Jared D. Smith[1], Julianne D Quinn[1], and Lawrence E Band[1,2]

[1]Civil and Environmental Engineering, University of Virginia
[2]Environmental Science, University of Virginia

July 31, 2024

# COMPARING ROBUST OPTIMIZATION APPROACHES FOR ADDRESSING HYDROLOGIC MODEL UNCERTAINTY IN INFRASTRUCTURE PLANNING: A GREEN INFRASTRUCTURE EXAMPLE

Jared D. Smith, Ph.D.[1*], Julianne D. Quinn, Ph.D.[2], and Lawrence E. Band, Ph.D.[2,3]

[1]University of Virginia, Civil and Environmental Engineering, Charlottesville, VA, USA. Email: jared.d.smith485@gmail.com

[*]Currently employed at: U.S. Geological Survey, Reston, VA, USA

[2]University of Virginia, Civil and Environmental Engineering, Charlottesville, VA, USA

[3]University of Virginia, Environmental Science, Charlottesville, VA, USA

## ABSTRACT

Water resources planning depends upon hydrologic models to estimate flows and storage under candidate engineering designs. However, such models are calibrated with limited flow data relative to the many model parameters. This may result in different equifinal parameterizations that imply different optimal designs. To assess if and how this uncertainty should be considered, we compare three methods for multi-objective optimization of green infrastructure (GI): one that designs to the most likely parameterization and two robust alternatives that use several likely parameterizations with 1) likelihood-weighted objective functions, and 2) min-max objective functions. To evaluate these methods, we set synthetic true values for model parameters, use them to simulate "observed" streamflow, and then use Bayesian calibration to estimate parametric uncertainty. We compare results from optimization to the synthetic parameterization against the three alternatives. The GI optimizations aim to minimize flooding, low flow intensification, and cost. We find the two robust methods provide objective values and decisions that are closer to those optimized to the synthetic

truth, demonstrating value in considering hydrologic model uncertainty in water resources system designs.

## INTRODUCTION

Water resources engineers develop plans for new water management infrastructure and operations using hydrologic models that estimate water flows, storage, and water quality for candidate designs. Hydrologic models are subject to a variety of uncertainty sources (Vrugt 2016), including spatially limited observations for model calibration, which often leads to equifinality in the estimated model structure and parameters (Beven and Freer 2001). Equifinal parameterizations provide similar model output values at locations with observations yet can provide different values elsewhere (Smith et al. 2022). This could have proximate consequences on engineering designs if equifinal parameterizations imply different optimal sites or design specifications due to the uncertainty in model outputs across a watershed. We present and compare different optimization methods to discover engineering designs for hydrological models with uncertain parameters.

Characterizing model uncertainty has been studied for decades in hydrological sciences (Herrera et al. 2022), yet it is rarely considered in robust optimization studies within water resources planning and management literature (Herman et al. 2020; Tebyanian et al. 2022), wherein the focus of robustness analyses is more commonly on discovering designs that perform well under the “deep uncertainty" of future changes (e.g., climate, population growth) (Webber and Samaras 2022). While uncertainty in future change often dominates the total model output uncertainty at longer timescales (Vaghefi et al. 2019), model uncertainty can have a non-negligible contribution to the total uncertainty in future simulations, particularly in the near term (Lehner et al. 2020) and for extremes (Steinschneider et al. 2012).

Robust optimization over equifinal model parameterizations has been used in a couple studies to address model uncertainty in decision making. Jia and Culver (2006) used generalized likelihood uncertainty estimation (GLUE) (Beven and Binley 1992) to characterize hydrologic model parameter uncertainty, followed by a fuzzy-likelihood weighted multi-objective optimization of spatial pollutant load distributions to meet water quality requirements. Williams et al. (2020) used

multi-objective optimization to calibrate an agent based model to maximize fit and diversity in the parameter space, then evaluated alternative management plans across the diverse parameterizations. Neither of these studies compare their optimizations to alternative approaches and therefore it is unclear if using an approach that is robust to different model parameterizations leads to better decisions than optimizing to a single most likely parameterization. We therefore aim to address the following research question: *Do optimization approaches that are robust to modeled streamflow uncertainties provide better solutions than approaches that ignore streamflow uncertainties?*

To address this question, we rely on synthetic true values for hydrologic model parameters so that we can evaluate the performance of Pareto optimal solutions that are discovered from traditional and robust optimization approaches relative to the solutions discovered using the true parameterization. We define traditional optimization as using only the most likely parameter values that are discovered during model calibration. We consider two competing robust optimization strategies: multi-objective robust optimization (MORO) (Bartholomew and Kwakkel 2020; Shavazipour et al. 2021) across a likelihood-weighted average of the objectives obtained from several likely parameterizations, and a min-max robust optimization (MinMax) that considers minimizing the worst obtained objective value across the likely parameterizations (Wald 1949). Note that the MinMax approach is also a MORO approach; it simply differs in its objective weighting scheme. These robust optimization choices are motivated by previous studies that demonstrate meaningful differences in discovered solutions from these types of problem formulations (Quinn et al. 2017; Giuliani and Castelletti 2016; McPhail et al. 2018).

We consider as an example a common multi-objective problem of controlling flooding in a watershed using green infrastructure (GI) (Fletcher et al. 2015). We use a hydrologic model to estimate flows in the watershed, and we are interested in GI portfolios (spatial allocations) that are robust to the uncertainty in simulated streamflows. Numerous studies have been published on the topic of multi-objective GI optimization, often for use as water quantity or quality control measures; however, a recent review of such studies by Tebyanian et al. (2022) found only two examples that employed multi-objective robust optimization and neither considered robustness to hydrologic

model uncertainty. GI and other low impact development strategies have become popular solutions for flood control while increasing green space in communities, often at a reduced cost compared to more traditional gray infrastructure (Golden and Hoghooghi 2018). GI helps to reduce the impact of urbanization on increased impervious surfaces and subsequent increases in flood peaks, with a recent study estimating that a 1% increase in urban land leads to an expected 2 - 5% increase in flood magnitudes (Blum et al. 2020). Partly in response to urbanization, municipalities across the United States have created goals to make the percentage of runoff that flows through stormwater GI about 10-20%, including Baltimore County, MD, USA where our example watershed is located (Pennino et al. 2016).

We consider reforestation as a type of GI that provides flood control (Penning et al. 2023). In this study, reforestation with a native needleleaf evergreen tree species, Virginia Pine (*Pinus virginiana*), is the only type of GI considered, so we use GI to mean reforestation throughout the manuscript. A possible consequence of reforesting a watershed is exacerbating low flows (Buechel et al. 2022), which could contribute to ecological consequences (Poff et al. 1997) and earlier onset of hydrologic drought conditions (Penning et al. 2023). We therefore expect tradeoffs between these two objectives and a cost objective. An overview of the modeling objectives for our example optimization problem is provided in Fig. 1. Further details about the objectives and the experimental design of the optimizations are provided in the following section.

## METHODS

### Experimental Design: Optimization Approaches

We aim to benchmark the performance of two robust optimization approaches and a traditional optimization approach using an example GI optimization problem. We allocate the same computational time to each optimization approach to allow for a fair comparison. To enable benchmarking, we select synthetic true parameter values for a hydrological model (described in Section 2), and run a GI optimization using this parameterization to estimate a synthetic true Pareto front of optimal solutions. The other approaches first develop estimates of these true parameters, and then optimize GI portfolios using model runs with one or more estimated parameterizations. We evaluate how

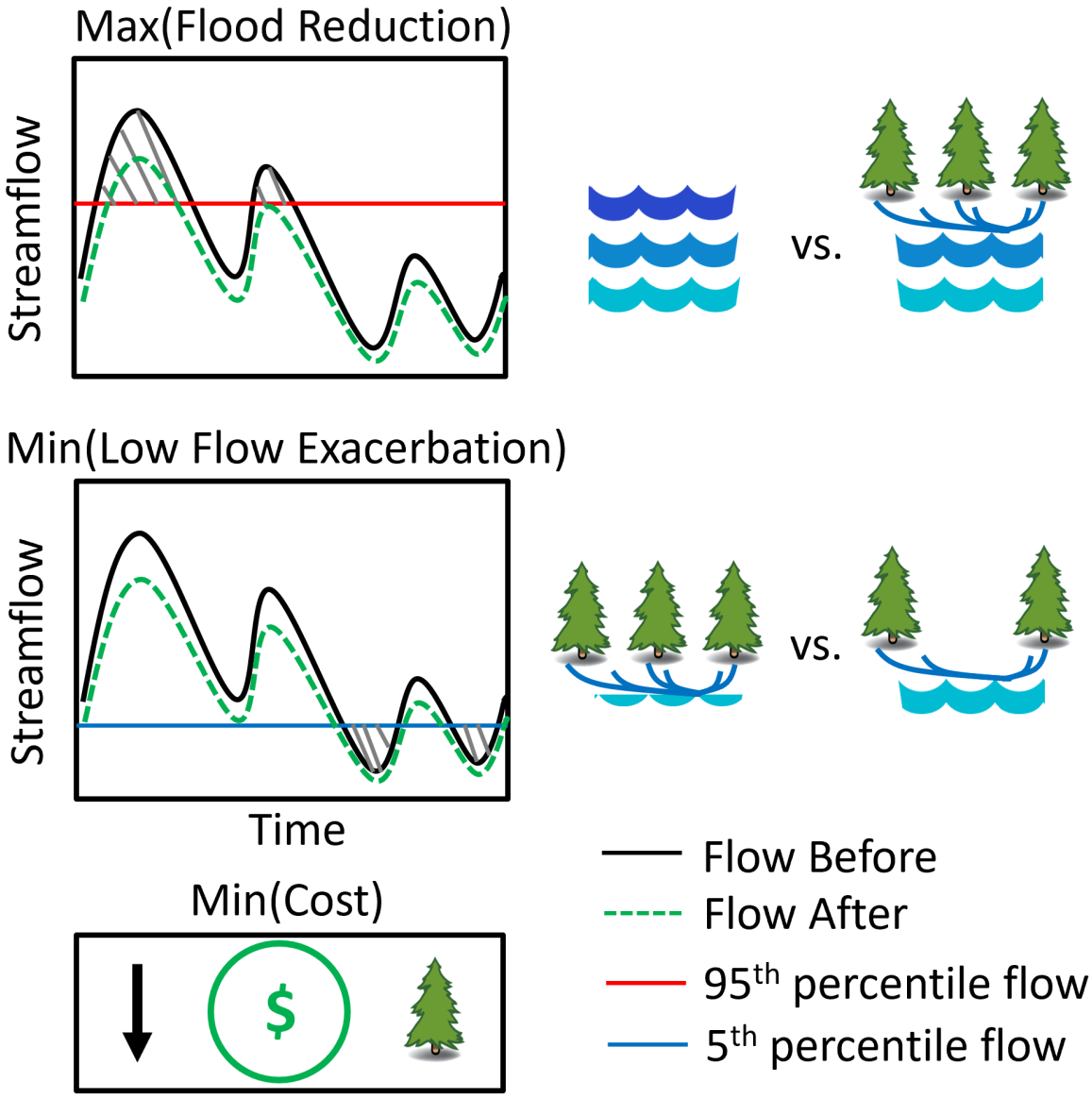


**Fig. 1.** Illustration of the three objectives used for our optimization study (left), and the expected impact of reforestation on streamflows (right).

much the solutions discovered by the optimization approaches deviate from the synthetic true solutions in terms of both their decisions (amount of GI and GI spatial locations) and objectives (flood reduction and low flow exacerbation). We compare objective values for each solution by computing the deviation from the value obtained for the parameterizations used in the optimization, and the value obtained from running the solution on the synthetic true parameterization. Each optimization can be described generally by equation 1:

$$\mathbf{D}^* = \underset{\mathbf{D}}{\operatorname{argmin}} \quad \mathbf{J^f}(\mathbf{D}) \tag{1}$$

where $\mathbf{D}^*$ is the Pareto-optimal set of decision variable vectors (GI portfolios) from the set of possible vectors $\mathbf{D}$ (defined in Section 2) and $\mathbf{J^f}$ is a vector of objectives for optimization formulation $f$.

The objectives for the Synthetic and traditional optimizations have a similar format. The Synthetic optimization seeks to find the Pareto set using the synthetic true parameter set, while the traditional optimization seeks to find the Pareto set for the estimated maximum a-posteriori (MAP) parameterization. These objectives are described mathematically by Equations 2 and 3:

$$J_i^{MAP} = J_{i,\theta_{MAP}} \tag{2}$$

$$J_i^{Syn} = J_{i,\theta_{Syn}} \tag{3}$$

where $J_{i,\theta_{MAP}}$ and $J_{i,\theta_{Syn}}$ are the values of the $i^{th}$ objective obtained from running the hydrologic model with the MAP parameterization, $\theta_{MAP}$, and Synthetic parameterization, $\theta_{Syn}$, respectively.

