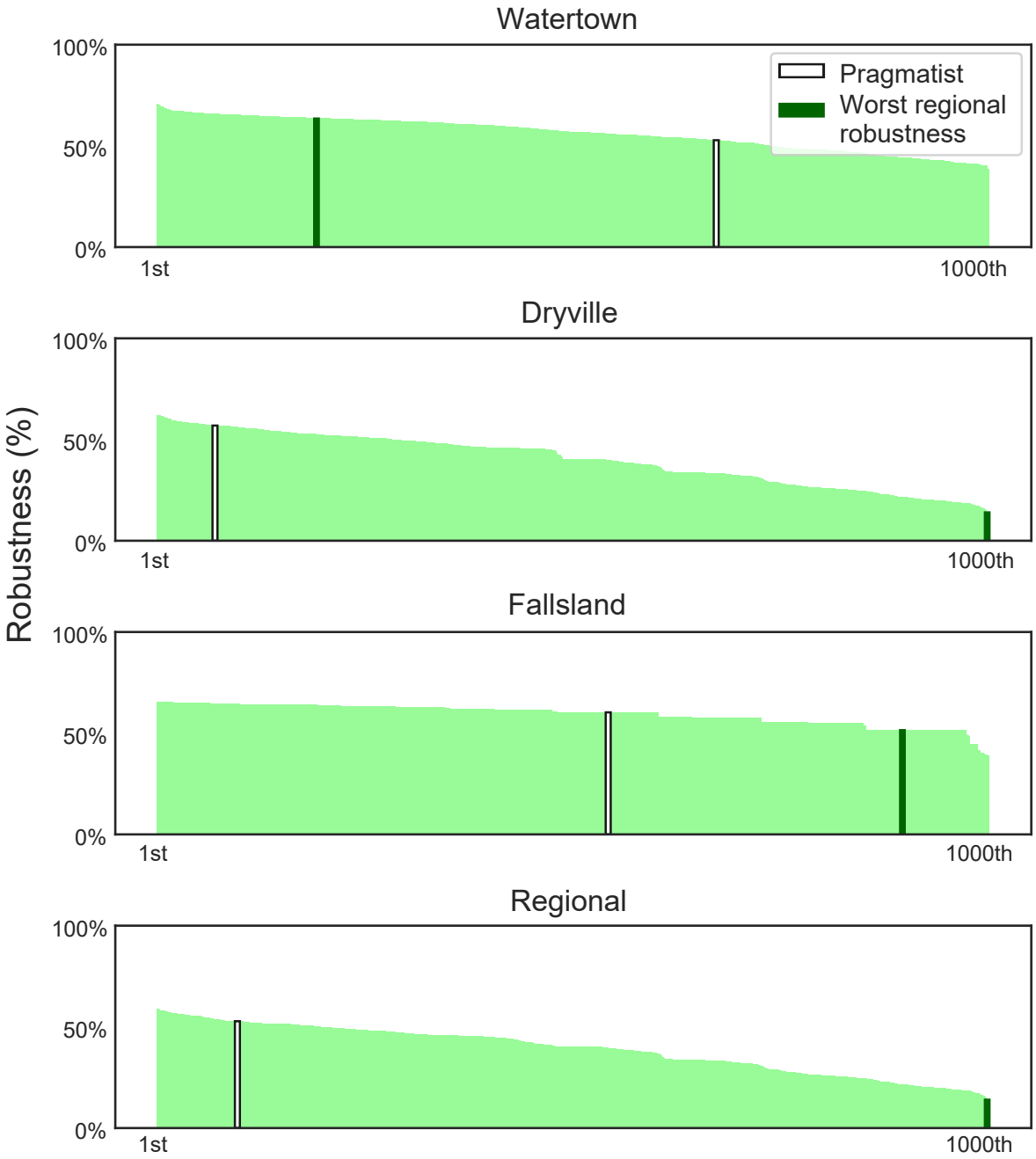

Figure S7. Ranked robustness of the original Pragmatist compromise portfolio and that of all its perturbed instances. The white bar indicates the original compromise portfolio, and the green solid bar denotes the least-robust perturbed instance across the entire region. Each row denotes all utilities and the region.

