Calibrating a large-domain land/hydrology process model in the age of AI: the SUMMA CAMELS emulator experiments

Mozhgan A. Farahani[1], Andrew W. Wood[1,2], Guoqiang Tang[3], and Naoki Mizukami[4]

[1]Climate and Global Dynamics, National Center for Atmospheric Research, Boulder, CO, USA
[2]Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, USA
[3]State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan, China
[4]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

Correspondence to: Mozhgan A. Farahani (mozhgana@ucar.edu); Andrew W. Wood (andywood@ucar.edu)

Abstract. Process-based (PB) hydrological modeling is a long-standing capability used for simulating and predicting complex water processes over large, hydro-climatically diverse domains, yet PB model parameter estimation (calibration) remains a persistent challenge for large-domain applications. New techniques and concepts arising in the artificial intelligence (AI) context for hydrology point to new opportunities to tackle this problem in complex PB models. This study introduces a new scalable calibration framework that jointly trains a machine learning emulator for model responses across a large-sample collection of watersheds while leveraging sequential optimization to iteratively refine hydrological model parameters. We evaluate this strategy through a series of experiments using the Structure for Unifying Multiple Modeling Alternatives (SUMMA) hydrological modeling framework coupled with the mizuRoute channel routing model for streamflow simulation. This 'large-sample emulator' (LSE) approach integrates static catchment attributes, model parameters, and performance metrics, and yields a powerful new strategy for large-domain PB model parameter regionalization to unseen watersheds. The LSE approach is compared to using a more traditional individual basin calibration approach, in this case using a single-site emulator (SSE), trained separately for each basin. The jointly trained LSE framework achieves comparable or better performance to traditional individual basin calibration, while further enabling potential for probabilistic parameter regionalization to out-of-sample, unseen catchments. Motivated by the need to optimize complex hydrology models across continental-scale domains in support of applications in water security and prediction, this work demonstrates a strategy to leverage new insights from AI era hydrology research can help to surmount old challenges in the calibration and regionalization of large-domain PB models.

Short summary. We present a new strategy to calibrate large-domain land/hydrology models over diverse and extensive regions. Using the Structure for Unifying Multiple Modeling Alternatives (SUMMA) and mizuRoute models, our approach integrates catchment attributes, model parameters, and performance metrics to optimize streamflow simulations. By leveraging recent innovations in machine learning methods and concepts for hydrology, we improve calibration outcomes and enable regionalization to ungauged basins, which is valuable for national-scale water security studies.

Style Definition: Heading 2
Style Definition: Heading 3
Style Definition: Heading 4
Style Definition: Bullets: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0" + Tab after: 0.5" + Indent at: 0.5"
Formatted: Font: 16 pt, Bold, Font color: Black
Formatted: Normal, Line spacing: Multiple 1.83 li, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)
Formatted: Font: 16 pt, Bold, Font color: Black
Deleted: Tang[1]
Deleted: Mizukami[3]
Formatted: Font: 12 pt, Font color: Black
Formatted: Font: 12 pt, Font color: Black
Formatted: Font: 12 pt, Font color: Black
Formatted: Font color: Black
Deleted: [3]Research
Formatted: Default Paragraph Font
Deleted: mozhgana@ucar.edu
Formatted: Font color: Black
Formatted: Normal, Space Before: 6 pt, After: 18 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)
Deleted: andywood@ucar.edu
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: strategy…apability used for simulating and predicting complex water processes over large, hydro-climatically diverse domains, yet PB model parameter estimation (calibration) remains a persistent challenge for large-scale…omain applications. New techniques and concepts arising in the artificial intelligence (AI) context for hydrology point to new opportunities to tackle this problem in process-based…omplex PB models. This study presents…ntroduces a new scalable calibration framework that jointly trains a machine learning (ML) based calibration strategy for large-domain modeling, implemented…mulator for model responses across a large-sample collection of watersheds while leveraging sequential optimization to iteratively refine hydrological model parameters. We evaluate this strategy through a series of experiments using the Structure for Unifying Multiple Modeling Alternatives (SUMMA) land/hydrology model …ydrological modeling framework coupled with the mizuRoute channel routing model. We explore various ML methods to develop and evalu ... [2]
Deleted: 1¶
Formatted: Font color: Black
Formatted ... [1]

1 Introduction

Hydrological modeling advances have significantly expanded our capacity to simulate and predict complex water-related processes. Such models provide critical information for water resource management and planning, flood hazard prevention, and climate resilience studies, among other applications. Accurate hydrologic simulations are vital in regions as expansive and diverse as the contiguous United States (CONUS), if not the globe, where variability in climate, land cover, and hydrological responses can be a challenge for the seamless implementation of land/hydrology models (LHMs: i.e., hydrologic models and/or the hydrologic components of land models). Traditional single-site calibration approaches that involve tuning model parameters for individual basins can be time-intensive, spatially non-generalizable and computationally costly, which limits their suitability for large-domain (national, continental, global) applications (Shen et al., 2023; Tsai et al., 2021; Herrera et al., 2022). Because parameter estimation is vulnerable to sampling and input uncertainty and input errors, such basin-specific methods often lead to spatial inconsistencies in parameter estimates, limiting the model's generalizability across broader regions (Wagener and Wheater, 2006).

Recent advances and applications in artificial intelligence (AI) -- a family of methods including machine learning (ML), deep learning (DL), large language models (LLMs) and other methods -- have been demonstrated to provide not only a skillful strategy for simulating hydrology (Kratzert et al., 2024; Nearing et al., 2024; Arsenault et al., 2023; Feng et al., 2020), but also for process-based (PB) hydrology model calibration. Calibration methods in hydrology are numerous and have a long history, advancing hand-in-hand with the proliferation of models ranging in complexity from low-dimensional conceptual schemes (commonly used in engineering applications and operational forecasting) to more explicit high resolution PB models used in watershed and Earth System science. The greater complexity of such models drove calibration method innovations such as surrogate modeling for individual basins (Gong et al., 2016; Adams et al., 2023), which enabled a less-costly interrogation of the model parameter space despite the models' increased computational demand. Such techniques have even more recently been discovered by the Earth System modeling (ESM) community, which previously calibrated complex ESM components (e.g., ocean, atmosphere, land) through ad hoc manual parameter sensitivity testing and tuning. AI-based methods including model emulators are now increasingly used for exploring land model parameter uncertainty and constraining model implementations (Dagon et al., 2020; Watson-Parris et al., 2021; Bennett et al., 2024).

ML hydrology modeling applications have yielded the remarkable (and perhaps in retrospect, unsurprising) finding that joint model training across many watersheds can learn robust, heterogeneous hydrometeorological relationships that enable them to predict hydrological behavior for unseen watersheds and time periods -- which represents a large step forward in solving the longstanding hydrological prediction-in-ungauged-basins challenge (PUB; Wagener et al., 2007; Hrachowitz et al., 2013). Mai et al. (2022) clearly demonstrated the superior performance of Long Short-Term Memory (LSTM) networks in out-of-sample temporal and spatial hydrologic simulation compared to a range of results from non-ML models. Such regionalization ability had not been achieved previously with conceptual and PB hydrology models, where joint multi-site or regional training more

Deleted: modelings

Deleted: .

Deleted: .

Deleted: physics-based

Deleted: tuning and

Deleted: tests.

Deleted: related

Deleted: 2¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

often comes at a cost to individual basin model performance (Mizukami et al., 2017; Samaniego et al., 2010; Tsai et al., 2021; Kratzert et al., 2024), notwithstanding some gains in regional parameter coherence. Samaniego et al. (2010) achieved moderate success at parameter regionalization using a joint large-domain training solution involving calibrating the coefficients of transfer functions relating geophysical attributes ('geo-attributes') to model parameters, expanding on common pedotransfer concepts for soil parameters. Since then, ML and DL approaches, including differentiable modeling -- e.g., embedding of DL elements within conceptual models converted to differentiable form (Feng et al., 2020; Shen et al., 2023) -- and hybrid ML/conceptual models (Frame et al., 2022) have continued to advance, outperforming traditional models and showing new potential for generalizing to ungauged basins with diverse hydroclimatic conditions (Kratzert et al., 2024; Feng et al., 2020).

Generally, emulator strategies have evolved along two primary lines: (i) emulating model performance by directly relating model parameters to one or more performance objective functions, without explicitly modeling the dynamic behavior of the system (Gong et al., 2016; Herrera et al., 2022; Maier et al., 2014; Razavi et al., 2012; Sun et al., 2023), and (ii) emulating key dynamic model states or fluxes, then using the resulting emulator outputs (e.g., time series) to cheaply explore parameter-output sensitivities (Bennett et al., 2024; Maxwell et al., 2021). Importantly, this study explicitly focuses on the first strategy, emulation of model performance metrics, which originated primarily within hydrological modeling contexts. This choice greatly reduces the need to run the full hydrological model iteratively during calibration, substantially lowering computational expense and enabling scalable optimization for increasingly complex, large-domain hydrology models.

The aim of the research described in this paper is to surmount traditional basin-specific calibration challenges by leveraging insights from recent AI-related progress in hydrology. The specific objective (and research sponsor motivation) for the study is to calibrate a PB LHM, the Structure for Unifying Multiple Modeling Alternatives (SUMMA; see Sect. 2.2) over the entire CONUS for use in generating a large ensemble of future climate-informed hydrologic scenarios for use by US federal water agencies and others in water security applications -- e.g., agency guidance and long-term planning studies. Prior experience with individual basin calibration followed by regionalization, and associated performance limitations, motivated this investigation of possibilities for a more scalable and powerful approach.

To this end, we present an ML-based model calibration and regionalization strategy and associated method evaluation experiments for the CONUS-wide implementation of SUMMA, which was also demonstrated for calibrating the hydrology component of an ESM land model in a companion paper by Tang et al. (2025). The large-sample emulator (LSE) approach employs a joint training strategy that combines model performance (i.e., response surface) emulation and parameter optimization scheme to estimate parameters jointly across diverse catchments, building on recent advances in the ML hydrologic modeling community. By training the emulator on a large sample catchment dataset to predict model performance as a function of catchment geo-attributes and parameters, we build the capacity for identifying optimal parameter sets across large, varied and unseen domains. We compare the LSE results with traditional single-site emulator (SSE) calibration, and comment on avenues for further advances in this direction. This study evaluates whether the LSE framework can improve

Moved (insertion) [1]

Deleted: physics-based

Deleted: is

Deleted: a

Deleted: 2024

Deleted: combined

Deleted: from

Deleted: an

Deleted: The following sections describe

Deleted: 3¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

model calibration performance over the SSE, and whether the LSE enables effective regionalization of parameters to unseen basins through spatial cross-validation. The following sections describe and discuss the methods and results of a series of experiments with this approach as applied to a large collection of US watersheds.