The two robust optimization approaches are multi-objective robust optimization (MORO) and min-max optimization (MinMax). The MORO approach minimizes the likelihood-weighted average of the objective values obtained from a set of likely hydrological model parameterizations, $\Theta$, as shown in Equation 4:

$$J_i^{MORO} = \sum_{\theta \in \Theta} \frac{\mathcal{L}_\theta}{\sum_{\theta \in \Theta} \mathcal{L}_\theta} J_{i,\theta} \tag{4}$$

where $J_{i,\theta}$ is the value of the $i^{th}$ objective for the parameterization $\theta$, and $\mathcal{L}_\theta$ is the likelihood of the parameterization. The MinMax approach minimizes (Equation 1) the maximum (worst case) objective values obtained across several likely hydrological model parameterizations, as shown in Equation 5

$$J_i^{MinMax} = \max(\mathbf{J_{i,\Theta}}) \quad (5)$$

where $\mathbf{J_{i,\Theta}}$ is a vector of values for the $i^{th}$ objective that is obtained for a single candidate GI portfolio that is evaluated on each of the hydrologic model parameterizations in $\Theta$.

The MORO approach considers a wider sample of possible parameterizations than the MAP approach while being weighted towards the MAP, whereas the MinMax approach is only concerned with the worst case scenario. We consider the worst case scenario independently for each objective, such that different parameterizations could provide the worst case flooding and low flow objective values. We explore which parameterizations provide the worst case scenario for each MinMax Pareto solution to guide interpretation of the optimization results.

*Selection of Likely Parameterizations*

Robust optimizations are limited by available computational resources, so considering a large number of posterior samples in $\Theta$ may be infeasible. We used a method to reduce the estimated posterior to several representative yet unique samples, which is similar in concept to diverse model calibration (DMC) presented in (Williams et al. 2020). We use a two step approach: 1) model performance metric-based screening of parameterizations, and 2) k-means clustering to discover clusters of parameterizations with similar characteristics. We first selected higher performing model parameterizations from the posterior using a threshold of at least 0.7 Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe 1970) for the log of streamflow. We used the log of streamflow because the NSE for logs has lower variability and is less sensitive to poorly fitting extremes (Lamontagne et al. 2020). This step only removed 14% of the posterior samples that were largely in the beginning of the Markov chains used for posterior estimation, so this step effectively served as a performance-based burn-in method.

We then applied k-means clustering to the remaining parameterizations and evaluated the clustering performance using the total within-cluster sum of squared errors for one (no clustering) to 40 clusters (supplementary information [SI] Fig. S1). About 50% reduction in error relative

to using one cluster was achieved using twelve clusters, and about 67% reduction was achieved using 40 clusters. There were small improvements from 20 to 40 clusters. We performed the optimization on the University of Virginia's Rivanna High Performance Computing (HPC) cluster. When we began this study, the Rivanna HPC had 20 cores per node. To maximize computational efficiency of the robust optimizations under the configuration of this cluster, we considered using either nine or nineteen parameterizations (one or two worker cores running nineteen or nine parallel computations, respectively). The runtime of one simulation was close to one hour for our problem, so we determined that about 7000 function evaluations could be completed with nine parameterizations and 3500 could be completed with nineteen parameterizations within the time limits of the HPC cluster. We decided to use nine parameterizations to increase the chance of Pareto front convergence within the allocated time. Nine clusters reduced the within-cluster variance by 45% compared to a single cluster.

To select the nine parameterizations, we computed the Euclidean distance from each parameterization in a cluster to the cluster center and selected the parameterization with the smallest distance. The cluster with the MAP sample used the MAP instead of the cluster center.

*Computational Implementation*

We used the multi-manager (master) implementation of the Borg multi-objective evolutionary algorithm for optimization (Hadka and Reed 2013; Hadka and Reed 2015). Our approach used two managers that provided simultaneous searches for Pareto solutions, which has been shown to outperform methods that utilize one manager when a large number of cores are available (Zatarain Salazar et al. 2017). For this approach, a controller core assembles global Pareto solutions from local solutions that are provided by each manager. The managers assemble local Pareto solutions that are provided from jobs that they allocate to workers that run the hydrological model for each candidate GI portfolio. The computational resource we used had nodes with 20 total cores when we began this study, and we could use a maximum of 45 nodes. We needed one core as a controller, two cores as managers, and the remaining cores were available as workers (448 and 449 per manager). The Synthetic and MAP optimizations utilized all available workers to

run the hydrological model, compute objectives, and send their values back to their managers. The robust optimizations used workers to spin off parallel jobs that ran the hydrological model with the selected parameterizations (one job per parameterization) and sent the resulting objective values back to their manager. We made minor modifications to the Borg algorithm to enable the functionality of the robust optimization method (e.g., passing the worker ID number, manager ID number, and the number of function evaluation (NFE) to the solver). Because each evaluation of a GI solution using the robust optimization approaches required nine model runs, one for each parameterization, we could complete about $1/9^{th}$ as many evaluations of GI solutions with these approaches for a fixed computational time relative to the Synthetic and MAP optimizations. We used a 3-day computational time on 900 cores for all experiments, which was the maximum allowed on the Rivanna HPC at the time of this study. All job submission scripts are available in our code repository.

**Optimization Objectives**

The flooding and low flow objectives rely on the summation of flows below and above a quantile threshold, as shown in Equations 6 and 7

$$Q_{Flood}(Q) = \sum_{t=1}^{T} Q_t \mathbb{1}(Q_t \geq Q_{\theta_{95}}) \tag{6}$$

$$Q_{LowFlow}(Q) = \sum_{t=1}^{T} Q_t \mathbb{1}(Q_t \leq Q_{\theta_5}) \tag{7}$$

where $Q_{Flood}$ and $Q_{LowFlow}$ are the summation of simulated streamflows, $Q_t$, over $T$ time steps that meet the flooding and low flow thresholds defined by $Q_{\theta_{95}}$ and $Q_{\theta_5}$, the 95th and 5th percentiles, respectively. The indicator function, $\mathbb{1}$, is equal to one when $Q_t$ meets the threshold and zero otherwise. $Q_\theta$ is the streamflow from a hydrologic model parameterization, $\theta$, without GI added. Therefore, the values of these quantile thresholds vary for each model parameterization. The Synthetic optimization uses the synthetic true parameterization, the MAP optimization uses the MAP parameterization, and the MORO and MinMax optimizations use several parameterizations,

with each objective computed relative to the $Q_\theta$ for that parameterization.

The flooding and low flow objectives describe the relative change caused by adding GI, as shown in Equations 8 and 9

$$J_{Flood,\theta} = \frac{Q_{Flood}(Q_{GI}) - Q_{Flood}(Q_\theta)}{Q_{Flood}(Q_\theta)} \tag{8}$$

$$J_{LowFlow,\theta} = \frac{Q_{LowFlow}(Q_{GI}) - Q_{LowFlow}(Q_\theta)}{Q_{LowFlow}(Q_\theta)} \tag{9}$$

where $J_{Flood,\theta}$ and $J_{LowFlow,\theta}$ are the flooding and low flow objectives for parameterization $\theta$, respectively, and $Q_{GI}$ is the simulated streamflow with GI added. Finally, we consider the GI cost to be directly proportional to the number of trees planted, $N_{GI}$, as shown in Equation 10.

$$J_{Cost} = N_{GI} \tag{10}$$

The cost is the same for all hydrologic model parameterizations.

**Decision Variables: Possible Reforestation Locations**

Our example study uses the Baisman Run watershed, whose land cover and possible GI reforestation locations are provided in Fig. 2. The lower two-thirds of the watershed is primarily undeveloped deciduous forest, while the upper third of the watershed has exurban development (Fig. 2A). There are two power lines that appear as Southwest-Northeast trending linear grass cover. Baisman Run is located about 20 km North-Northwest of Baltimore, Maryland, USA and is part of the larger Chesapeake Bay watershed. Baltimore County plans to have 10% of runoff drain through GI by 2030 (Pennino et al. 2016), which partly motivates the choice of this watershed.

We consider possible GI locations to be areas with at least 9 $m^2$ square, contiguous grass cover within a 900 $m^2$ model pixel (Fig. 2B). Additional constraints placed on the GI locations include: 1) buffers around power lines (15 m), major roads (8 m) (O'Neil-Dunne and Grove. 2004), and any existing infrastructure (2 m) based on local guidelines, and 2) a minimum of 150 $m^2$ total

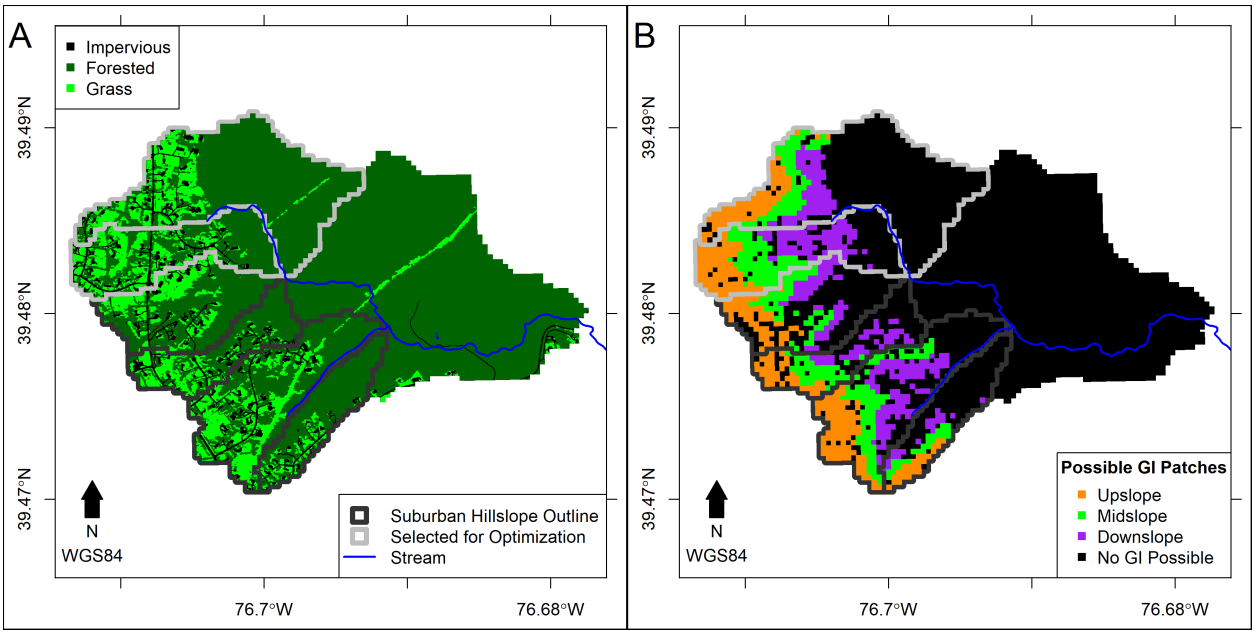


**Fig. 2.** The Baisman Run watershed near Baltimore, Maryland, USA showing (A) 1 $m^2$ resolution land cover from 2013-2014 (Chesapeake Conservancy 2017) and (B) 900 $m^2$ modeling resolution showing possible green infrastructure (GI) reforestation locations. The exurban sub-watershed selected for our optimization experiment is outlined in light gray.

area for septic system leach fields (model pixels near houses) based on typical water loadings for a five-person household (InspectAPedia 2024).

While we could use every model pixel with a possible GI location as a decision variable in the optimization, it would likely be challenging to implement any discovered set of GI locations at the modeled spatial scale due to factors not considered in the optimization (e.g., local buy-in from homeowners, permitting, etc.). Additionally, we thought the computation time needed to run a physically-based spatially distributed hydrologic model in an optimization with hundreds of decision variables could result in a Pareto front that would not converge within our resource limits. So, we decided to reduce the optimization domain to the upper two hillslopes (gray outlines in Fig. 2), which have similar relative land cover proportions as the full watershed. We further simplified the number of decision variables using elevation-based quantile cutoffs that approximate upslope, midslope, and downslope areas (Fig. 2B). These areas represent different local ecoregions within which the effectiveness of reforestation is expected to differ (Hung et al. 2020; Fan et al. 2019).

The optimized decision variables are the proportion of the area to reforest in each of the hillslope ecoregions, resulting in a total of six decision variables. This formulation of the decision variables allows for more flexibility to implement the resulting solutions. However, using this approach, the same combination of decision variables could lead to different objective values. We evaluated the uncertainty in flooding and low flow objectives (Equations 8 and 9) that results from 100 random allocations of GI within each of the ecoregions using several hypothetical decisions (20%, 40%, 60%, and 80% GI; SI Fig. S2) and account for uncertainty in these objectives in the optimization using an "epsilon tolerance" level (Laumanns et al. 2002; Hadka and Reed 2013) on the objectives. The tolerance level describes the difference in objective values that is practically similar, in this case due to uncertainty in the GI locations. Based on this experiment, we set the epsilon tolerance of the flooding and low flow objectives to 0.015. We also used an epsilon tolerance of one tree for the cost objective (Equation 10).