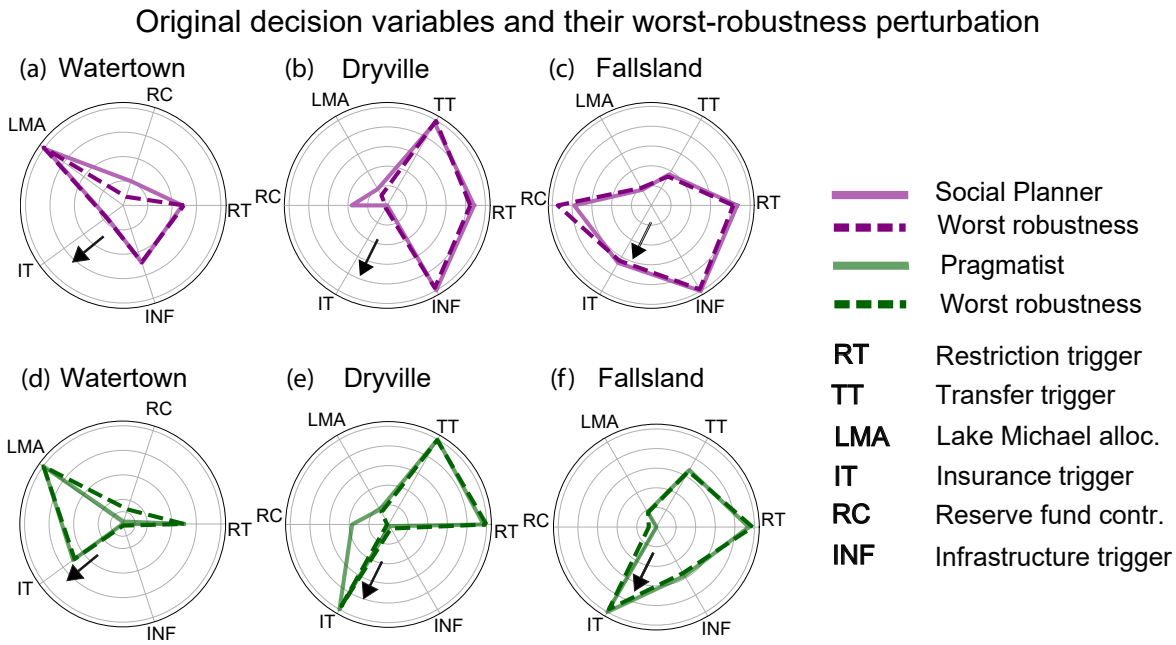

Figure S8. The decision variables of the Social Planner (panels (a) to (c)) and Pragmatist (panels (d) to (e)). The solid line indicates the decision variables of the original compromise portfolio, and the dotted line denotes the decision variables that result in each utility's least-robust perturbed instance. The direction of the black arrow denotes increased used of the decision variable.