2 Methods

2.1 Study domain

The study focuses on CONUS, a region encompassing diverse hydrological conditions due to its varied climate, landforms, and vegetation types. To represent such variability, we utilize a subset of watersheds from the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) dataset, which combines static catchment attributes with hydrometeorological time series for use in benchmarking hydrological modeling applications (Newman et al., 2015; Addor et al., 2017). Such datasets are well-suited for large-domain modeling due to their rich suite of attributes, including climate indices, soil properties, land cover, and streamflow observations, which provide a comprehensive basis for model calibration, evaluation and regionalization over a diverse range of hydroclimate settings. We selected 627 headwater basins from the 671 CAMELS basins, excluding those with nested interior basins to ensure independence and avoid overlapping drainage areas. Catchment boundaries for the modeling were updated from those provided in the original CAMELS dataset, correcting inaccuracies in boundary and drainage areas by using the original boundaries from the Geospatial Attributes of Gages for Evaluating Streamflow, version II dataset (Falcone, 2011), which are consistent with U.S. Geological Survey (USGS)-estimated drainage areas. The spatial unit for the calibration experiments is each CAMELS watershed. A comprehensive summary of the CAMELS basin characteristics is provided in Addor et al. (2017) and is not reproduced here.

Deleted: , followed by discussion and conclusions

Deleted: scale

Deleted: or upstream areas

Formatted: Font: Bold, Font color: Black

Deleted: 4¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

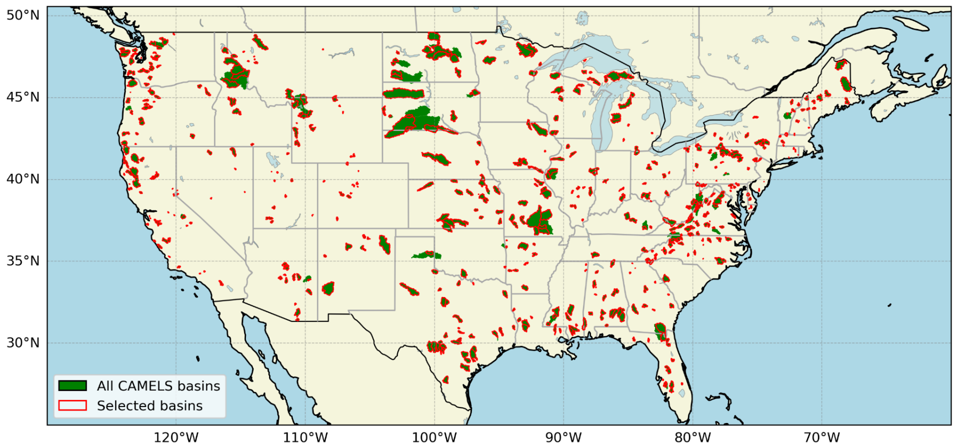
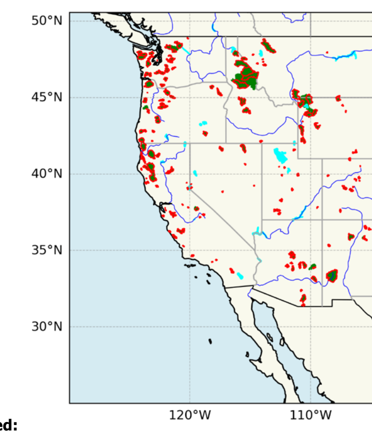


Figure 1. Spatial distribution of selected headwater basins (red outlines) from the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) dataset (green areas) across the contiguous United States (CONUS).

2.2 Process-based modeling with SUMMA and mizuRoute

SUMMA is a PB LHM framework designed for flexibility in representing hydrological processes across diverse catchments (Clark et al., 2015a, 2015b, 2021). SUMMA solves generalized mass and energy conservation equations, offering multiple parameterization schemes for hydrological fluxes, and enabling flexible advanced numerical techniques to optimize solution performance. SUMMA represents watersheds with a hierarchical spatial organization centering on Grouped Response Units (GRUs) that are divisible into one or multiple Hydrologic Response Units (HRUs). GRU geometry is user-defined and has varied in usage from mesoscale catchment boundaries to fine or mesoscale resolution grid, as well as point-scale simulations. Such configurations allow SUMMA to represent the natural topography of the domain to the extent warranted by a given application, thereby improving the interpretability of model results (Gharari et al., 2020).

Here as in other SUMMA modeling studies, runoff and subsurface discharge outputs from SUMMA simulations are subsequently input to the mizuRoute channel routing model (Mizukami et al., 2016), a flexible framework supporting multiple hydrologic routing methods to provide streamflow estimates at gage locations. MizuRoute organizes the routing domain using catchment-linked HRUs connected by stream segments (Mizukami et al., 2016, 2021). Currently five methods are offered in MizuRoute, of which the Diffusive Wave (DW) routing scheme, as implemented by Cortés-Salazar et al. (2023), was adopted here. Both SUMMA and mizuRoute model codes are open source and their development has been extensively sponsored by



Deleted:

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Deleted: .

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: The Structure for Unifying Multiple Modeling Alternatives (SUMMA) is a process-based

Deleted: SUMMA

Deleted: have

Deleted: 5¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

the US water agencies (the Bureau of Reclamation and US Army Corps of Engineers, USACE) with growing support from other agencies in the US and internationally.

For this study, SUMMA and mizuRoute are run at a nominal 3-hourly simulation timestep. The associated sub-daily forcing, including precipitation, temperature, specific humidity, shortwave and longwave radiation, wind speed, and air pressure, were derived from gridded datasets but spatially aggregated across each basin area, resulting in basin-averaged input time series. Specifically, precipitation and temperature forcings were derived from the Ensemble Meteorological Dataset for Planet Earth (EM-Earth), which provides hourly data with 0.1° spatial resolution, merging ground-station data with reanalysis for enhanced accuracy (Tang et al., 2022). EM-Earth integrates gap-filled ground station data with reanalysis estimates, offering improved accuracy over standalone reanalysis products. Similarly, wind speed, air pressure, shortwave and longwave radiation, and specific humidity inputs were derived from ERA5-Land, also at 0.1° spatial and hourly temporal resolutions (Muñoz-Sabater, 2021). To match the EM-Earth spatial configuration (a 0.05° offset), the ERA5-Land grids were interpolated, and the combined basin forcings dataset spanned the period 1950-2023.

The initial (default) SUMMA configuration and parameters used in this study were developed in prior SUMMA and mizuRoute applications projects (e.g., Broman and Wood, 2021; Wood et al., 2021; Wood and Mizukami, 2022), based on expert judgment involving consultation with other model developers, evaluation of previous modeling experiments and sensitivity analyses, and model process algorithms that directly influence runoff generation. These choices include model physics selections, soil and aquifer configuration, spatial and temporal resolution, an *a priori* parameter set and target calibration parameters. The SUMMA model configuration adopted a single HRU per GRU, in which the GRU was the entire lumped area of each catchment. A maximum of 5 layers was specified for snow, and the subsurface included 3 soil layers with a total layer depth of 1.5m, underlain by an aquifer (bucket) with a maximum water holding capacity of 1.5m. For the routing network, the MERIT-Hydro (Lin et al., 2019) stream channel topology was chosen.

The SUMMA calibration parameters and physics configuration choices are summarized in Table 1 and Table A1 (appendix), respectively. Default values shown in Table 1 are taken from a representative basin (e.g., Basin ID: 05120500) for reference. Some default parameter values (e.g., soil or vegetation-related) vary across basins based on local attributes, and the values in this table may not be globally consistent across the domain. The first phase of the calibration process, via a Latin Hypercube Sampling (LHS) of the parameter space and model response, supported sensitivity analysis and refinement of parameter search bounds to focus trial values into likely behavioral areas and/or to avoid model convergence issues (e.g., the lowest theoretically possible values of the vGn_n parameter in SUMMA produces non-physical behavior). The calibration 'trial' parameter selection was designed to control major hydrologic process phenomena -- e.g., infiltration, evapotranspiration, soil storage and transmission, snow accumulation and melt, hillslope runoff attenuation, aquifer storage and release -- though identifying an efficient number of controlling parameters, versus conducting comprehensive parameter sensitivity assessment and selection optimization.

Deleted: , and the
Deleted: include
Deleted: . Precipitation
Deleted: and hourly temporal
Deleted: Wind
Deleted: 2019
Deleted: and review of
Deleted: parameterizations (i.e.,
Deleted:) to assess their
Deleted: on
Deleted: 0m
Deleted: 2
Deleted: Yamazaki
Deleted: 6¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

We note that the 'default' parameter values used here reflected prior study calibration efforts from a site-specific CAMELS-based SUMMA streamflow calibration project conducted by authors Wood and Mizukami (unpublished). The earlier effort used the Dynamically Dimensioned Search (DDS) algorithm (Tolson and Shoemaker, 2007) and calibrated many of the same parameters. However, the work used an earlier version of SUMMA and did not include mizuRoute routing, thus it forms only a baseline reference for our current parameter choices in this study. The prior effort's SUMMA-CAMELS dataset, DDS calibration workflow and parameter selections later contributed to a SUMMA sensitivity study (Van Beusekom et al., 2022) and was published in associated repositories.

Table 1. Selected calibration parameters with default values and ranges.