**Hydrologic Model Description: RHESSys**

We used the Regional Hydro-Ecologic Simulation System (RHESSys) to simulate streamflows across the watershed (Tague and Band 2004). RHESSys provides spatially distributed water routing and uses coupled physically-based process models of water, carbon, and nitrogen cycles within vegetation and soil. RHESSys models forest stands (areas) instead of individual tree stems, so our reporting of results as number of trees (9 $m^2$ canopy areas) reforested is for convenience and may not accurately reflect the total number of trees in a forested area. We used a version of RHESSys adapted for urban watersheds that includes water routing for storm drains (Lin 2019b; Lin 2021). We used GIS2RHESSys (Lin 2019a) to process spatial data to the RHESSys model grid. The workflow that we used for running GIS2RHESSys and RHESSys on the University of Virginia's Rivanna HPC cluster is provided in the code repository (Smith 2021).

We used the "static" vegetation mode of RHESSys based on the findings reported in Smith et al. (2022) that the "dynamic" mode can generate unrealistic vegetation growth and mortality for randomly sampled parameter values that could occur in our Bayesian calibration. We also relied on the findings of a spatial global sensitivity analysis of RHESSys within the Baisman Run watershed

in Smith et al. (2022) to inform the selection of parameters to calibrate in this study's two hillslopes. A total of 19 parameters were selected, as shown in Table 1, which also gives the bounds of the uniform distribution used in the Bayesian calibration and references for those bounds. Hillslope and zone parameters control processes over the entire modeling domain (e.g., groundwater and atmospheric properties), while land use, vegetation, and soil parameters are specified uniquely by location. To reduce the number of parameters to calibrate, we modeled all forests that were not GI as broadleaf deciduous trees and did not consider specific tree species and their composition across the watershed (Lin et al. 2019, e.g.,). We split soil parameters into undeveloped or developed to account for possible compaction. A supplementary table (SI S3) provides all other RHESSys parameter values and structural process equation choices that we used.

*Bayesian Calibration*

We assume that an observed streamflow time series can be predicted by a model of the form shown in Equations 11-12:

$$\mathbf{Q} = \hat{\mathbf{Q}}\boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{11}$$

$$\mu_t = \exp\left(\mu_b \hat{Q}_t\right) \tag{12}$$

where $\mathbf{Q}$ is the vector of observations over all time steps, $t$, $\hat{\mathbf{Q}}$ is the vector of simulated streamflows from the hydrological model, $\boldsymbol{\mu}$ is the vector of bias corrections defined by parameter $\mu_b$, and $\boldsymbol{\epsilon}$ is the vector of residuals. We chose to model the residuals using a generalized normal distribution, called the skew exponential power (SEP) distribution, with a heteroskedasticity and lag-1 autoregressive adjustment (Schoups and Vrugt 2010). The normal distribution is a special case of the SEP distribution, which allows for skewness and kurtosis to deviate from normal, as is common in hydrologic model simulations. The adjusted residuals follow a standard SEP distribution (center of zero and scale of 1), shown in Equation 13

**TABLE 1.** Calibrated RHESSys parameters and uniform priors. Categories: L - land use, H - hillslope, Sd - developed soil, Su - undeveloped soil, Z - zone, V - vegetation.

| Parameter | Description | Lower Bound | Upper Bound |
|---|---|---|---|
| L: Septic Water Load | Septic system water load ($m/m^2/yr$) (InspectAPedia 2024) | 0 | 2 |
| H: GW Loss Coef. | Proportion of groundwater storage lost to the stream (Lin 2021) | 0 | 0.3 |
| Sd: Ksat | Saturated hydraulic conductivity at surface (m/d) (Dingman 1994; García-Gutiérrez et al. 2018) | 0.1 | 5 |
| Sd: Soil Thickness | Maximum soil depth (m) (Froelich 1975; Reybold and Matthews 1976) | 1 | 30 |
| Sd: m | Exponential decay parameter for saturated hydraulic conductivity with increasing saturation deficit. (Lin 2021) | 0.03 | 1.5 |
| Sd: Air Entry Pres. | Air entry pressure (m water) (Dingman 1994; Clapp and Hornberger 1978) | 0.15 | 1.5 |
| Sd: Sat. to GW Coef. | Proportional amount of water moving from the saturated soil storage to the groundwater storage (Lin 2021) | 0 | 0.3 |
| Sd: vKsat | Vertical saturated hydraulic conductivity at surface (m/d) (Dingman 1994; García-Gutiérrez et al. 2018) | 0.05 | 5 |
| Sd: Pore Size | Soil pore size index (Dingman 1994; Clapp and Hornberger 1978) | 0.11 | 0.606 |
| Su: Ksat | Saturated hydraulic conductivity at surface (m/d) (Dingman 1994; García-Gutiérrez et al. 2018) | 0.3 | 7 |
| Su: Soil Thickness | Maximum soil depth (m) (Froelich 1975; Reybold and Matthews 1976) | 1 | 30 |
| Su: m | Exponential decay parameter for saturated hydraulic conductivity with increasing saturation deficit (Lin 2021) | 0.05 | 2 |
| Su: Air Entry Pres. | Air entry pressure (m water) (Dingman 1994; Clapp and Hornberger 1978) | 0.15 | 1.5 |
| Su: Sat. to GW Coef. | Proportional amount of water moving from the saturated soil storage to the groundwater storage (Lin 2021) | 0 | 0.3 |
| Su: vKsat | Vertical saturated hydraulic conductivity at surface (m/d) (Dingman 1994; García-Gutiérrez et al. 2018) | 0.2 | 7 |
| Z: Wind speed | Wind speed (m/s) (NREL 2012) | 0 | 5 |
| V: Tree Max. Stomatal Cond. | Maximum stomatal conductance (m/s) (Lin 2021; Leonard et al. 2017; Miles 2015a; Miles 2015b) | 0.001 | 0.01 |
| V: Tree Stomatal Fraction | Stomatal fraction (Lin 2021; Leonard et al. 2017; Miles 2015a; Miles 2015b) | 0.3 | 1 |
| V: Tree Rainwater Capacity | Interception storage capacity (m/leaf area index) (Muleta and Nicklow 2005) | $2.5E^{-4}$ | $7.5E^{-4}$ |

$$a_t = \frac{\epsilon_t - \epsilon_t \phi_1}{\sigma_t} \sim \text{SEP}(0, 1, \xi, \beta) \tag{13}$$

where $\sigma_t$ is the heteroskedasticity adjusted standard deviation of the residuals at time $t$ that is described by a linear function of $\hat{\mathbf{Q}}$ with intercept and slope parameters $\sigma_0$ and $\sigma_1$, $\xi$ and $\beta$ respectively are the skewness and kurtosis parameters of a standard SEP distribution, and $\phi_1$ is the lag-1 autoregressive parameter. The log-likelihood for these six parameters relies on several statistics of these parameters, as described in Schoups and Vrugt (2010) and shown in Equations 14 and 15:

$$\mathcal{L}(\boldsymbol{\Theta}, \boldsymbol{\eta} = \{\sigma_0, \sigma_1, \xi, \beta, \phi_1, \mu_b\} | \mathbf{Q}) = n \log \frac{2\sigma_\xi \omega_\beta}{\xi + \xi^{-1}} - \sum_{t=1}^{n} \log \sigma_t - c_\beta \sum_{t=1}^{n} |a_{\xi,t}|^{2/(1+\beta)}, \tag{14}$$

$$a_{\xi,t} = \xi^{-\text{sign}(\mu_\xi + \sigma_\xi a_t)} (\mu_\xi + \sigma_\xi a_t) \tag{15}$$

where $\mu_\xi$ and $\sigma_\xi$ are the mean and standard deviation of $\xi$, and $\omega_\beta$ and $c_\beta$ are standard exponential power parameters that are functions of $\beta$. Please refer to Schoups and Vrugt (2010) for a complete derivation of this equation.

Using the priors in Table 1 and the likelihood in Equation 14, we use Bayesian calibration to estimate the joint posterior distribution of hydrologic model parameters as shown in Equation 16

$$p(\boldsymbol{\Theta}, \boldsymbol{\eta} | \mathbf{Q}) \propto \mathcal{L}(\boldsymbol{\Theta}, \boldsymbol{\eta} | \mathbf{Q}|) p(\boldsymbol{\Theta}, \boldsymbol{\eta}) \tag{16}$$

where $p(\boldsymbol{\Theta}, \boldsymbol{\eta} | \mathbf{Q})$ is the joint posterior distribution of hydrological model parameters and residual error model parameters and $p(\boldsymbol{\Theta}, \boldsymbol{\eta})$ is the prior distribution for both sets of parameters. Priors for $\boldsymbol{\eta}$ are provided in SI Table S1. To limit the number of parameters to calibrate, we simplified the estimation of $\boldsymbol{\eta}$ by using maximum likelihood estimation (MLE) for each candidate set of $\Theta$ values that was proposed during our Bayesian calibration routine. The MLE we used was slightly modified from the “generalizedLikelihoodFunction" in the *spotpy* Python package (Houska et al.

2015) and is provided in our code repository. We used Sequential Least SQuares Programming (SLSQP) (Kraft 1988) in Python's `scipy` library with 500 starting locations and selected the $\boldsymbol{\eta}$ values that provided the maximum likelihood.

For Bayesian calibration, we used the DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm with archive (z) and snooker update (s) adaptations ($\text{DREAM}_{\text{zs}}$) (Laloy and Vrugt 2012) for Markov Chain Monte Carlo (MCMC) estimation of the posterior distribution. DREAM has been shown to perform well for high-dimensional and multi-modal parameter estimation problems (Kavianihamedani et al. 2024), which are common in hydrology. We used 4 sets of 10 independent chains for a total of 48,000 posterior samples. About a third of the total samples were used to adapt the transition probabilities used within the DREAM algorithm, and nearly all those samples were removed using the NSE-based burn-in of the chains. Other $\text{DREAM}_{\text{zs}}$ hyperparameter settings used the recommended default settings in the BayesianTools R package (Hartig et al. 2017). We made minor edits to this package to work for our problem and provide the edited code in our code repository. A table of $\text{DREAM}_{\text{zs}}$ hyperparameter values is provided in SI Table S2.

For all optimizations except the Synthetic optimization, we computed the flooding and low flow objectives as the mean value from 1000 randomly sampled realizations of the residual errors, $\boldsymbol{\epsilon}$, using the estimated $\boldsymbol{\eta}$. We assume that the relationship between model parameters and error model parameters is stationary with changing landscape (GI), even though land cover change may alter the error distribution. Future work could explore conditioning the error model on such landscape characteristics.

*Generation of Synthetic True Observed Streamflow*

We randomly generated the synthetic true parameterization as an independent sample of each parameter from within the bounds shown in Table 1. We then ran RHESSys with observed precipitation from Oregon Ridge Park (Welty and Lagrosa 2018) and NOAA Global Historical Climatology Network Daily temperature data from the Maryland Science Center station (obtained using the R package *rnoaa* (Chamberlain 2019)) from 1999-11-15 to 2013-09-30 to generate a synthetic true streamflow time series for the outlet of our modeled portion of the watershed. This

period experienced several large precipitation events for which flood control could be evaluated. We consider about five years (1999-11-15 to 2004-09-30) as a spin-up period during which RHESSys internal states can reach an equilibrium, and we used the remaining portion of the time series for calibration.

In model calibrations that use real observation data, model parameterizations may not be able to exactly replicate an observed time series, so we generated the final synthetic observed time series by adding error to the modeled time series. We used the SEP distribution in Equation 13 to generate normally distributed autoregressive error with a 0.7 lag-1 correlation, no bias, no heteroskedasticity, and a constant standard deviation $\sigma_0$ of 0.05. We added one randomly sampled error time series to the RHESSys simulated time series to get the final synthetic true observed time series, $\mathbf{Q}$, for the Bayesian calibration.

## RESULTS

### Calibrated Parameter Sets for Optimization Experiments

The Bayesian estimated marginal distributions for two example RHESSys parameters are provided in Fig. 3 and distributions for other RHESSys parameters and likelihood parameters are provided in SI Figures S3 and S4. Most of the maximum a-posteriori (MAP) estimates are close to the synthetic true parameter values and have posterior distributions that have narrowed in range to more likely values, (e.g., septic water load in Fig. 3A). Other posterior distributions did not change much from their priors or have multiple modes (e.g., tree rainwater capacity in Fig. 3B). For the vegetation parameters, differences from the synthetic true value could compound uncertainty in modeled streamflows as the GI optimization adds evergreen trees to the watershed that interact with the broadleaf deciduous trees. This effect is considered by the robust optimization approaches but not by the MAP approach. For example, the eight parameterizations used by our robust optimizations span the range of likely values for all parameters and sample multiple modes of the posterior distributions (Fig. 3). These results demonstrate that hydrologic model calibration can lead to complex posterior distributions, which further justifies our comparison of traditional and robust optimization approaches.