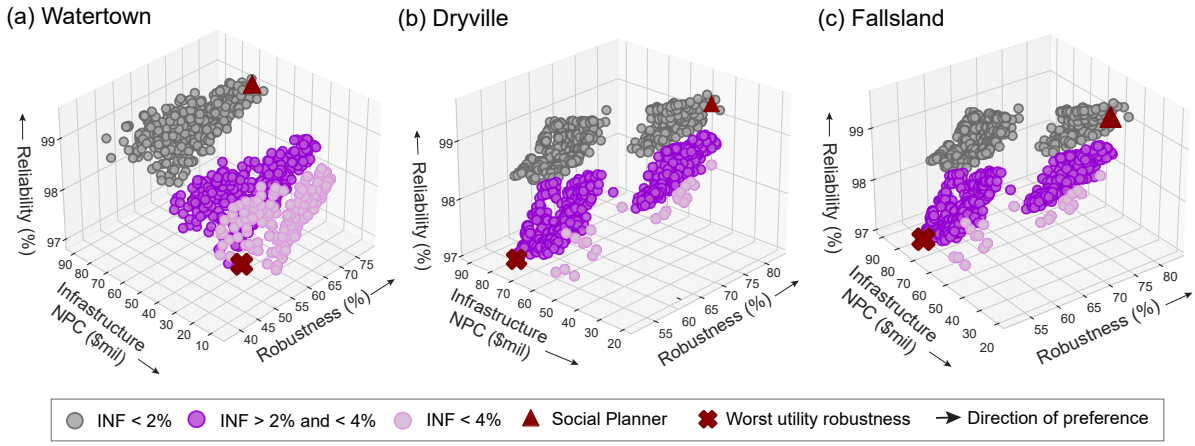

Figure S9. Clusters of infrastructure investment triggers and their associated outcomes for the three Sedento Valley utilities in panels (a) to (c). The colors distinguish three infrastructure trigger ranges (less than 2%, between 2% and 4%, and more than 4%). The triangle denotes the original Social Planner compromise pathway policy, while the cross denotes its perturbed instance that results in the lowest robustness. These axes plot the individual utilities' infrastructure NPC, robustness, and reliability. The arrows indicate the direction of preference along the axes.

References

Caldwell, C., & Characklis, G. W. (2014). Impact of contract structure and risk aversion on interutility water transfer agreements. , *140*(1), 100–111. (Publisher: American Society of Civil Engineers) doi: 10.1061/(ASCE)WR.1943-5452.0000317

Gold, D. F., Reed, P. M., Gorelick, D. E., & Characklis, G. W. (2022). Power and pathways: Exploring robustness, cooperative stability, and power relationships in regional infrastructure investment and water supply management portfolio pathways. , *10*(2), e2021EF002472. doi: 10.1029/2021EF002472

Herman, J. D., Zeff, H. B., Lamontagne, J. R., Reed, P. M., & Characklis, G. W. (2016). Synthetic drought scenario generation to support bottom-up water supply vulnerability assessments. , *142*(11), 04016050–04016050. doi: 10.1061/(asce)wr.1943-5452.0000701

Kirsch, B. R., Characklis, G. W., & Zeff, H. B. (2013). Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: Practical improvement for generating synthetic streamflows. , *139*(4), 396–406. doi: 10.1061/(ASCE)WR.1943-5452.0000287

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, (2011). Scikit-learn: Machine learning in python. , *12*(85), 2825–2830.

Trindade, B., Gold, D. F., Reed, P. M., Zeff, H. B., & Characklis, G. W. (2020). Water pathways: An open source stochastic simulation system for integrated water supply portfolio management and infrastructure investment planning. , *132*, 104772. doi: 10.1016/j.envsoft.2020.104772

Trindade, B., Reed, M., Patrick, & Characklis, G. W. (2019). Deeply uncertain pathways: Integrated multi-city regional water supply infrastructure investment and portfolio management. , *134*. doi: 10.1016/j.advwatres.2019.103442

Zatarain Salazar, J., Reed, P. M., Quinn, J. D., Giuliani, M., & Castelletti, A. (2017). Balancing exploration, uncertainty and computational demands in many objective reservoir optimiza-

tion. *Advances in Water Resources*, *109*, 196–210. doi: 10.1016/j.advwatres.2017.09.014

Zeff, H. B., Herman, J. D., Reed, P. M., & Characklis, G. W. (2016). Cooperative drought adaptation: Integrating infrastructure development, conservation, and water transfers into adaptive policy pathways. , *52*(9), 7327–7346. doi: 10.1002/2016WR018771