Parameter name	Default	Minimum	Maximum	Process importance
k_soil	7.5e-06	1e-07	0.0001	Hydraulic conductivity: regulates transmission of water through soil layers
theta_sat	0.55	0.2	0.7	Soil porosity: influences water storage capacity
aquiferBaseflowExp	2.0	1.0	4.0	Controls aquifer discharge
aquiferBaseflowRate	0.001	0.0001	0.1	Controls aquifer discharge
qSurfScale	5.0	1.0	20.0	Affects partitioning of direct runoff versus infiltration
summerLAI	3.0	0.2	10.0	Regulates transpiration
frozenPrecipMultip	1.0	0.5	2.5	Snow undercatch factor, scales winter precipitation
routingGammaScale	18000.0	360.0	86400.0	Controls GRU combined runoff attenuation and delay
routingGammaShape	2.5	1.0	5.0	Controls GRU combined runoff attenuation and delay
Fcapil	0.06	0.01	0.11	Affects refreeze of snowmelt within pack, timing of snowmelt runoff
tempCritRain	273.16	270.16	276.16	Temperature threshold to discriminate rain from snow
heightCanopyTop	20.0	2.0	50.0	Impacts turbulent heat fluxes (sensible, latent); influences snow cycle timing and magnitude
heightCanopyBottom	2.0	0.000	5.0	Not directly calibrated; scaled proportionally to heightCanopyTop

Deleted: The …e note that the 'default' parameter values u … [5]
Formatted … [6]
Formatted … [8]
Formatted … [7]
Deleted: Name
Deleted: Importance
Formatted … [9]
Formatted … [10]
Formatted … [11]
Deleted: hydraulic
Formatted … [14]
Formatted … [12]
Formatted … [13]
Deleted: soil
Formatted … [15]
Formatted … [16]
Formatted … [17]
Deleted: controls
Formatted … [18]
Formatted … [19]
Formatted … [20]
Formatted … [21]
Formatted … [22]
Deleted: controls
Formatted … [23]
Formatted … [24]
Formatted … [25]
Deleted: affects
Formatted … [26]
Formatted … [27]
Formatted … [28]
Deleted: regulates
Formatted … [29]
Formatted … [30]
Formatted … [31]
Deleted: snow
Formatted … [32]
Deleted: controls
Formatted … [33]
Formatted … [34]
Formatted … [35]
Deleted: controls
Formatted … [36]
Formatted … [37]
Formatted … [38]
Deleted: affects
Formatted … [41]
Deleted: 10
Formatted … [39]
Formatted … [40]
Deleted: temperature
Formatted … [42]
Formatted … [43]
Formatted … [44]

windReductionParam	0.28	0.05	1.0	Impacts turbulent heat fluxes (sensible, latent)
vGn_n	2.0	1.3	4.0	Van Genuchten 'n': regulates retention of water in soil layers

2.3 ML-Based parameter estimation approach

We apply and assess the LSE and related parameter estimation techniques introduced in Tang et al. (2025), a companion paper that focused on calibrating the hydrological components of the Community Terrestrial Systems Model (CTSM; Lawrence et al., 2019). For this study, the approach was further tailored to calibrate the combined SUMMA-mizuRoute model (e.g., channel routing was not used with CTSM). As noted in the Introduction, basin-specific calibration approaches can be computationally intensive and result in spatially discontinuous parameter fields, limiting their scalability and generalizability to large, diverse domains like CONUS. To assess whether the LSE can offer a more effective calibration strategy for large-domain SUMMA modeling, we run several experiments using the ML-based emulation strategy to optimize model parameters, focusing on two contrasting variations: the basin-specific single-site emulator (SSE) and the combined-basin joint LSE, which simultaneously calibrates multiple basins. The calibration period spans six water years, from October 1982 to September 1989, with the first year treated as spin-up and excluded from model evaluation. This period was selected based on its consistent data availability across basins and its use in previous large-sample studies, allowing for comparability and minimizing confounding effects from land use change or climate trends. Figure 2 provides a schematic overview of the iterative ML-based calibration workflow using SSE or LSE, including time series input data and selected parameters for each basin, initial sampling ('iter-0'), emulator-based iterative optimization, and final parameter selection. Details are provided in the following sections.

Deleted: impacts
Formatted: Left, Space After: 0 pt
Formatted: Centered, Space After: 0 pt
Formatted: Left, Space After: 0 pt
Deleted: van
Formatted: Left, Space After: 0 pt
Formatted: Centered, Space After: 0 pt
Formatted: Left, Space After: 0 pt
Deleted: 'large-sample emulator' (
Deleted:)
Deleted: 2024
Deleted: two
Formatted: Font: Bold
Deleted: 8¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

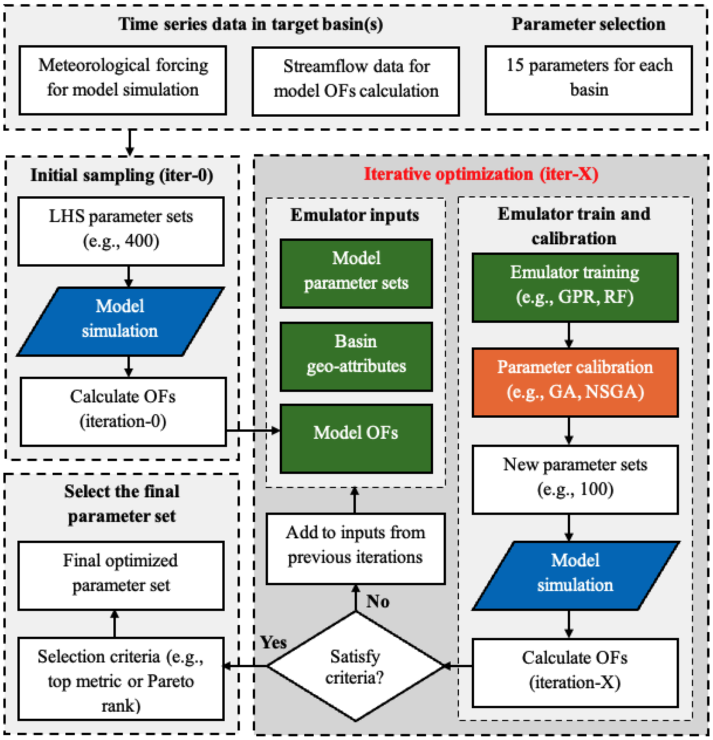


Figure 2. The flowchart of the parameter optimization in this study, including target parameter selection, initial sampling (iter-0), iterative optimization cycles (iter-X) using emulators, and the selection of the final parameter set. For emulator inputs, the basin geo-attributes are only used for the large-sample emulator (LSE) approach.

2.3.1 SSE-based calibration

The SSE calibration approach optimizes model parameters for each basin separately. Our approach is based on other single site optimization approaches that use surrogate modeling (e.g., the MO-ASMO method of Gong et al., 2016) to represent the relationship between model parameters and objective function (OF) results. The initial step involves generating a large set of parameter combinations (400) using LHS for each basin. These parameter sets are used to run SUMMA-mizuRoute simulations, and their performance is quantified using one or more OFs, which serve as the minimization target for calibration.

Moved (insertion) [2]

Deleted: (400)

Deleted: Latin Hypercube Sampling (

Deleted:)

Deleted: 9¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

These initial semi-random sampling and model OF evaluations are first used in selecting the form of the emulator to be used for each basin. Based on insights from Tang et al. (2025), two emulators—Gaussian process regression (GPR) and random forests (RF)—are assessed here via a five-fold training and cross-validation procedure on the initial LHS sample. The emulator with better performance is selected and then retrained on the complete initial parameter set, for each basin separately. We note that the emulator can also play a role in identifying and selecting optimal parameters for calibration as described in Tang et al. (2025), which provides details on that usage.

Following this step, the main iterative calibration process begins. For each basin, the trained emulator is used within an optimization algorithm to explore the parameter space, searching for improved parameter sets that minimize the OF. In this study, a Genetic Algorithm (GA; Mitchell, 1996) is employed for single-objective calibration, whereas the Non-dominated Sorting Genetic Algorithm II (NSGA-II; Deb et al., 2002) is used for multi-objective calibration. Each iteration involves generating a new suite of emulator-predicted parameter sets (100), which are then used to run the SUMMA-mizuRoute model and calculate model OFs. These results are added to the existing previous parameter sets to retrain the emulator for further optimization, leading to the next iteration. This iterative process continues until a specified stopping criterion is met, such as achieving a performance threshold or completing a predetermined number of iterations. In this study, limited iterations (six following the initial LHS iter-0) combined with a greater number of trials per iteration helped reduce noise while improving calibration efficiency. The number of trials per iteration and other hyperparameters (Table A4) of the process were selected through sensitivity testing and also informed by the experimental outcomes of Tang et al. (2025).

2.3.2 LSE-based calibration

In contrast to the SSE, the LSE calibration approach includes all basins jointly in a single calibration process that estimates optimal parameters in all basins at once. The initial phase is similar to the SSE approach, where a large number of LHS selected parameter sets (e.g., 400 in this study) are generated for each of the 627 basins, and their performance in SUMMA-mizuRoute is evaluated. As in the SSE, the LHS parameter sets are unique for each basin to afford the maximum diversity in parameter trials across the associated model simulations (as in Baker et al., 2021). In contrast to the SSE, however, the emulator relies on additional static basin attributes to learn the parameter-performance responses of different types of basins. In this study, we choose 27 such 'geo-attributes' representing basin-specific geographic and climatic characteristics, such as soil properties, vegetation, and climate indices (Table A2). Including geo-attributes enables the emulator to estimate model performance (i.e., the OFs) conditioned on both parameter and attribute values, which means the LSE can be used to predict potentially optimal parameter sets for unseen basins where the performance is not known or cannot be measured, enabling its potential use in parameter regionalization, i.e., prediction in ungauged basins (PUB).

We use the RF emulator for the LSE due to its speed and performance relative to GPR, which struggles to train on the much larger joint basin parameter trial dataset -- i.e., 400*627 or 250,800 samples for the initial sampling step, growing by 62,700 new trials with each iteration. To start each calibration iteration, the trained RF emulator from the initial step (iter-0) or prior

Deleted: performance
Deleted: 2024
Deleted:
Deleted: A
Deleted: 1998
Deleted: (100)
Deleted: ,
Deleted: iteration '0'
Deleted:
Deleted: 2024
Deleted: set (400)
Deleted: must also use
Deleted: distinguish between
Deleted: these
Deleted: .
Deleted: Figure 2 provides a schematic overview of the ML-based calibration workflow used in this study.
Deleted: as the
Deleted: form
Deleted: iteration
Deleted: 10¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

iteration is used by an optimization algorithm to predict potentially improved parameter sets (100) for each basin individually. In multi-objective optimization, the NSGA-II inherently produces a pareto-set of optimized parameters, whereas for single-objective optimization, we achieve a pareto-set through randomized initializations of the GA. For each new trial in an iteration, static geo-attributes are held constant by restricting their search ranges to the basin-specific values, avoiding the specification of geo-attribute values that do not match those of the study basins. SUMMA is then run with the predicted trial parameters and new OF values are obtained, which are added to the emulator training sample to be used in subsequent iterations.

In collaboration with the effort described in Tang et al. (2025), the development of this calibration approach presented several challenges, which were tackled through extensive (albeit ad hoc) testing of different choices in the implementation. For example, one concern was hyperparameter selection, which required balancing the complexity of the emulator to prevent overfitting while ensuring adequate generalization. Hyperparameters for the RF and GA models were tuned using a combination of grid search and cross-validation. The computational demand of the LSE approach was significant; even using an emulator, it still requires conducting a large number of simulations to generate parameter sets based on optimization algorithms, as well as testing them in a computationally expensive LHM. To address this, the number of iterations was minimized while the number of parameter trials per iteration was increased, which we found improved efficiency without sacrificing accuracy. Additionally, we relied on parallel and high-performance computing (HPC) resources from the National Center for Atmospheric Research (NCAR) and engineered the HPC-specific workflow using load balancing in the emulator training and parameter optimization phases to reduce the overall computational cost and wall-clock time. Further discussion of these hyperparameter experiments and workflow development is beyond the scope of this paper and may be tackled in a subsequent publication.