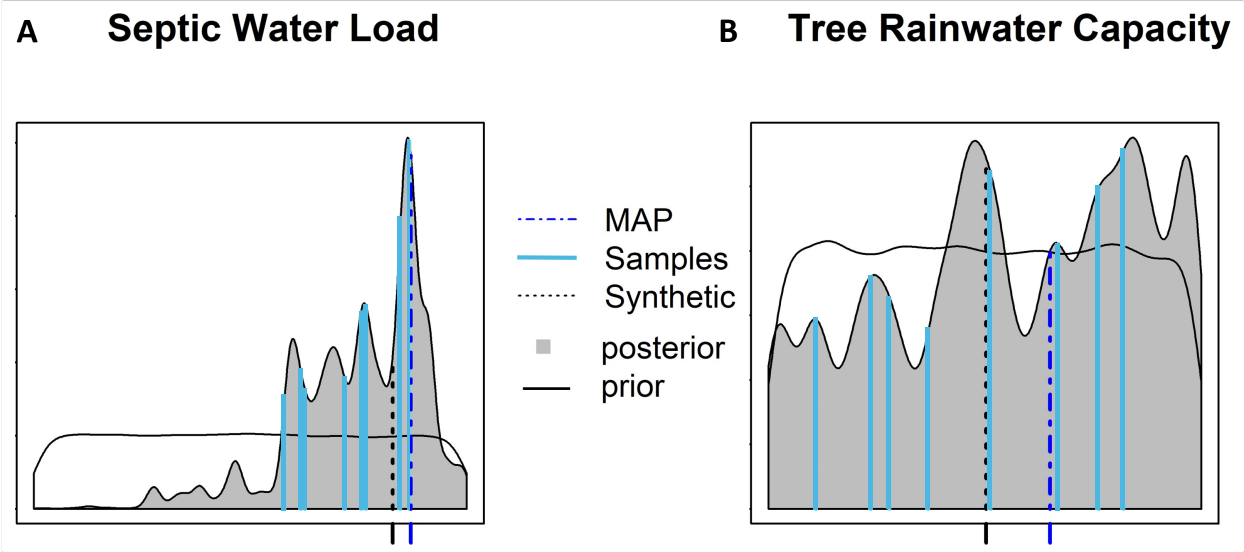


**Fig. 3.** Marginal posterior distributions of (A) septic water load and (B) tree rainwater capacity, two RHESSys parameters estimated from the DREAM$_{ZS}$ method. Distributions are overlain with the synthetic true parameter values, the maximum a-posteriori (MAP) estimate and the other eight parameterizations that are selected for the robust optimization experiments. The uniform priors are shown for reference.

## Optimization Results

The Pareto fronts for the four optimization experiments are provided in Fig. 4. The first row (A-D) shows the objective values of GI solutions from each approach when evaluated on the parameterizations to which they were optimized. The second row shows the objective values of the solutions from the MAP and robust optimizations when evaluated on the synthetic true parameter set. Hypothetical compromise thresholds of >20% flood reduction and <-20% low flow change are shown, with solutions to the upper right of these achieving objective values that meet both thresholds. Histograms of the number of trees planted by solutions meeting these thresholds in Fig. 4A-D are shown in Fig. 4E.

As expected, increasing the number of trees reforested generally results in a decrease in flooding and low flows. All four optimizations show an almost piecewise linear relationship between the flooding and low flow objectives, with an inflection point and higher variability of solutions around 30-40% flood reduction. The Synthetic optimization shows that there are some solutions that provide up to about 10% reduction in flooding while also increasing low flows for small amounts

of trees reforested (Fig. 4A). The MAP, MORO, and MinMax methods also find solutions in this space when evaluated on the synthetic parameterization, but not in their original space (Fig. 4B-D,F-H). Similar changes in perceived objective values are seen with respect to the compromise thresholds. Each of the MAP, MORO, and MinMax approaches estimate that more solutions meet these thresholds than using the Synthetic optimization (Fig. 4E). The MinMax finds the fewest compromise solutions, yet they are all concentrated around the solutions found for the Synthetic optimization, suggesting these may provide the most accurate objective estimates on this problem. The MORO and MAP approaches have higher variability around the solutions found for the Synthetic approach, with the MORO solutions being slightly more concentrated around the number of trees found by the Synthetic optimization. Therefore, the two robust optimization approaches appear better than the MAP approach for matching the total number of trees needed to meet these compromise thresholds (lower variance and bias).

**Analysis of Pareto Solutions**

Among solutions found to be within the compromise region based on their respective definitions of the objectives, we investigate where these solutions allocate trees within the watershed (Fig. 4). The MAP, MORO, and MinMax approaches all show the downslope areas as the least forested, which is in agreement with the Synthetic optimization. The MAP, MORO and MinMax approaches correctly allocate trees to the northern midslope area, and over-allocate the southern midslope area compared to the Synthetic optimization. For upslope areas, the MinMax optimization gets the correct spatial allocation while the MORO approach flips the North-South allocation of trees compared to the Synthetic approach; however, the total amount of trees allocated to upslope areas is similar for both robust optimizations and the Synthetic optimization. The MAP optimization over-allocates trees to upslope areas. Based on these results, the allocation of trees to upslope, midslope and downslope areas among compromise solutions from the robust optimizations are also more optimal than under the MAP optimization.

Switching to evaluations of the objectives, Fig. 6 shows the distributions of the change in objective values between the optimization and evaluation on the synthetic truth across solutions

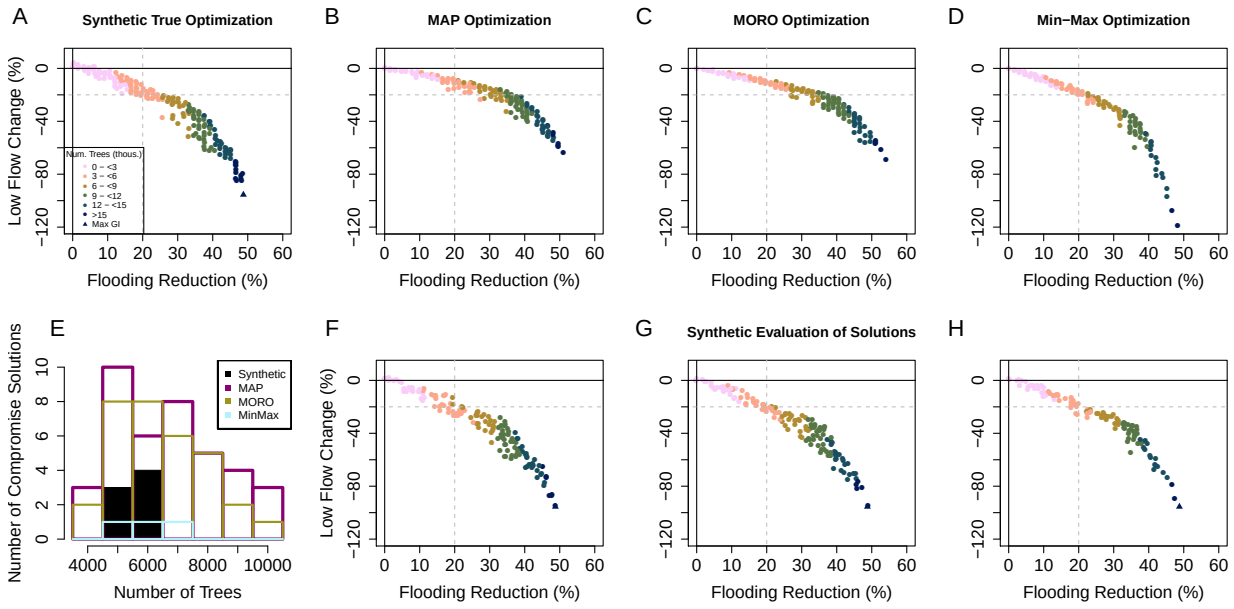


**Fig. 4.** (A-D) Pareto fronts for the four optimization experiments. (F-H) The Pareto solutions in panels B-D evaluated on the synthetic true parameterization. For reference, solid lines are shown in each panel at objective values of 0% low flow change and 0% flood reduction. Gray dashed lines on these figures show the locations of example compromise thresholds for the flooding and low flow objectives. The solution with the maximum possible GI is plotted with a triangular symbol. Note that the y-axis shows flow increases as positive values and decreases as negative values. (E) Histograms of the number of trees (9 $m^2$ areas) planted by solutions that meet compromise thresholds in panels A-D. The MAP and MORO lines are transparent to see where both are present.

from each formulation. The MinMax approach provides objective values that are most similar to the values obtained when the solutions are evaluated with the synthetic true parameterization; however, it is also the most skewed for low flows. The skew is in the direction of being robust to less favorable objective values, which makes intuitive sense because the MinMax approach represents the worst case possibility for each decision. The MAP and MORO approaches have similar variability and they are notably biased compared to the MinMax approach. The MAP approach is slightly less biased on the flooding objective than the MORO approach, and the MORO approach is slightly less biased on the low flow objective. Both MAP and MORO tend to estimate that more flood reduction and less low flow reduction will occur than actually would occur, which could lead to under designing the system for flood control and underestimating the impact of GI on low flows. Based on this analysis, risk-averse decision makers would likely prefer the MinMax

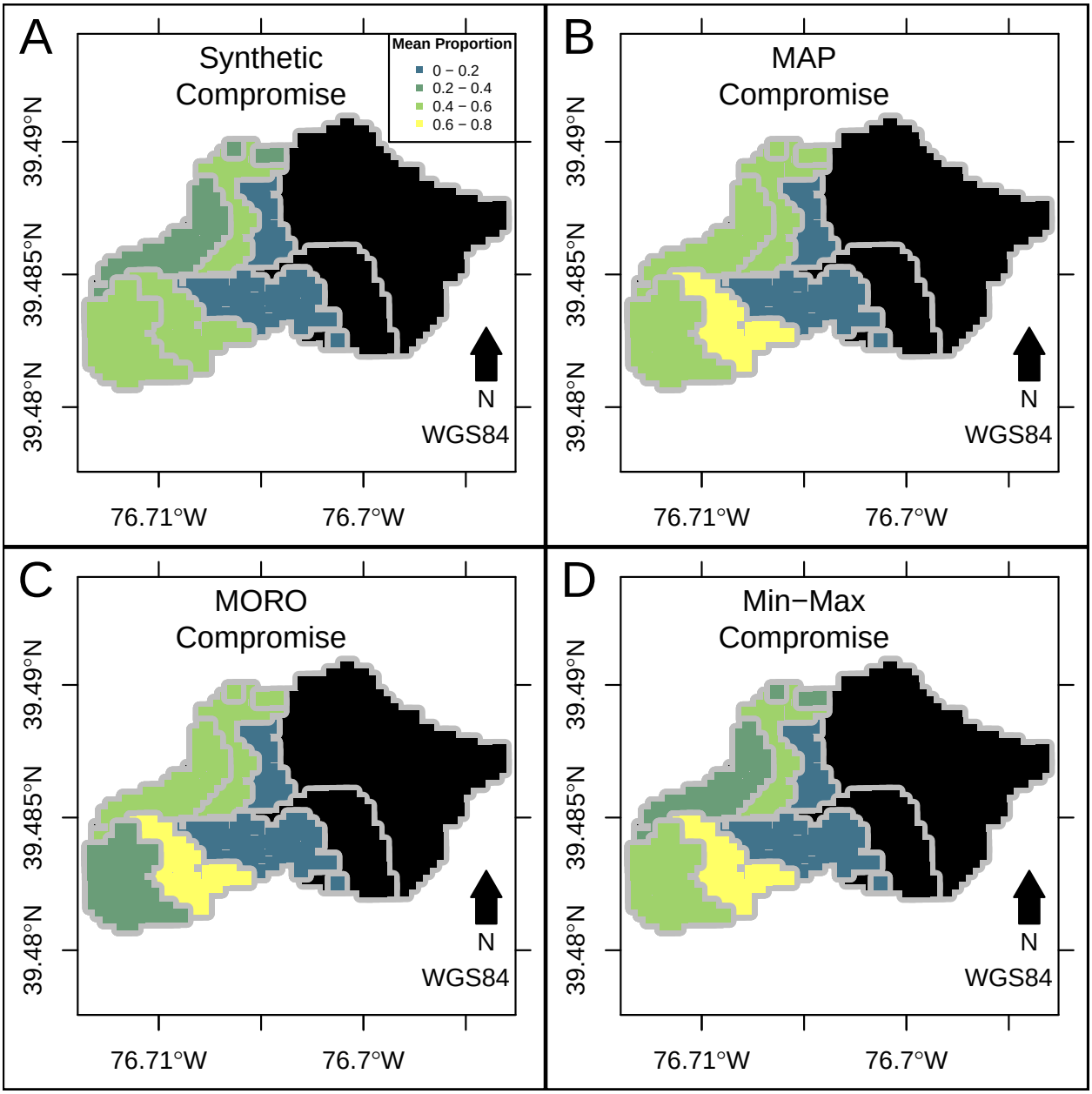


**Fig. 5.** The average proportion of each decision variable's area that is reforested for all Pareto solutions that meet the compromise thresholds in Fig. 4A-D. Black areas were not considered as possible reforestation locations. Please refer to Fig. 2B for reference to upslope, midslope and downslope areas.