2.4 Experimental design and evaluation approach

2.4.1 Experiments

Our evaluation provides insight into the performance of different aspects of the approach through several experiments. First, we assess the emulator accuracy -- the agreement between the OFs predicted by the emulator and those simulated by the model. Next, we use a small set of metrics to assess the approach in three ways: (1) with the emulator trained on all of the study basins ('all-basin') for the calibration period; (2) with all-basin training for a temporally separate validation period; and (3) with a spatial cross-validation, in which the emulator is trained and separately tested on different parts of the study basin dataset.

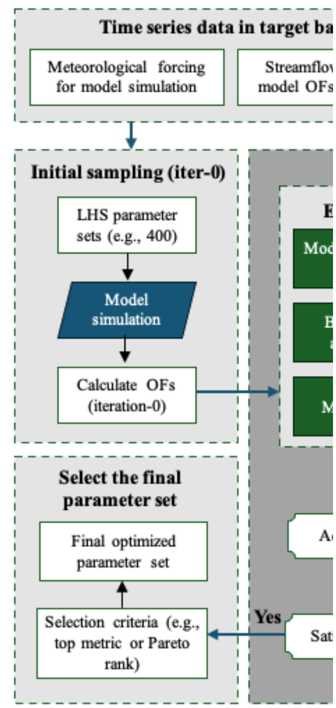
The first experiment compares the LSE and SSE approach results across all 627 basins during the calibration period. In a second experiment, temporal validation was performed by selecting the best performing parameter sets from the calibration period to simulate streamflow for an independent time period (October 2003 to September 2009) to assess the temporal robustness of the calibrated parameters and their ability to generalize under varying meteorological conditions. This period has a similar length to the calibration period and is separated by multiple years, without further considerations imposed.

Deleted: 2024

Deleted: a major

Formatted: Font: Bold, Font color: Black

Moved up [2]: **Figure 2. The flowchart of the parameter optimization in this study, including target parameter selection, initial sampling (**



Deleted:

Deleted: **iteration-0), iterative optimization (iteration-X) using emulators, and the selection of the final parameter set. For emulator inputs, the basin geo-attributes are only used for LSE.**

Deleted: ,

Deleted: 11

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

A third experiment (termed LSE_CV) evaluates the LSE's capability for regionalization in unseen basins through using a spatial cross-validation training and testing approach. The basin dataset was divided through random sampling into five spatially distinct and (roughly) equally-sized folds. In each iteration of the LSE calibration process, four folds (80% of the basins) were used for training the emulator and the remaining fold (20% of the basins) was used for testing. Model parameters (and the emulator-predicted OF values) were predicted by the emulator for the test fold basins based solely on their geo-attributes. The parameter sets with the best emulator-predicted OFs were then selected for model simulation in the test fold basins, and OF results for the five test fold simulations were pooled after each iteration for assessment. Note, due to emulator error, the parameter set selected based on emulator-estimated performance for a test basin was rarely the best performing parameter set among the iteration trial options for that basin (a point discussed further below). The LSE_CV experiment is approximately 5 times more expensive than the others, given that 5 emulators must be trained, and after iteration 0, each iteration involves 5 rather than one model simulation per parameter trial number.

2.4.2 Evaluation metrics and application

The emulator's performance was evaluated using cross-validation techniques and statistical metrics to quantify its ability to predict the OF values based on parameter sets and catchment geo-attributes. As in Tang et al. (2025), we adopted the normalized Kling-Gupta Efficiency ($NKGE'$), a version of the modified Kling-Gupta Efficiency (KGE') (Kling et al., 2012; Beck et al., 2020), as the OF for calibration. $NKGE'$ was chosen to mitigate the influence of outliers, which disproportionately affect the emulator's performance due to the amplified (unbounded) range of KGE' values from poorly performing basins and/or trials, and because the joint evaluation requires standardization across diverse basin flow error ranges. All KGE' and $NKGE'$ values reported in this study are computed based on daily streamflow.

The formulations for KGE' and $NKGE'$ are:

$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (1)$$

$$NKGE' = \frac{KGE'}{(2-KGE')} \quad (2)$$

where r is the linear correlation, β is the bias ratio, and γ is the variability ratio. KGE' ranges from $-\infty$ to 1, whereas $NKGE'$ normalizes this range to $[-1,1]$, which is necessary to balance the information weight of each basin during training. This nonlinear rescaling prevents extremely poor-performing trials from dominating the learning process while preserving their rank order. In contrast to a hard cap on low KGE' values, this smooth rescaling avoids discontinuities in the objective function, improves emulator training stability, and provides a more interpretable optimization surface.

For each of the experiments, we take stock of the calibration performance after each calibration iteration (of 100 trials). This can be done by calculating the evaluation metrics considering each iteration separately, or by calculating them after each iteration and including prior iterations. The iteration-specific evaluation gives insight into the path of the calibration as it seeks

Deleted:
Deleted: set
Deleted: 2024
Deleted: NKGE'),
Deleted:)
Deleted: NKGE'
Formatted: Font: 9 pt
Deleted: KGE'
Deleted: NKGE'
Deleted: ,
Deleted: $KGE'/(2-KGE')$,
Deleted: while
Formatted: Font: Gungsuh
Formatted: Font: +Headings (Times New Roman)
Deleted: 12
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

improved parameters, while the cumulative evaluation shows the overall achievement of the calibration in finding optimal parameters by the end of each iteration. Most model performance results are shown in terms of the KGE' metric (not the optimization metric, which is less familiar to practitioners), and in the form of spatial maps and cumulative distribution functions (CDFs) of KGE' values from all the basins. Both the LSE and SSE calibrations start with the same default parameter configurations (iter-0) to ensure a consistent baseline for comparison. Improvements in KGE' relative to the default model results are analysed in some cases to illustrate the impact of both methods.

3 Results

3.1 Emulator performance

To gauge whether the emulator is successful in learning the model performance metric response to variations in input parameters (and for the LSE, in geo-attributes), we first evaluate emulator performance by comparing its OF predictions ($NKGE'$) to the actual SUMMA model OF values across successive calibration iterations. For each iteration, the emulator-based parameters and OF values are estimated from the simulations of the previous iteration, thus the simulations on which they are tested are independent (were not seen by the emulator previously). Scatter density plots for iterations 1-6, illustrating SSE, LSE calibrations and the LSE_CV experiment, aggregated across all basins, are shown in Figures 3-5, respectively. Figure 3 demonstrates that the SSE approach struggles to improve accuracy in predicting the model OF values progressively across subsequent iterations. While some improvements are observed (e.g., iteration 3), overall performance remains suboptimal, with low to moderate correlations and substantial scatter around the 1:1 line. These results suggest that the SSE approach is limited by a small training sample, resulting in emulator noise and only weakly capturing the relationship between parameters and OFs, without improving significantly as training progresses. It does tend to lead to overall model performance improvements (selected from the ensemble post hoc) as a result of a relatively broad (though inefficient) range of predicted parameter sets in each new iteration, the best of which yield strong performance.

Deleted: modified Kling-Gupta Efficiency (

Deleted: ')

Deleted: iteration

Deleted: portray

Deleted: benefit

Deleted: NKGE')

Deleted: Figs.

Deleted:

Deleted: accurately predict

Deleted:

Deleted:

Deleted: values

Deleted: (a few

Deleted: are often performative).

Deleted: 13¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

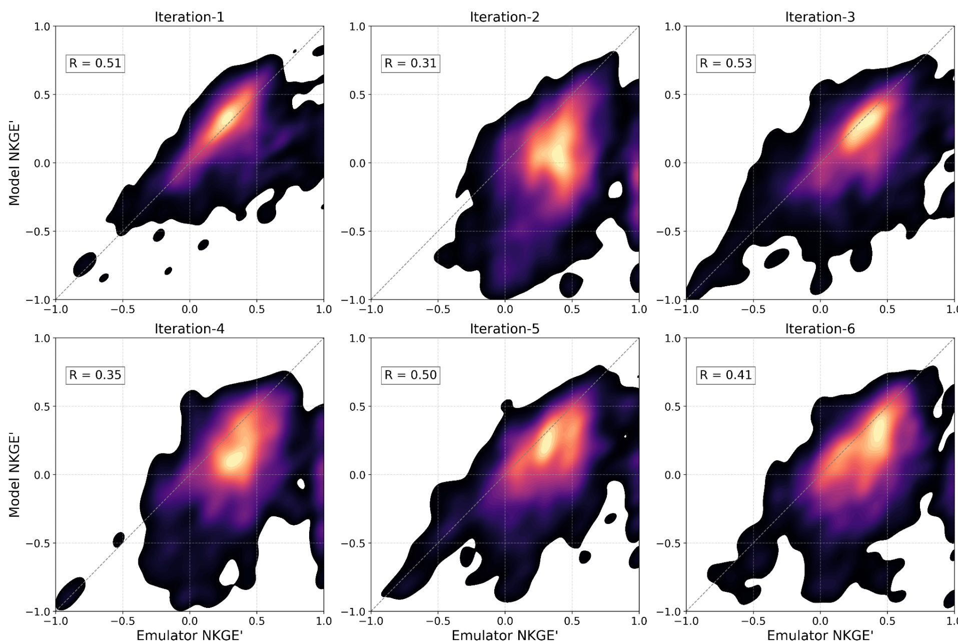
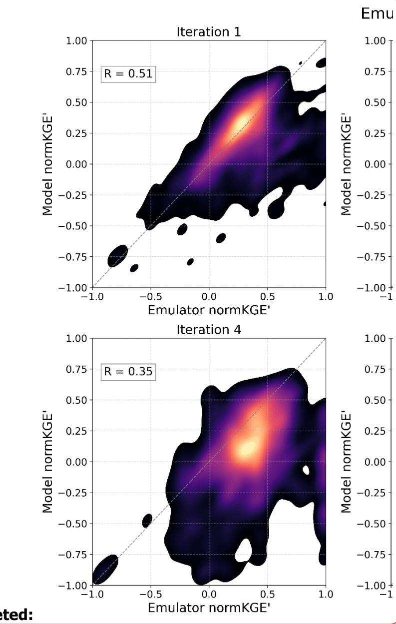


Figure 3. A scatter density plot of emulator-predicted objective function (OF) values (normalized modified Kling-Gupta Efficiency; $NKGE'$) versus real model OF values for the single-site emulator (SSE) approach across six iterations, aggregated across all basins. The Pearson correlation coefficient (R) is shown for each iteration (as in Figures 4 and 5).