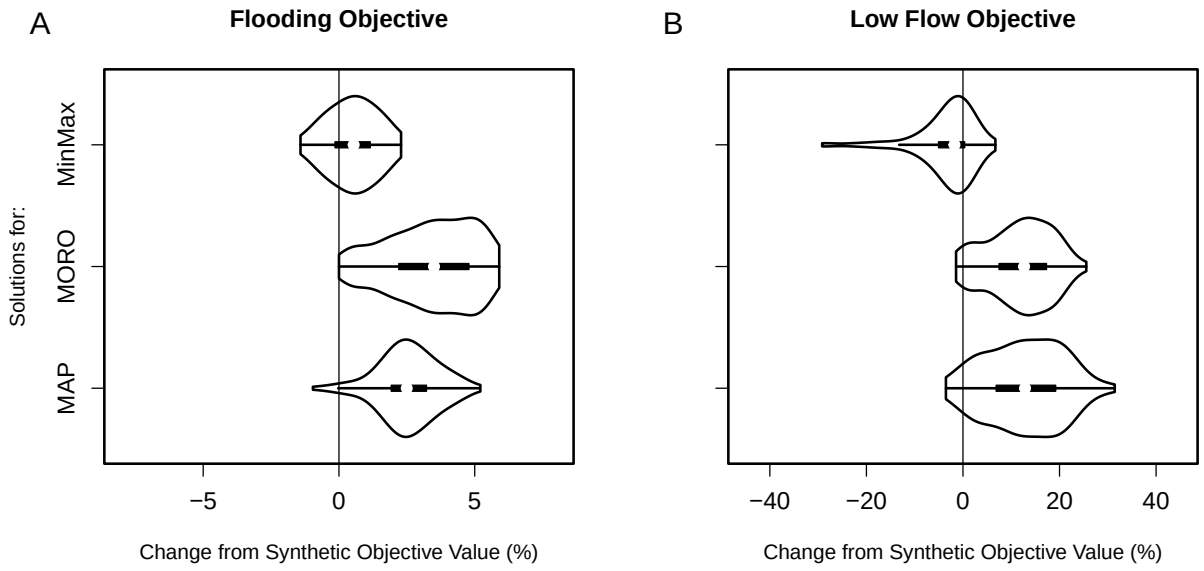


**Fig. 6.** Change of the flooding and low flow objective values when the Pareto solutions are evaluated using the synthetic parameterization. Decreases indicate more conservative worse-case estimates (less flood reduction, more low flow reduction).

approach because it provides solutions that are less variable and less biased than the MORO and MAP approaches, even though the MinMax approach could lead to overestimating the reduction of low flows.

To further understand the MinMax approach, Fig. 7A and 7B show which hydrologic model parameterizations provided the worst-case flooding and low flow objective values for each Pareto solution. The parameterizations differ by Pareto solution and show a pattern with increasing number of trees (increased flood reduction). This finding suggests that we could utilize perceptual hydrological models of the system (McMillan et al. 2023) and select parameterizations for robust optimizations based on the effects parameterizations have on specific hydrological processes. We see in Fig. 7C that using our process-agnostic clustering approach to select parameterizations provides unique parameter combinations that could reflect different impacts on hydrological processes.

For the flooding objective (Fig. 7A), parameterizations P1 (the MAP) and P9 provide the majority of worst-case values. These parameterizations have different effects on flood flows. Soils

in P1 are thick, have low vertical and lateral hydraulic conductivity and small pore sizes (Fig. 7C). These characteristics could lead to low soil infiltration rates that could result in quick runoff and flood generation during precipitation events. The septic water load is also high for P1, which provides a higher baseflow value and wetter soils that could increase the chance of flooding. However, the source of septic water is assumed to be from deep groundwater (a well), which could limit contributions of deep groundwater to baseflow and offset the greater available shallow soil water impacts on saturation runoff. This may explain why P9 provides slightly more worst-case flood objective values than P1. P9 also has the lowest tree rainwater capacity across the nine parameterizations, which would lead to less interception and more water available as runoff. This result may explain why, as more trees are added, P9 is more commonly providing the worst-case flood objective values.

For the low flow objective (Fig. 7B), parameterization P7 provides the majority of worst-case values, followed by P4, P3 and P8. Parameter values in P7 tend to be located in the extremes of the posterior distribution and these extreme values have subsequent effects on low flows (Fig. 7C). For example, there is a low contribution from septic loads that lowers the baseflow and soil moisture, a low saturation to groundwater coefficient that allows water to be retained in soils and available for evapotranspiration, and a high groundwater loss coefficient that quickly transports water out of the system. P4 also has a high groundwater loss coefficient, low saturation to groundwater coefficient, and low tree rainwater capacity, but these parameter values are not as extreme as P7. Solutions where P3 and P8 provided the most reduction in low flows had a small amount of trees.

**Impact of Optimization Approach on Making Informed Decisions**

We also evaluate the discovered Pareto optimal objective values as a function of the decisions made. Each of the optimization approaches show remarkably similar trends in how the allocation of trees affects the low flow and flooding objectives, and in how the trees are allocated to upslope, midslope and downslope areas as the total number of trees increases (Fig. 8). Initially, there is a preference to allocate tress to midslope and then upslope areas (Fig 8B), which provides the largest reduction in flood flows for the least reduction in low flows (slopes of lines in Fig. 8A). Upslope

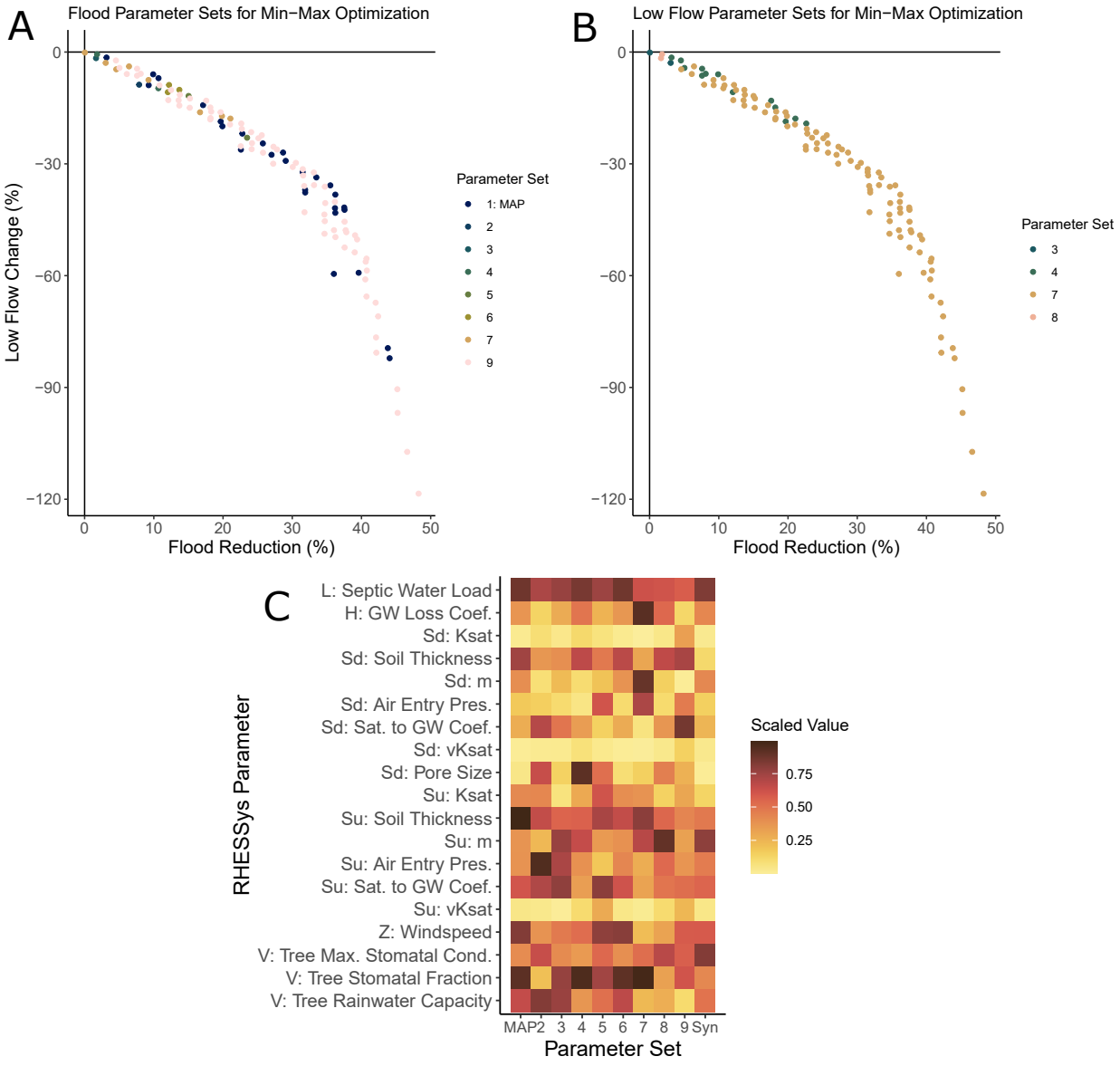


**Fig. 7.** MinMax Pareto solutions colored by the parameterization that provided the worst objective values: (A) least flood reduction, and (B) most low flow reduction. Parameterization colors are the same in each panel. (C) The relative parameter values for each parameterization, to aid interpretation of the results.

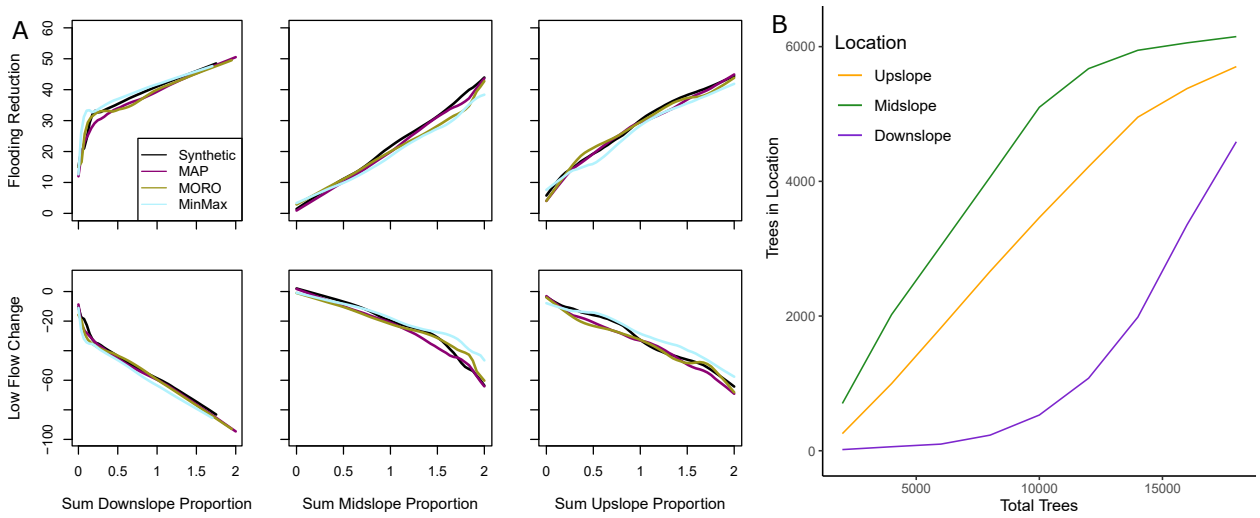


**Fig. 8.** A: Flooding and low flow objectives for Pareto solutions as a function of the GI reforestation decisions. B: Total number of trees (9 $m^2$ areas) by location in the watershed, aggregated over all Pareto solutions from all models.

areas are reforested in a nearly constant proportion across Pareto solutions, as indicated by a constant slope in Fig. 8B. Adding trees to downslope areas begins after about 5000 trees are reforested, at the expense of a greater relative reduction in low flows (inflection point in low flows in Fig. 8A). After about 10000 trees are planted across the midslope and upslope areas, the midslope areas nearly reach their full capacity and downslope areas begin to be reforested in higher proportions. As a result, the relative reduction in low flows becomes more extreme and the relative reduction in flood flows is reduced. These results suggest that any of these optimization approaches would be able to discover that reforesting in midslope, then upslope, and then downslope areas could provide the best compromise between flood reduction and low flow exacerbation, but robust approaches, particularly MinMax, better estimate the resulting impact of reforestation on streamflows.