Figure 4 demonstrates the LSE approach's superior ability to progressively improve the prediction of the model OF values across iterations. Starting with moderate agreement in iteration 1 (R = 0.71), the emulator steadily improves, achieving strong correlations and reduced scatter as training progresses. Iterations 4-6, with R fluctuating between 0.90 and 0.94, suggest that the emulator's performance may have reached a limit in the available information. The more progressive upward trend of performance underscores the potential benefit of the jointly trained LSE approach relative to the SSE in more comprehensively learning and using the model relationship between parameters and OFs to more effectively search for parameter sets in the individual basins.



Deleted:

Deleted: NKGE')

Formatted ... [51]

Deleted: with their

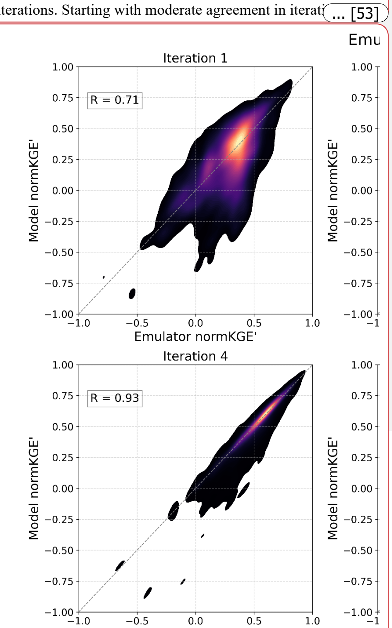
Formatted ... [52]

Deleted: inset.

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: progressive and …uperior ability to accurately predict…rogressively improve the prediction of the model OF values across iterations. Starting with moderate agreement in iterati ... [53]



Deleted:

Deleted: 14

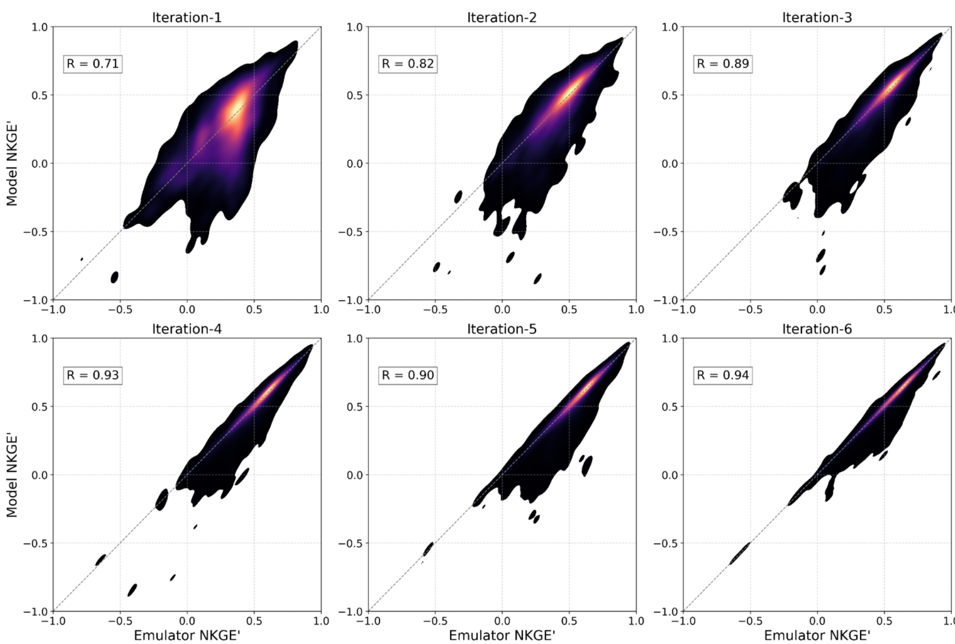


Figure 4. A scatter density plot between emulator-predicted OF values ($NKGE'$) versus real model OF values for the LSE approach across six iterations, aggregated across all basins. The lower-right scatter regions in early iterations reflect emulator overestimation, where predicted performance is high, but actual model performance is poor. This misalignment diminishes as the emulator improves over successive iterations.

Figure 5 evaluates the LSE_CV approach, which represents parameter regionalization in unseen basins through spatial cross-validation (CV). From the initial iterations (e.g., iteration 1, R = 0.43), the emulator's predictive performance rises more slowly, with some vacillation, and plateaus at much lower skill levels than for the LSE of Fig. 4, peaking at R = 0.56 by iteration 5. As expected, the LSE_CV performs worse than the LSE, which sees all basins in calibration, but nonetheless outperforms the SSE approach (Fig. 3). This indicates that the information gained from the large sample of basins provides additional stability in estimating parameters in an unseen basin over the information gained by the SSE (with the same number of trials) -- even when the SSE has learned directly (but only) from that basin.

Deleted: objective function (

Deleted:)

Deleted: NKGE')

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Normal, Space After: 10 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Deleted: with their Pearson correlation coefficient (R) inset

Formatted: Font color: Black

Deleted:

Deleted: appears to plateau

Deleted: Figure

Deleted:

Deleted: Figure

Deleted:

Deleted: 15¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

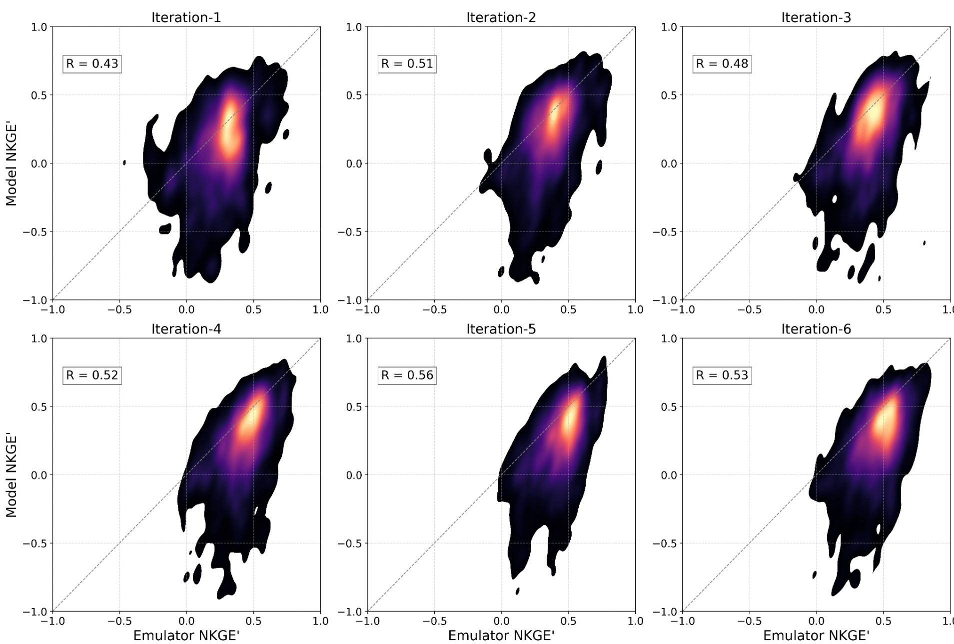


Figure 5. A scatter density plot comparing emulator-predicted OF values ($NKGE'$) to real model OF values for the LSE_CV approach across six iterations, aggregated across all basins.

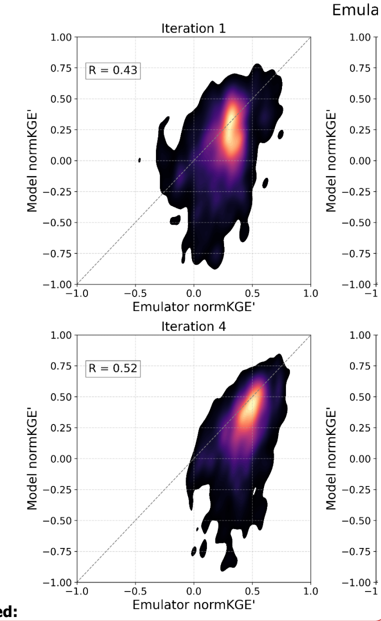
3.2 Calibration performance

The following analyses assess the implications of the apparent emulator strengths and weaknesses for the associated model simulations during calibration. Figure 6 illustrates the CDFs of KGE' for both SSE and LSE calibration approaches across all basins, comparing their evolution from the default parameter set through multiple iterations of calibration. The default configuration provides a baseline with a median KGE' of 0.30, represented by the blue line, and each curve comprises the best 'cumulative' basin model KGE' across all basins after each iteration, including prior iteration results.

The SSE approach (Fig. 6a) shows the SSE calibration results, where the median KGE' improves from the default value of 0.30 to 0.69 by iteration 6 ('iter-6'). The largest gains are observed during iteration 0 (KGE' of 0.52), after which the improvement rate slows. In Figure 6b, the LSE approach begins with significant improvement in iteration 0, attaining a median KGE' across all basins of 0.52. Over subsequent iterations, the LSE approach gains skill, culminating in iteration 6 with a

16

Emula



Deleted:

Deleted: The

Deleted: between

Deleted: objective function (

Deleted:)

Deleted: NKGE')

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted ... [55]

Deleted: with their Pearson correlation coefficient (R) inset.

Formatted: Font color: Black

Deleted: -…6.…('iter-6'). The largest gains are observed during iteration- ... [56]

Deleted: …n Figure 6b, the LSE approach begins with significant improvement in iteration- ... [57]

Deleted: -

Deleted: 16

Formatted: Font color: Black

Formatted ... [54]

median KGE' of 0.76. Overall, the LSE approach outperforms SSE in all iterations. By iteration 6, LSE achieves a median KGE' of 0.76 compared to 0.69 for SSE. Both approaches show diminishing returns after early iterations, but the LSE's joint multi-basin calibration is more effective in progressively identifying performative parameter sets through each iteration. Notably, this superior performance is achieved through training the joint model emulator only 6 times, versus training 627*6 = 3,763 emulators to estimate calibration parameters when using the SSE.

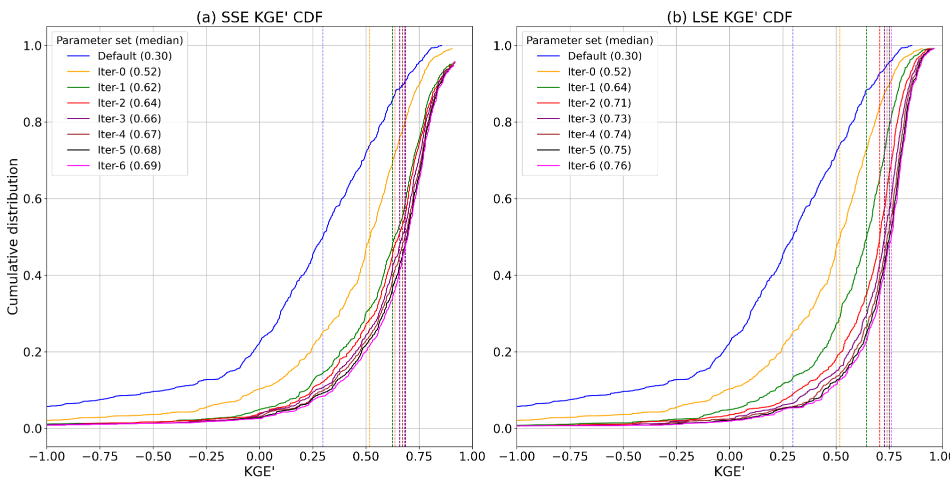


Figure 6. Comparison of calibration performance: cumulative distribution function (CDFs) of modified Kling-Gupta Efficiency (KGE') for (a) SSE and (b) LSE calibration across all basins over six iterations, with median KGE' values noted in the legend. The blue line represents CDF and median KGE' based on the default parameter set for all 627 basins. Both LSE and SSE approaches start with the same iter-0. The x-axis range is set to [–1, 1] for visual clarity, and no normalization or scaling has been applied to KGE'.

The geographic distribution of model performance is shown in Fig. 7, which compares individual basin KGE' values for the calibration period across the CONUS domain for the default parameter configuration, the SSE-based calibration, and LSE-based calibration. The default configuration results of Fig. 7a reflect a degree of indirect calibration (as noted in Section 2.2) and provide a benchmark for this study's calibration improvements. The general pattern of performance, with central US basins showing lower KGE' values than west coast, eastern, and intermountain west basins, is consistent with results shown in numerous other studies based on the CAMELS-US basin dataset, including the first (e.g., Newman et al., 2015).

Deleted: across
Deleted: -
Deleted: a
Deleted: <object>
Deleted: CDF) Comparison
Deleted: KGE'
Formatted: Font: Bold, Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: KGE'
Deleted: .
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: KGE'
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: parameters of
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: iteration
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Font: Not Bold
Deleted: (which
Deleted: Sect.
Deleted: context
Deleted: bains
Deleted: modeling
Deleted: 17
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

The SSE-based calibration KGE' values (Fig. 7b) show marked improvements in hydrological accuracy relative to the default parameters (Fig. 5a), especially in the eastern United States, where high KGE' values are achieved in many basins. However, the results reveal significant spatial variability, with several western basins showing KGE' values below zero. The LSE (Fig. 7c) demonstrates further improvement across nearly all basins, perhaps most notable in the Appalachian basins of the eastern U.S, and fewer basis with KGE' values below zero (as is also clear from Fig. 6). A timeseries-based illustration of the performance of LSE-based calibration for two basins is provided in Appendix Fig. B1 and Fig. B2. In Figure B1, calibration achieves a high-quality simulation of observed streamflow at daily and seasonal mean levels, while Fig. B2 shows a basin with notable daily and monthly improvements from calibration over the default simulation, but errors remaining in the seasonality of simulated streamflow.

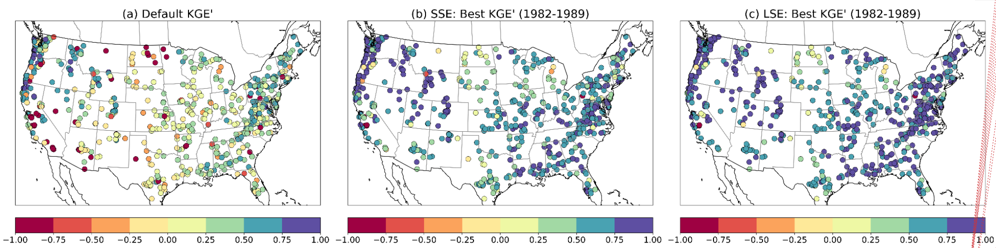
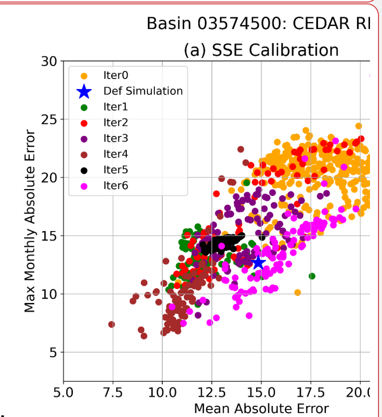


Figure 7 Comparison of KGE' values for (a) default configuration, (b) SSE and (c) LSE calibration across the CONUS (1982-1989).

Although such CONUS-wide summaries of performance are useful, the contrast between SSE and LSE calibration performance can be stark when reviewed at the level of individual basins. Figure 8 shows an example of the parameter sampling trajectories of the SSE and LSE calibrations, as assessed using two model diagnostic performance metrics that were not used as the calibration objective: the mean absolute daily streamflow error, and a seasonality metric defined as the maximum long-term mean monthly absolute streamflow error -- i.e., the largest error in long-term mean monthly flow. The LHS-generated 'iter-0' parameters provide the starting point for searching the parameter space, which the SSE and LSE duly improve upon, each recommending parameters that lead to superior model performance in iteration 1. In subsequent iterations, however, the SSE parameter set performance regresses and vacillates, while the LSE parameter recommendations generally advance. For the SSE, the successive iterations of the parameter search have a broader performance spread and are more scattershot, while the LSE search tends to yield steady refinement in model performance. This behavior certainly varies by location, but the general character of each approach shown in Fig. 8 is consistent with the rates of improvement shown in Fig. 6.

Deleted: oriented
Deleted: .
Deleted: KGE'
Formatted: Font: Bold, Font color: Black
Formatted: Font: Bold, Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Formatted ... [59]
Deleted: *<object>*
Formatted: Font color: Black
Deleted: Iter
Deleted: foundation
Deleted: Iteration
Deleted: recommendations regress and vacillate in model
Deleted: tend to
Deleted: are
Deleted: progress
Deleted: corroborated by



Deleted:
Formatted ... [60]
Formatted: Font: Bold, Font color: Black
Deleted: 18
Formatted: Font color: Black
Formatted ... [58]

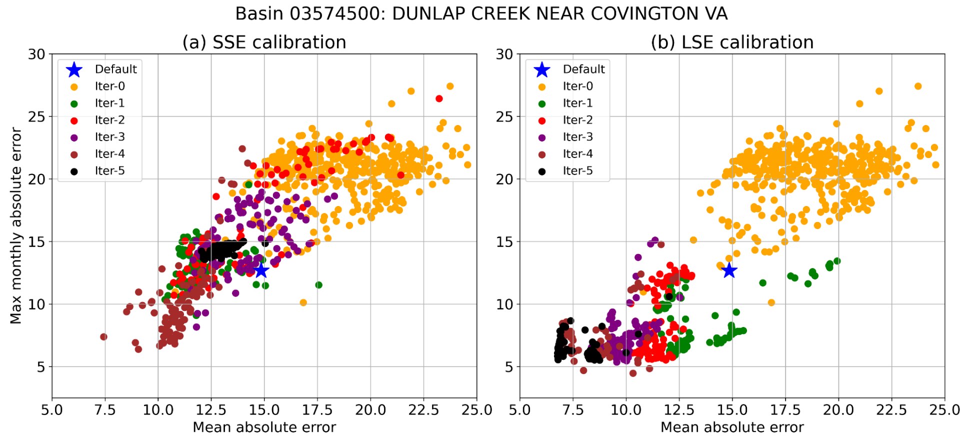


Figure 8 Illustration of the (a) SSE and (b) LSE parameter calibration progress across successive iterations, as measured by two metrics not used in calibration. Better model performance for both metrics (i.e., lower error) is found in the bottom left corner of each plot. Iter-0 contains 400 parameter sets, and subsequent iterations contain 100. The default simulation is from a previous SUMMA application after individual basin optimization with the DDS algorithm.

Incidentally, Fig. 8 also illustrates that despite recent critiques against calibrating hydrology models to integrated streamflow metrics such as the Nash Sutcliffe Efficiency (NSE) and KGE (e.g., Brunner et al., 2021; Knoben et al., 2019), such metrics can be effective in jointly optimizing hydrology model performance across multiple metric dimensions, hence their long-standing popularity in practice.

3.3 Temporal validation of the SSE and LSE approaches

Temporal validation (during the independent 2003–2009 period) of the SSE and LSE approaches shows that, as expected, calibration parameter performance for both falls relative to the calibration period. Figure 9 shows the KGE' CDF curves with the median of the distribution of each basin's best performance (including all prior iterations) reaching 0.65 and 0.69 or the SSE and LSE, respectively, by the 6th iteration (best scores include the best of all prior iterations). The lower reduction in validation scores for the SSE than for the LSE may or may not be notable (i.e., it may be a study-specific result); if significant, it suggests that the best selected SSE parameters, even with lower overall performance in both calibration and temporal validation, may be slightly more robust to meteorological variability than those from the LSE.

Deleted: .
Formatted: Font: Bold, Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: Iteration
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: Figure
Deleted: NSE
Deleted: KGE
Deleted: ('shrinkage')
Deleted: 19
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

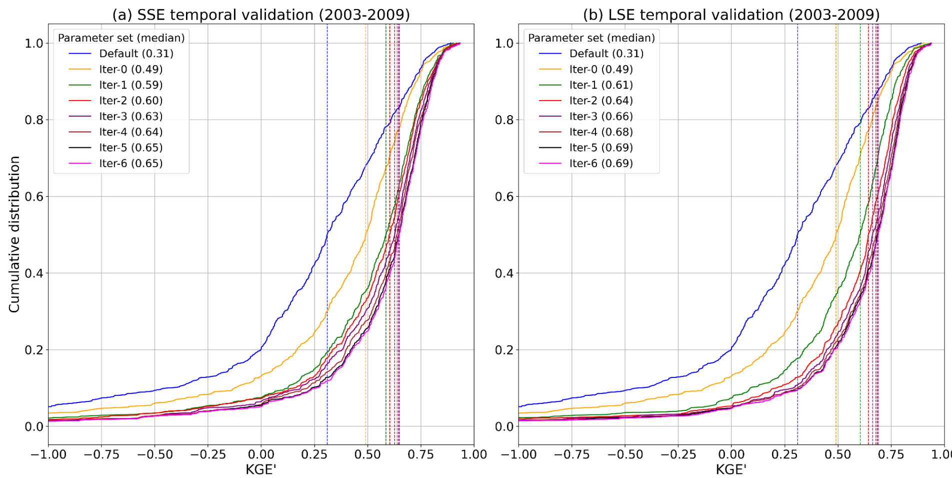
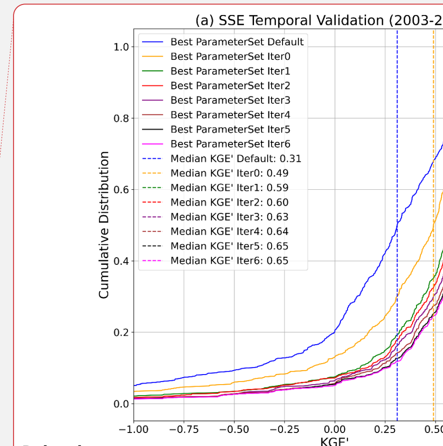


Figure 9 Comparison of temporal validation (during the independent 2003–2009 period) performance: CDFs of the best KGE' for (a) SSE and (b) LSE calibration across all basins over six iterations. The blue line represents the CDF and median KGE' based on the default parameter set over all basins.