## DISCUSSION

Based on the above results, the robust optimization approaches, particularly the MinMax approach, seem better than the MAP approach for estimating objectives and decisions. Here, we explore the key components of the multi-objective robust optimization approaches and provide guidance on extensions of this work. The key components of the multi-objective robust optimization approaches in this study are: 1. a model that provides uncertainty estimates, 2. a method to select

realizations from the model's uncertainty distribution for use in the robust optimization, and 3. a method to aggregate results from each realization to compute the objectives.

1. We used an ecohydrological process model to simulate hydrologic flows, and relied on Bayesian model calibration to estimate a posterior distribution from which streamflow realizations could be generated. We assumed that the relationship between hydrologic model parameters and error model parameters was stationary with a changing landscape (GI), and future work could relax this assumption by conditioning the error model on such landscape characteristics, similar to the dynamic error model introduced by Brodeur et al. (2024) to capture non-stationary model errors under climate change. To do so would likely require calibrating the model on a larger number of sites in order to learn the effect of landscape change on the model outputs and their errors. Process models with long runtimes may become less feasible as larger areas or multiple watersheds need to be calibrated, so adopting machine learning (ML) models or integrating an ML solver for the process model (Tsai et al. 2021) may be advantageous. For example, directly estimating parameters of a probability distribution for streamflow, as in mixture density neural networks (Klotz et al. 2022), could allow for training on a large number of sites while learning the relationship between changing catchment characteristics and hydrological errors simultaneously.

2. Robust optimization for a select few realizations from the uncertainty distribution helps reduce computational complexity. Our diagnostics of the MinMax approach suggest that the set of uncertain model parameterizations to use in robust optimizations can be based on the effects each parameterization would have on hydrological processes that affect the study objectives. This could be informed by clustering to discover similar parameterizations, as in this study, diverse model calibration (Williams et al. 2020), or model sensitivity analyses for the retrospective and future periods of interest for the design life of the infrastructure (Karimi et al. 2022).

3. We explored a likelihood weighted approach (MORO) and min-max (MinMax) approach to aggregate results. We found that the MORO approach tended to estimate more flood reduction and less low flow reduction than actually would occur, which could lead to under-designing the system for flood control and underestimating the impact on low flows. These results are similar to

not using robust optimization, and suggest that using a MORO approach that weights the objective values towards the most likely model parameterization may not be preferable for the discovery of robust solutions. Comparing the solutions discovered from a MinMax approach and a likelihood-weighted approach could be good practice for informed decision making. For example, in this study, we found that each of the optimization approaches showed remarkably similar trends in how the allocation of trees affects the low flow and flooding objectives, and in how the trees are allocated to upslope, midslope and downslope areas as the total number of trees increases. However, only the MinMax approach provided unbiased objective values relative to synthetic true values, and had low variability. Comparing solutions from multiple optimization experiments is further supported by each method (MAP, MORO, and MinMax) contributing approximately equal numbers of non-dominated Pareto-optimal solutions to a combined Pareto front of all discovered solutions (SI Figure S5).

We provide a simple multi-objective optimization example to evaluate the potential value of considering robustness of water systems designs to hydrological model uncertainties. Our findings suggest that considering robustness can lead to better decisions and better estimation of their impacts on hydrological flows. Therefore, this study supports additional studies on more complex problems. For example, while we focus on only one type of GI in this study, the multi-objective robust optimization methods can be used with other types of green and gray infrastructure (e.g., (Giacomoni and Joseph 2017)) or for other types of water resources systems problems. There are also opportunities to include additional objectives for co-benefits of GI or other infrastructure that we did not consider, such as water quality and societal benefits (Zhang et al. 2023; Zhang et al. 2020; van Meerveld et al. 2021). Our watershed did not have a variety of demographic groups, so considering a social vulnerability or environmental justice objective would not have been meaningful, but other studies have found that including such objectives can achieve similar hydrological benefits as approaches that ignore them (Herbst et al. 2023). Uncertainties for models that simulate each of the considered objectives (e.g., water quantity and quality models) can be integrated within the robust optimization. Uncertainty in future changes like climate and land

use over time could also be considered as components of scenarios within a robust optimization (Webber and Samaras 2022). To date, there are few studies that utilize multi-objective robustness to deep uncertainty in GI optimizations (Piscopo et al. 2021) and no known studies that additionally integrate hydrologic model uncertainties into such optimizations (Tebyanian et al. 2022). Future research could explore when, where, and for what types of water systems planning and management problems the integration of hydrologic model and future scenarios (e.g., climate, land use) are important to consider for more robust designs.

Regarding the example application, our finding that the preferred reforestation locations for flood control are midslope and upslope areas before downslope areas is consistent with recent studies that evaluate the impact of deforestation on flooding (Hurtado-Pidal et al. 2022), yet is counter to implementation guidelines (Bentrup 2008) that suggest downslope riparian buffers may provide better flood control. Reducing upslope runoff contributions during major storm events limits the dynamic contributing area of the watershed, which should reduce flooding at the watershed outlet, as we found. Downslope areas will generally be more saturated and more prone to contribute storm runoff, even if they have a forested riparian buffer. Therefore, reforesting downslope areas may be less effective for flood control than reforesting midslope and upslope areas.

## CONCLUSIONS

Water resources planning and management studies typically do not consider the uncertainty in hydrologic models to inform multi-objective engineering designs, and the few examples that are available in literature do not demonstrate if considering such uncertainty results in better decisions. We explored the question *Do optimization approaches that are robust to modeled streamflow uncertainties provide better solutions than approaches that ignore streamflow uncertainties?*. We found that the robust methods, particularly the MinMax method, do provide better objective values and decisions compared to methods that only consider the single most likely model parameterization. Furthermore, we demonstrated that ignoring model uncertainty could lead to under-performance of designs for flood control and worse than expected low flows. These results suggest that integrating model uncertainty into water systems planning and management decision support analyses can be

beneficial for discovering more robust engineering designs and operations.

## DATA AVAILABILITY STATEMENT

All data, models, and code generated during the study are available in a repository online in accordance with funder data retention policies (GitHub and Zenodo code repository: (Smith 2024); HydroShare data repository (should be accessible to reviewers and we will assign a DOI when the article is accepted): https://www.hydroshare.org/resource/0f6be4446d0f47519753d85d0e908d42). The multi-master Borg algorithm is freely available to academic and non-commercial users and may be requested from http://borgmoea.org/#download. The "passNFE" branch of the code repository was used for this study.


## ACKNOWLEDGMENTS

The authors thank Laurence Lin and members of the Quinn research group for constructive feedback on this work. The authors acknowledge Research Computing at The University of Virginia for providing computational resources and technical support that have contributed to the results reported within this publication. URL: https://rc.virginia.edu. Data were supported by the Baltimore Ecosystem Study.


## SUPPLEMENTAL MATERIALS

Word document SI.pdf provided with this submission contains supplemental figures (S1-S5) and tables (S1-S2) that are referenced within the manuscript. A supplemental table named S3_Tables_RHESSysParameterDescriptions.xlsx contains definitions of all RHESSys parameters and is provided in the data repository.

## REFERENCES


Bartholomew, E. and Kwakkel, J. H. (2020). "On considering robustness in the search phase of robust decision making: A comparison of many-objective robust decision making, multi-scenario many-objective robust decision making, and many objective robust optimization." *Environmental Modelling & Software*, 127, 104699.

Bentrup, G. (2008). "Conservation buffers: Design guidelines for buffers, corridors, and greenways." *Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC*, <https://www.fs.usda.gov/nac/buffers/docs/conservation_buffers.pdf> (9).

Beven, K. and Binley, A. (1992). "The future of distributed models: Model calibration and uncertainty prediction." *Hydrological Processes*, 6, 279–298.

Beven, K. and Freer, J. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology." *Journal of Hydrology*, 249, 11–29.

Blum, A. G., Ferraro, P. J., Archfield, S. A., and Ryberg, K. R. (2020). "Causal effect of impervious cover on annual flood magnitude for the united states." *Geophysical Research Letters*, 47.

Brodeur, Z., Wi, S., Shabestanipour, G., Lamontagne, J., and Steinschneider, S. (2024). "A hybrid, non-stationary stochastic watershed model (swm) for uncertain hydrologic simulations under climate change." *Water Resources Research*, 60(5), e2023WR035042.

Buechel, M., Slater, L., and Dadson, S. (2022). "Hydrological impact of widespread afforestation in great britain using a large ensemble of modelled scenarios." *Communications Earth & Environment*, 3, 6.

Chamberlain, S. (2019). *rnoaa: 'NOAA' Weather Data from R*, <https://cran.r-project.org/package=rnoaa>. R package version 0.8.4.

Chesapeake Conservancy (2017). "High-resolution land cover dataset 101." *Chesapeake Conservancy Conservation Innovation Center*, <https://www.chesapeakeconservancy.org/wp-content/uploads/2017/01/LandCover101Guide.pdf>.

Clapp, R. B. and Hornberger, G. M. (1978). "Empirical equations for some soil hydraulic properties." *Water Resources Research*, 14, 601–604.

Dingman, S. L. (1994). *Physical Hydrology*. Macmillan Publishing Company, <https://books.google.com/books?id=MRMSAQAAIAAJ>.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P.

598 C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara,
599 J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan,
600 J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., van Verseveld,
601 W., Volk, J., and Yamazaki, D. (2019). "Hillslope hydrology in global change research and earth
602 system modeling." *Water Resources Research*, 55, 1737–1772.

603 Fletcher, T. D., Shuster, W., Hunt, W. F., Ashley, R., Butler, D., Arthur, S., Trowsdale, S., Barraud,
604 S., Semadeni-Davies, A., Bertrand-Krajewski, J.-L., Mikkelsen, P. S., Rivard, G., Uhl, M.,
605 Dagenais, D., and Viklander, M. (2015). "Suds, lid, bmps, wsud and more – the evolution and
606 application of terminology surrounding urban drainage." *Urban Water Journal*, 12, 525–542.

607 Froelich, A. J. (1975). "Bedrock map of montgomery county, maryland." *Department of the Interior,*
608 *USGS*, <https://ngmdb.usgs.gov/Prodesc/proddesc_16412.htm>.

609 García-Gutiérrez, C., Pachepsky, Y., and Ángel Martín, M. (2018). "Technical note: Saturated
610 hydraulic conductivity and textural heterogeneity of soils." *Hydrology and Earth System Sciences*,
611 22, 3923–3932.

612 Giacomoni, M. H. and Joseph, J. (2017). "Multi-objective evolutionary optimization and monte
613 carlo simulation for placement of low impact development in the catchment scale." *Journal of*
614 *Water Resources Planning and Management*, 143.

615 Giuliani, M. and Castelletti, A. (2016). "Is robustness really robust? how different definitions of
616 robustness impact decision-making under climate change." *Climatic Change*, 135, 409–424.

617 Golden, H. E. and Hoghooghi, N. (2018). "Green infrastructure and its catchment-scale effects: an
618 emerging science." *Wiley Interdisciplinary Reviews: Water*, 5(1), e1254.

619 Hadka, D. and Reed, P. (2013). "Borg: An auto-adaptive many-objective evolutionary computing
620 framework." *Evolutionary Computation*, 21, 231–259.

621 Hadka, D. and Reed, P. (2015). "Large-scale parallelization of the borg multiobjective evolution-
622 ary algorithm to enhance the management of complex environmental systems." *Environmental*
623 *Modelling & Software*, 69, 353–369.

624 Hartig, F., Minunno, F., and Paul, S. (2017). *BayesianTools: General-Purpose MCMC and SMC*

*Samplers and Tools for Bayesian Statistics*, <https://github.com/florianhartig/BayesianTools>. R package version 0.1.4.

Herbst, R. S., Culver, T. B., Band, L. E., Wilson, B., and Quinn, J. D. (2023). "Integrating social equity into multiobjective optimization of urban stormwater low-impact development." *Journal of Water Resources Planning and Management*, 149.

Herman, J. D., Quinn, J. D., Steinschneider, S., Giuliani, M., and Fletcher, S. (2020). "Climate adaptation as a control problem: Review and perspectives on dynamic water resources planning under uncertainty." *Water Resources Research*, 56.

Herrera, P. A., Marazuela, M. A., and Hofmann, T. (2022). "Parameter estimation and uncertainty analysis in hydrological modeling." *WIREs Water*, 9.

Houska, T., Kraft, P., Chamorro-Chavez, A., and Breuer, L. (2015). "Spotting model parameters using a ready-made python package." *PLOS ONE*, 10, e0145180.

Hung, F., Harman, C., Hobbs, B., and Sivapalan, M. (2020). "Assessment of climate, sizing, and location controls on green infrastructure efficacy: A timescale framework." *Water Resources Research*, 56.