The associated maps in Fig. 10 show the geographic distribution of the temporal validation median KGE' scores for the SSE and LSE approaches, and their change in value relative to their respective calibration scores. The pattern of values for the validation scores (parts a and c) are broadly similar, which is notable given that the LSE represents their joint calibration in contrast to the individual attention that each basin receives in the SSE. Higher KGE' values are observed predominantly in the mountainous portions of the western US, and in the midwest and eastern US. Lower KGE' values are more prevalent in the southwestern US and northern plains region.

In general, areas that calibrated well under either method (Fig. 7) tended to hold up well in temporal validation. Plot parts (b, d) show the difference between the validation and best-calibrated KGE', and regions with smaller differences indicate that the calibrated parameter sets generalized well in time. The LSE-calibrated parameters led to slightly greater loss in skill in validation than did the SSE-calibrated parameters, and this effect was pronounced in the more challenging calibration regions noted above. Although the cause of this effect is unclear, a likely culprit is overtraining -- i.e., that the LSE harnesses more sequence-specific information than the SSE to gain a stronger calibration and validation performance. That said, if the objective of a modeling application is to calibrate a LHM for use over a large number of measured catchments, this analysis nonetheless suggests that the LSE would provide both efficiency and skill improvements over the traditional basin-specific calibration.



Deleted:

Deleted: .

Deleted: KGE'

Formatted: Font: Bold, Font color: Black

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: for the (a) SSE Calibration and (b) LSE calibration.

Deleted: KGE'

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: parameters

Deleted:

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: Figure

Deleted: obtained during calibration

Deleted: site

Deleted:

Formatted: Font: Bold

Deleted: 20

Formatted: Font color: Black

Formatted ... [61]

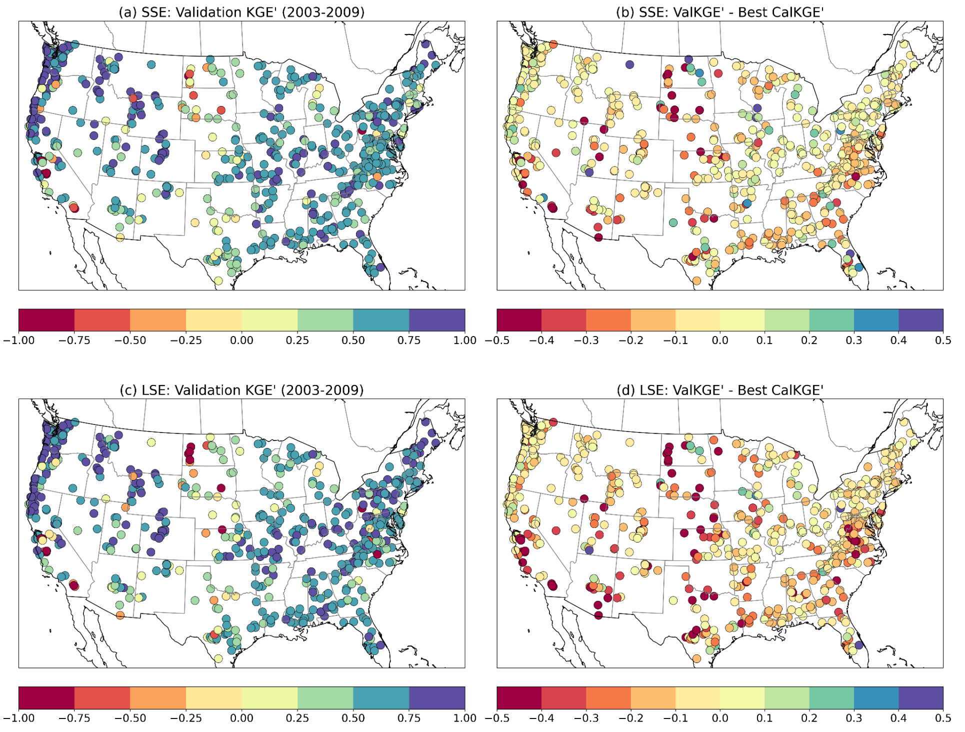


Figure 10 Temporal validation performance (during the independent 2003–2009 period) shown as spatial distribution across the CONUS. Panels illustrate (a) median KGE' values for SSE calibration, (b) difference between validation and calibration median KGE' values for SSE (c) median KGE' values for LSE calibration, and (d) difference between validation and calibration median KGE' values for LSE.

3.4 Spatial cross-validation

Figures 11 and 12 present results of the LSE_CV experiment, which tests SUMMA simulations using the parameter sets in each iteration (cumulative with prior iterations) that had the best emulator-predicted KGE' values in each basin. Each calibration iteration produces 100 recommended new parameter sets; thus, there is a need to decide *a priori* which set to select

Deleted: <*object*>**Figure 10. Temporal validation of KGE' for SSE and LSE calibration across the CONUS.**

Deleted: 21

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

for testing in the unseen basins (as noted in Sect. 2.4.1). Because the emulator has some skill in estimating model performance given different parameter sets, we use its performance estimate as a basis for the selection. We briefly experimented with alternative selection strategies, none of which were superior, and also evaluated whether transferring a small ensemble (top 5-20 parameter sets based on emulator ranking) leads to better mean performance (it does, but ensemble modeling is not the focus of this effort). We compare LSE_CV results to those from LSE calibration parameter sets selected in the same way (based on the highest emulator-predicted KGE' values), which leads to slightly lower performance than assessing the best actual model KGE' values, as shown in Fig. 6a.

In Figure 11, the contrast between the LSE_CV and LSE_all (from calibration) when using emulator-ranked parameter sets is striking. In both cases, median KGE' values generally rise over six iterations, indicating that the approaches can find improved parameters as they are run repetitively. Not surprisingly, the LSE_CV skill falls relative to the LSE all, and plateaus quickly. The LSE_all calibration achieves higher median emulator-predicted KGE' values compared to LSE_CV in every iteration, with a final median value of 0.73 for LSE_all compared to 0.43 for LSE_CV by iter-6. This discrepancy was larger than expected, and we review potential causes further in the Discussion section.

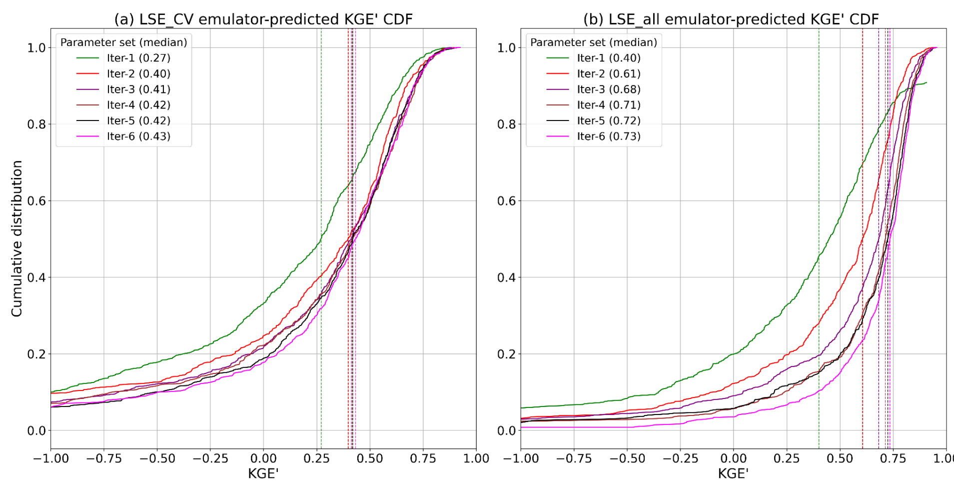
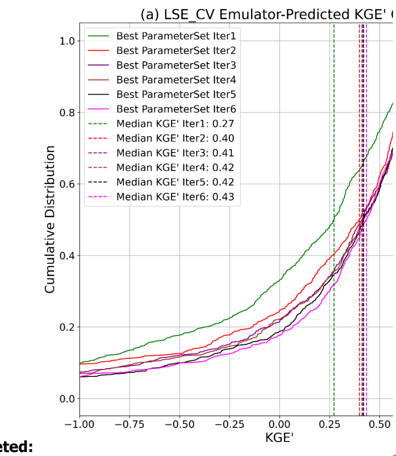


Figure 11. Comparison of calibration performance: CDFs of the best KGE' for (a) LSE_CV across all folds and (b) LSE_all across all basins over six iterations with median KGE' values.

Deleted:).

Deleted: iteration

Deleted: may arise for reasons which are explored



Deleted:

Deleted: .

Deleted: cumulative distribution function (CDF) Comparison

Deleted: KGE'

Formatted: Font: Bold, Font color: Black

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: KGE'

Formatted: Font color: Black

Deleted: 22¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Useful context for the emulator-predicted LSE_CV results is provided in Appendix Fig. B3, which shows the best model performance selected post facto for unseen test basins using LSE_CV parameter sets (recall that 100 parameter sets are estimated in each iteration, but only one is tested in Fig. 11a). Achieving a median KGE' value of 0.72 after six iterations, the best performing results suggest that the LSE is capable of finding competitive parameter sets in unseen basins; rather, the challenge is knowing in advance which of the estimated parameter sets are best to use in parameter transfer. For applications in which an ensemble of predicted, pre-trained parameter sets is useful, this finding is useful, especially if the fitness of an ensemble of emulator-predicted parameter sets could be judged through additional relevant criteria (such as performance at indirectly related or downstream gages, or assessment of consistence with related hydrologic statistics or signatures), aiding the regionalization task.