Hurtado-Pidal, J., Triana, J. S. A., Aguayo, M., Link, O., Valencia, B. G., Espitia-Sarmiento, E., and Conicelli, B. (2022). "Is forest location more important than forest fragmentation for flood regulation?." *Ecological Engineering*, 183, 106764.

InspectAPedia (2024). *Septic Drainfield Design: Septic Size Requirements Guide*, <https://inspectapedia.com/septic/Septic-Drainfield-Size.php>.

Jia, Y. and Culver, T. B. (2006). "Robust optimization for total maximum daily load allocations." *Water Resources Research*, 42.

Karimi, T., Reed, P., Malek, K., and Adam, J. (2022). "Diagnostic framework for evaluating how parametric uncertainty influences agro-hydrologic model projections of crop yields under climate change." *Water Resources Research*, 58.

Kavianihamedani, H., Quinn, J. D., and Smith, J. D. (2024). "New diagnostic assessment of mcmc algorithm effectiveness, efficiency, reliability, and controllability." *IEEE Access*, 12, 42385–

42400.

Klotz, D., Kratzert, F., Gauch, M., Sampson, A. K., Brandstetter, J., Klambauer, G., Hochreiter, S., and Nearing, G. (2022). "Uncertainty estimation with deep learning for rainfall–runoff modeling." *Hydrology and Earth System Sciences*, 26, 1673–1693.

Kraft, D. (1988). "A software package for sequential quadratic programming." *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*.

Laloy, E. and Vrugt, J. A. (2012). "High-dimensional posterior exploration of hydrologic models using multiple-try dream <sub>(zs)</sub> and high-performance computing." *Water Resources Research*, 48.

Lamontagne, J. R., Barber, C. A., and Vogel, R. M. (2020). "Improved estimators of model performance efficiency for skewed hydrologic data." *Water Resources Research*, 56.

Laumanns, M., Thiele, L., Deb, K., and Zitzler, E. (2002). "Combining convergence and diversity in evolutionary multiobjective optimization." *Evolutionary Computation*, 10, 263–282.

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins, E. (2020). "Partitioning climate projection uncertainty with multiple large ensembles and cmip5/6." *Earth System Dynamics*, 11(2), 491–508.

Leonard, L., Lin, L., and Band, L. E. (2017). "Deadrun definition files." *Hydroshare*, <https://www.hydroshare.org/resource/b60c5fa950ee401fac8a6ef0f1d0c441/>.

Lin, L. (2019a). *GIS2RHESSys*, <https://github.com/laurencelin/GIS2RHESSys>.

Lin, L. (2019b). *RHESSysEastCoast*, <https://github.com/laurencelin/RHESSysEastCoast>.

Lin, L. (2021). *RHESSys - EastCoast - rural urban catchment - Baisman Run, MD, U.S.*, <http://www.hydroshare.org/resource/424ff8bc247c43d09a168c2dbd808f52>.

Lin, L., Band, L. E., Vose, J. M., Hwang, T., Miniat, C. F., and Bolstad, P. V. (2019). "Ecosystem processes at the watershed scale: Influence of flowpath patterns of canopy ecophysiology on emergent catchment water and carbon cycling." *Ecohydrology*, 12(5), 1–15.

McMillan, H., Araki, R., Gnann, S., Woods, R., and Wagener, T. (2023). "How do hydrologists perceive watersheds? a survey and analysis of perceptual model figures for experimental

watersheds." *Hydrological Processes*, 37.

McPhail, C., Maier, H. R., Kwakkel, J. H., Giuliani, M., Castelletti, A., and Westra, S. (2018). "Robustness metrics: How are they calculated, when should they be used and why do they give different results?." *Earth's Future*, 6, 169–191.

Miles, B. (2015a). "Dead run 5, catonsville, md rhessys model." *Hydroshare*, <https://www.hydroshare.org/resource/895cc2fc795f4a63ab35c291d39977dc>.

Miles, B. (2015b). "Rhessys model of dead run 5 watershed, baltimore county, maryland, usa (with rain gardens)." *Hydroshare*, <https://www.hydroshare.org/resource/6dbb0dfb8f3a498881e4de428cb1587c/>.

Muleta, M. K. and Nicklow, J. W. (2005). "Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model." *Journal of Hydrology*, 306, 127–145.

Nash, J. and Sutcliffe, J. (1970). "River flow forecasting through conceptual models part i — a discussion of principles." *Journal of Hydrology*, 10, 282–290.

NREL (2012). "United states - annual average wind speed at 30 m." *WINDExchange*, <https://windexchange.energy.gov/maps-data/325>.

O'Neil-Dunne, J. and Grove., M. (2004). *GIS Shapefile - Transportation, Major Roads, Baltimore City - GDT Roads*, <https://doi.org/10.6073/pasta/3046b7191e6728ff387e27e40802eb5a>.

Penning, E., Burgos, R. P., Mens, M., Dahm, R., and de Bruijn, K. (2023). "Nature-based solutions for floods and droughts and biodiversity: Do we have sufficient proof of their functioning?." *Cambridge Prisms: Water*, 1, e11.

Pennino, M. J., McDonald, R. I., and Jaffe, P. R. (2016). "Watershed-scale impacts of stormwater green infrastructure on hydrology, nutrient fluxes, and combined sewer overflows in the mid-atlantic region." *Science of The Total Environment*, 565, 1044–1053.

Piscopo, A. N., Weaver, C. P., and Detenbeck, N. E. (2021). "Using multiobjective optimization to inform green infrastructure decisions as part of robust integrated water resources management plans." *Journal of Water Resources Planning and Management*, 147.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E.,

and Stromberg, J. C. (1997). "The Natural Flow Regime." *BioScience*, 47(11), 769–784.

Quinn, J. D., Reed, P. M., Giuliani, M., and Castelletti, A. (2017). "Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems." *Water Resources Research*, 53, 7208–7233.

Reybold, W. U. I. and Matthews, E. D. (1976). "Soil survey of baltimore county, maryland." *USDA Soil Conservation Service*, <https://www.nrcs.usda.gov/Internet/FSE_MANUSCRIPTS/maryland/baltimoreMD1976/baltimoreMD1976.p

Schoups, G. and Vrugt, J. A. (2010). "A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors." *Water Resources Research*, 46(10), 2009WR008933.

Shavazipour, B., Kwakkel, J. H., and Miettinen, K. (2021). "Multi-scenario multi-objective robust optimization under deep uncertainty: A posteriori approach." *Environmental Modelling & Software*, 144, 105134.

Smith, J. D. (2021). *RHESSys_ParamSA-Cal-GIOpt*, <https://github.com/jds485/RHESSys_ParamSA-Cal-GIOpt>.

Smith, J. D. (2024). "jds485/RHESSys_ParamSA-Cal-GIOpt: Robust Optimization Paper, <https://doi.org/10.5281/zenodo.11894592> (June).

Smith, J. D., Lin, L., Quinn, J. D., and Band, L. E. (2022). "Guidance on evaluating parametric model uncertainty at decision-relevant scales." *Hydrology and Earth System Sciences*, 26, 2519–2539.

Steinschneider, S., Polebitski, A., Brown, C., and Letcher, B. H. (2012). "Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change." *Water Resources Research*, 48.

Tague, C. L. and Band, L. E. (2004). "RHESSys: Regional Hydro-Ecologic Simulation System—an object-oriented approach to spatially distributed modeling of carbon, water, and nutrient cycling." *Earth Interactions*, 8(19), 1–42.

Tebyanian, N., Wu, H., Iulo, L., and Keller, K. (2022). "Uncertainty considerations in green

infrastructure optimization: A review." *Journal of Digital Landscape Architecture*, 549–560.

Tsai, W.-P., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., Liu, J., and Shen, C. (2021). "From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling." *Nature Communications*, 12, 5988.

Vaghefi, S. A., Iravani, M., Sauchyn, D., Andreichuk, Y., Goss, G., and Faramarzi, M. (2019). "Regionalization and parameterization of a hydrologic model significantly affect the cascade of uncertainty in climate-impact projections." *Climate Dynamics*, 53, 2861–2886.

van Meerveld, H. J. I., Jones, J. P. G., Ghimire, C. P., Zwartendijk, B. W., Lahitiana, J., Ravelona, M., and Mulligan, M. (2021). "Forest regeneration can positively contribute to local hydrological ecosystem services: Implications for forest landscape restoration." *Journal of Applied Ecology*, 58, 755–765.

Vrugt, J. A. (2016). "Markov chain monte carlo simulation using the dream software package: Theory, concepts, and matlab implementation." *Environmental Modelling and Software*, 75, 273–316.

Wald, A. (1949). "Statistical decision functions." *The Annals of Mathematical Statistics*, 165–205.

Webber, M. K. and Samaras, C. (2022). "A review of decision making under deep uncertainty applications using green infrastructure for flood management." *Earth's Future*, 10.

Welty, C. and Lagrosa, J. (2018). *Precipitation measurements at eight stations for the Baltimore Ecosystem Study*, <https://doi.org/10.6073/pasta/56a2cd35bdd8aa1978529fe0404573ab>.

Williams, T., Guikema, S., Brown, D., and Agrawal, A. (2020). "Assessing model equifinality for robust policy analysis in complex socio-environmental systems." *Environmental Modelling & Software*, 134, 104831.

Zatarain Salazar, J., Reed, P. M., Quinn, J. D., Giuliani, M., and Castelletti, A. (2017). "Balancing exploration, uncertainty and computational demands in many objective reservoir optimization." *Advances in Water Resources*, 109, 196–210.

Zhang, J., Yu, Z., Zhao, B., Sun, R., and Vejre, H. (2020). "Links between green space and public health: a bibliometric review of global research trends and future prospects from 1901 to 2019."

760 *Environmental Research Letters*, 15, 063001.

761 Zhang, R., Band, L. E., and Groffman, P. M. (2023). "Balancing upland green infrastructure and
762 stream restoration to recover urban stormwater and nitrate load retention." *Journal of Hydrology*,
763 626, 130364.

Supplemental Information for
COMPARING ROBUST OPTIMIZATION APPROACHES FOR ADDRESSING HYDROLOGIC MODEL UNCERTAINTY IN INFRASTRUCTURE PLANNING:
A GREEN INFRASTRUCTURE EXAMPLE

Jared D. Smith, Julianne D. Quinn, and Lawrence E. Band

## S1: Supplemental Figures

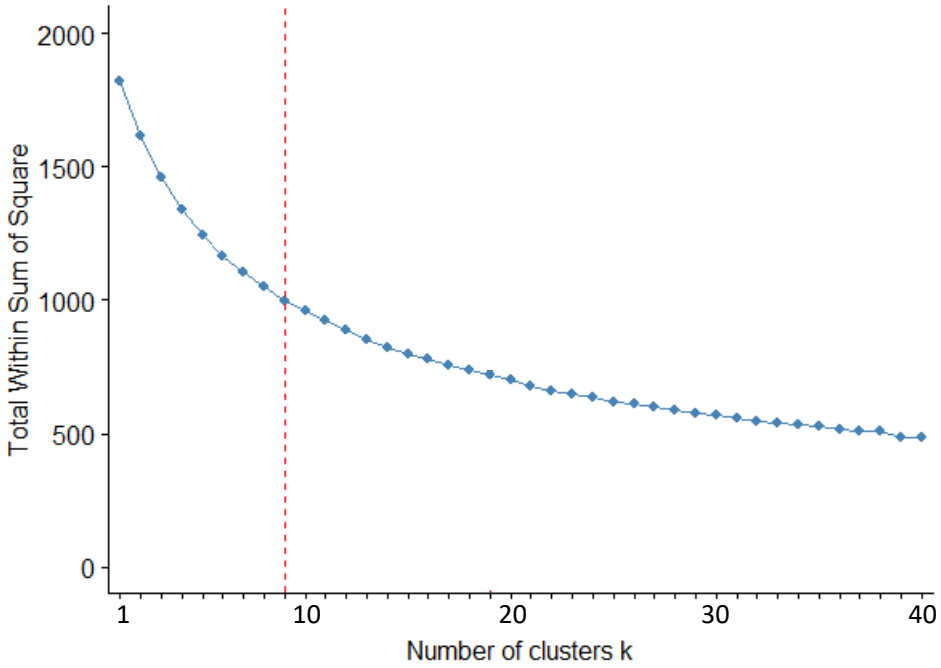


**Figure S1:** K-means clustering results for the hydrological model parameterizations discovered by Bayesian calibration, after removing parameterizations that provided an NSE of less than 0.7 for the log of streamflow at the watershed outlet. Our selected number of clusters is shown as a red dashed line.

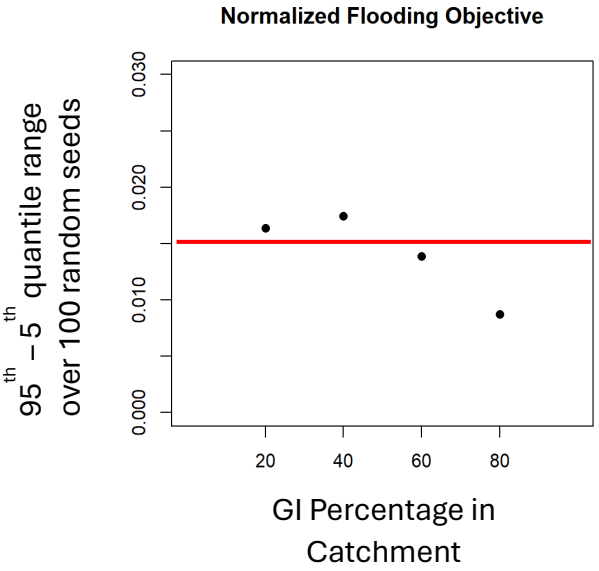


**Figure S2:** Average 5th to 95th percentile range of flooding and low flow objective values over 100 random seeds for different reforestation (GI) allocation percentages. The red line shows the selected epsilon tolerance level for the flooding and low flow objectives in the optimization.

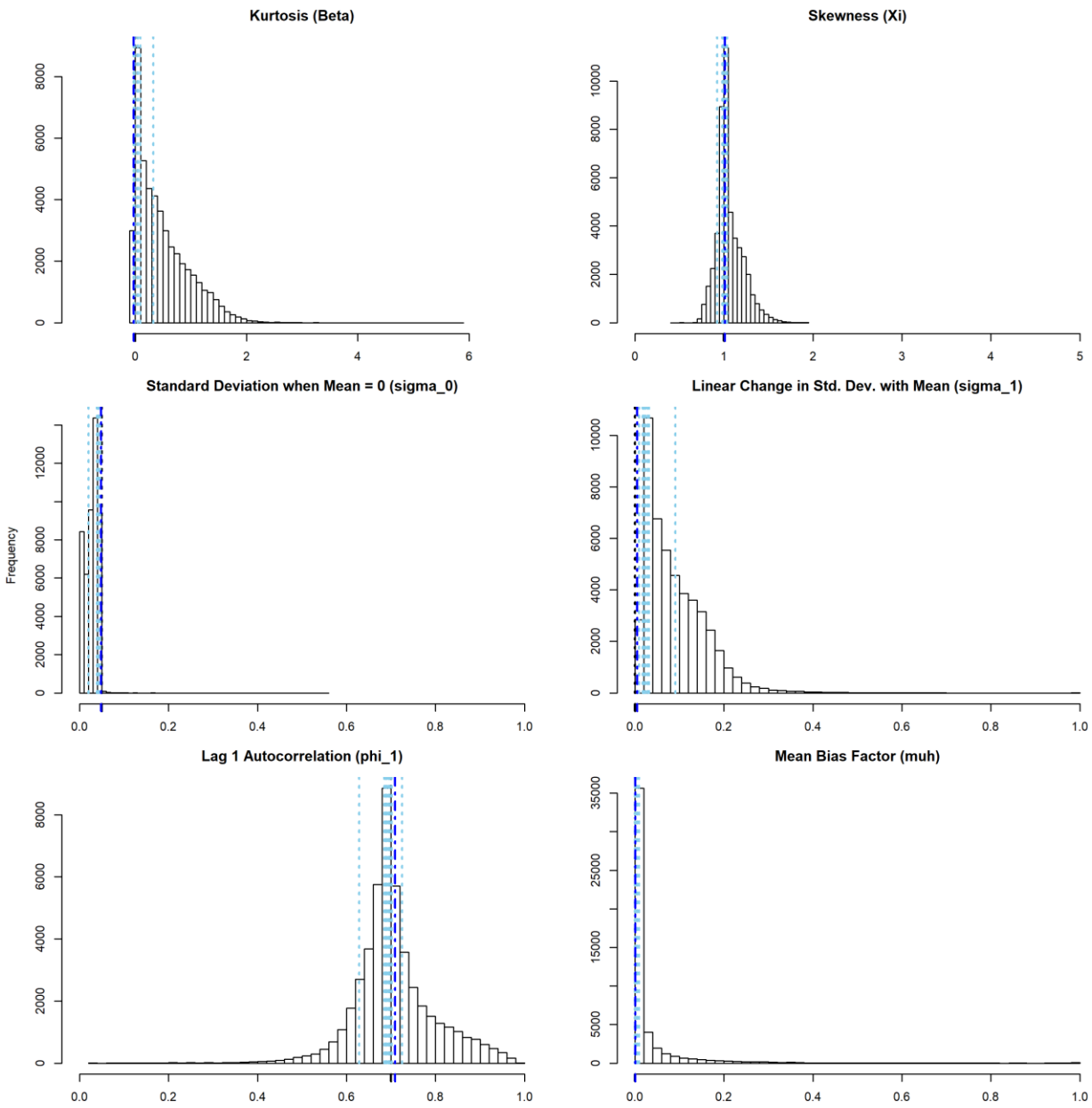


**Figure S3:** Posterior distributions of likelihood model parameters. The black thick tick mark on the x-axis indicates the true value, the blue thick dashed line is the MAP parameterization, and the light blue dotted lines are the other parameterizations selected for MORO and MinMax methods.

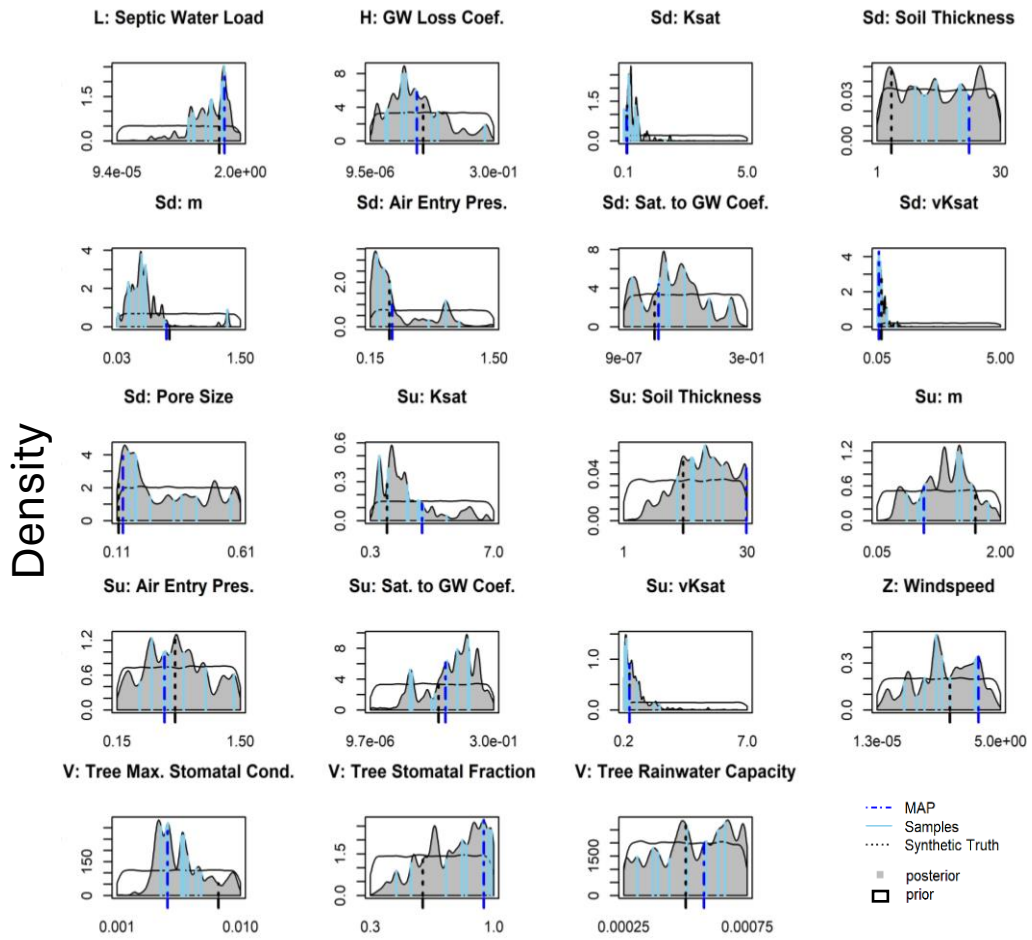


**Figure S4:** Posterior distributions of likelihood model parameters. The black thick dotted line indicates the true value, the blue thick dashed line is the MAP parameterization, and the light blue lines are the other parameterizations selected for MORO and MinMax methods.

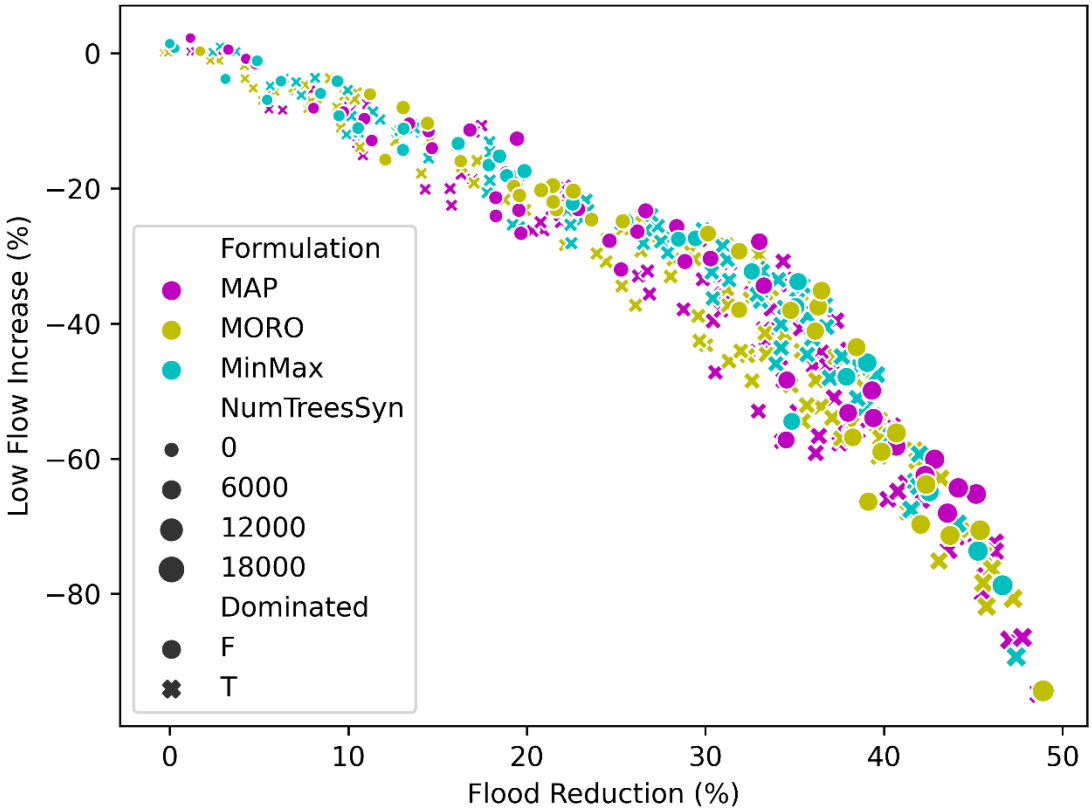


**Figure S5:** Dominated and nondominated solutions discovered for each of the optimization experiments, after evaluation on the synthetic true parameter set. Each method contributes about equally to the nondominated solutions.

## S2: Supplemental Tables

**Table S1**: Prior ranges for likelihood function parameters.

| Parameter | Uniform Prior |
|---|---|
| $\beta$ | (-1, 7] |
| $\xi$ | (0, 5] |
| $\sigma_0$ | [0.000001, 1] |
| $\sigma_1$ | [0, 1] |
| $\phi_1$ | [0, 1] |
| $\mu_h$ | [0, 1] |

**Table S2**: DREAM$_{(zs)}$ hyperparameter values.

| Hyperparameter | Description | Value |
|---|---|---|
| Iterations | Number of iterations per chain | 1200 |
| eps | Normally distributed noise to add to the proposal update | 0 |
| e | Multiplier used in the proposal update (default value used). | 0.05 |
| ZupdateFrequency | Iterations after which the archive (Z) is updated with solutions. | 3 |
| pSnooker | Probability of a snooker update | 0.1 |
| DEpairs | Number of chains used to compute the DE proposal update. | 2 |
| nCR | Number of crossovers | 10 |
| updateInterval | Number of iterations after which crossovers occur | 10 |
| burnin | Burn in for the chains. We used 0 within the algorithm and removed the adaptation iterations later. | 0 |
| adaptation | Number of iterations per chain for which adaptation occurs. | 400 |
| thin | Thinning rate. We did not use thinning (set to 1). | 1 |

## S3: RHESSys Parameter Values

All parameters that were not adjusted in the Bayesian calibration are provided as definition files that are used to run RHESSys. We also provide definitions of all variables in the file called S3_Tables_RHESSysParameterDescriptions.xlsx.