Figure 12 shows the geographic distribution of performance results for the LSE_CV, the LSE in calibration (which uses emulator-ranked best parameter sets, versus best model outcomes), and the LSE_CV difference from the default model performance (parts a-c, respectively). The calibrated LSE-based KGE' values are significantly high across most basins, demonstrating that substantial model performance may be achieved by directly calibrating parameters across all basins, benefiting from the full training dataset. The more uniform distribution of higher KGE' values across different regions, especially in the central and eastern U.S., highlights the LSE's ability to enhance accuracy over diverse hydroclimatic conditions. The LSE_CV approach, illustrated in Fig. 11a, shows improvement over the default model parameters for a majority of basins, with approximately 62% of basins achieving better values than the default configuration. However, there are multiple basins where LSE_CV underperforms compared to default parameters, particularly in complex, cold or arid regions.

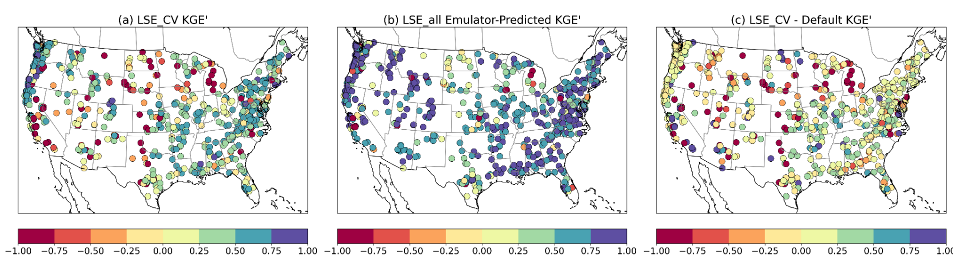


Figure 12. Comparison of performance (median sample KGE') between the LSE_CV, which uses spatial cross-validation for regionalizing parameters to unseen basins, and the LSE_all calibration, which uses information from all basins. The LSE_CV performance difference from the default parameter performance is also shown.

Deleted: Noteworthy
Deleted: that
Deleted: achieved in
Deleted: parameters
Deleted: not incapable
Deleted: *post-facto*
Deleted:
Deleted: *<object>*
Deleted: .
Deleted: KGE')
Formatted: Font: Bold, Font color: Black
Formatted: Font: Bold, Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Normal
Deleted:
Formatted: Font color: Black
Deleted: 23¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

4 Discussion and conclusions

This study investigates whether the challenge of calibrating large-domain implementations of complex, expensive PB land/hydrology models can be tackled through similar strategies to those now being advanced in the AI contexts (i.e., ML, DL and differentiable modeling). Such studies have shown that jointly training the DL or simple conceptual models (made differentiable) over large samples of catchments is not only viable but is recommended over individually calibrating such models on single basins. The reverse had generally been found for complex LHMs, but the recent emergence of ML model emulation strategies for complex models has provided an avenue for reassessing this consensus. In collaboration with a companion effort described in Tang et al. (2025), which focused on the CTSM land model, we develop and assess a large-sample emulator (LSE) based strategy for calibrating the SUMMA-mizuRoute modeling approach across CONUS watersheds. While our study focuses on small-to-medium basins in the CAMELS dataset, the LSE approach is being designed for application to large domains (regional to continental to global scale). Applying the emulator-guided calibration strategy to such larger regions may require adjustments to account for greater heterogeneity in factors such as spatial scale, dominant processes and land forms, flow routing complexity, meteorological input patterns, among others.

Our findings are, in short, promising. They suggest that a large-sample model response emulation approach has potential to become a preferred option for calibrating complex PB models over large domains as it has for ML and other AI-era modeling approaches. The generally higher performance achieved by the LSE relative to the SSE indicates that large-sample calibration can more effectively learn the model response surface to parameters, even for complex models, than is possible using local information alone. As noted by Kratzert et al. (2019) and others, the inclusion of static catchment attributes allows large-sample approaches to localize parameter influence and account for hydroclimatic variability across basins, which in turn leads to more efficient joint calibration and better overall model performance. Our SSE, while effective for some individual basin calibrations, did not reach the accuracy of the LSE when applied across diverse conditions, and pursued less efficient parameter search trajectories. In practice, the scalability of a strategy that jointly trains a single, low-cost model emulator for model calibration to yield usable parameter estimates for hundreds (or more) catchments at once is arguably attractive, given the main alternative for parameter regionalization is individually calibrating those catchments only as the first step toward training a separate parameter transfer scheme.

While the LSE strategy still requires a set of PB model simulations for training, it offers a substantial computational advantage over traditional calibration approaches by drastically reducing the number of required simulations in subsequent iterations. Rather than incurring the cost of repeated full-model evaluations across basins, the emulator enables efficient exploration of the parameter space with far fewer model runs. As described in Section 2.3.2, we further improved efficiency by increasing the number of parameter trials per iteration while reducing the total number of iterations—an approach that maintained accuracy while accelerating convergence. This balance between emulator fidelity and computational cost demonstrates the practicality of the method for large-domain hydrological modeling. Looking ahead, we are optimistic that future enhancements

Deleted: 5
Deleted: Conclusions
Deleted: process-based
Deleted: differentiablee
Deleted: ,
Deleted: true
Deleted: 2024
Formatted: Font: 9 pt
Deleted: struggled to
Deleted: of
Deleted: at least a fraction of
Deleted: 24
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

such as adaptive sampling, transfer learning, or cross-domain emulator reuse could further reduce the up-front simulation demand, opening new possibilities for applying this approach to even more complex or higher-resolution modeling systems.

The results of the LSE calibration (median KGE' = 0.76) and validation (0.69) in this study are competitive with published modeling studies using all or parts of the CAMELS catchment collection, though comparisons are inexact due to differences in factors such as basin selection, validation periods, and optimization objectives. For example, Feng et al. (2022) reported median temporal validation NSEs ranging from 0.62 to 0.75 for jointly calibrated DL and differentiable learning models, while Newman et al. (2015; 2017) achieved sample median NSE scores around 0.74 for calibration and 0.60–0.70 during temporal validation, using much simpler conceptual models (Sacramento and Snow-17), all individually trained.

Yet in other regards, such as advancing capabilities for prediction in ungauged basins, it is also clear from these experiments that further understanding and improvements are needed. The median KGE' score of 0.43 achieved in spatial cross validation is inadequate for use in many regionalization applications, though we believe it is on par with what is currently achievable for complex PB models using site-specific basin calibration followed by similarity-based regionalization. For instance, the performance is near the ungauged basin evaluation over CONUS reported in Song et al. (2025) for the US National Water Model 3.0, at KGE = 0.467. It moderately lags a new differentiable physics-informed ML model (δHBV2.0δUH) at 0.553 in the same study and considerably lags results from pure DL approaches -- e.g., the impressive median NSE of 0.69 achieved by the PUB LSTM of Kratzert et al. (2019). Such studies are not controlled comparisons with this one or each other but nonetheless provide useful context.

The lower performance of the LSE_CV relative to the LSE in spatial cross-validation validation may result from a combination of factors, including unexplained variability in the hydroclimatic settings and model response, and some overtraining to sample characteristics, which include meteorological input errors. This can also be attributed to (1) the reduced number of training samples in LSE_CV compared to LSE_all (20% fewer basins), and (2) the inherent difficulty of regionalizing parameters for ungauged locations, a well-documented challenge in hydrological modeling (Patil and Stieglitz, 2015). Overall, the sample size used in this study (627) may be inadequate for high-quality regionalization. Nonetheless, we are optimistic that with further exploration and development, the regionalization performance of calibration based on an LSE approach will improve. We have not yet investigated potential refinements such as the feature engineering and selection of static geo-attributes to enhance transferability. Here we adopted those used in Tang et al. (2025), and these contained inconsistencies (e.g., meteorological attributes were not based on the model forcing dataset climatology). Using more catchments in training with better screening for representativeness is likely to strengthen the regionalization, especially as some basins were later found to have erroneous streamflow observations. Supporting work in the study (not shown here) indicated that some parts of the US improve when restricting training to a similarity-based watershed selection, while others fare better when trained on the full sample – thus a blend of similarity-based and full-domain emulation may prove superior. The selection strategy for predicted parameter sets to transfer and various hyperparameter choices also warrant further investigation. The question of spatial scale

Deleted: Nash Sutcliffe Efficiency (
Deleted:)
Deleted: likely
Deleted: 2024
Deleted: KGE
Deleted: ,
Deleted: ,
Deleted: explored
Deleted: simply
Deleted: 2024
Deleted: This effort along with Tang et al. (2024) are initial attempts at implementing such a ML-based joint calibration strategy, and raise as many questions as they answer.
Deleted: 25¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

consistency between the training basins and the regionalization target basins may be critical to the success of its application in real-world large-domain uses; we suspect that the LSE approach remains robust under moderate scale inconsistencies, but further research is needed to understand these limits of spatial generalization.

This effort along with Tang et al. (2025) represent initial forays into implementing such a ML-based joint calibration strategy for process models, and each raises as a suite of compelling papers that are beyond the current paper's scope. This paper focuses on introducing, outlining and testing a new large-sample emulator framework, which necessitated substantial dataset, model and workflow development effort, while benchmarking the LSE against a logical baseline, the SSE, and qualitatively comparing it to other related studies using LSTMs, conceptual model and hybrid/differentiable models. We recognize the broader momentum within the ML/DL hydrology community toward methodological intercomparison and refinement, and look forward to undertaking such broader controlled comparisons and studies of methodological choices that were out of scope for this paper. We applied the method to lumped basin-scale PB model configurations for simulating streamflow, but the emulator framework itself is generalizable and could easily be adapted to models with different spatial structures, including gridded domains, levels of complexity, and to multivariate model fluxes and states.

Overall, we hope that these findings will update conventional wisdom about the ability of complex PBLHMs to compete with simpler conceptual models in performance, given that our temporal validation across hundreds of basins is on par with that of other published CAMELS-based conceptual modeling studies. Perhaps more importantly, we show that the power of large-sample model training underpinning recent advances in ML hydrology is extensible to complex PB hydrology models as well. We believe the work takes an important step toward addressing the longstanding challenge of applying such models of prediction in ungauged basins. With national water agencies and global modeling initiatives for land/hydrology and climate analysis and prediction continuing to seek unique multivariate insights from complex PB land/hydrology modeling approaches, we encourage further exploration of possibilities in this direction.

Appendix A. SUMMA configuration

Table A1 SUMMA default model decisions (physics configuration) for this study.

Model decision	Chosen option	Model decision description
soilCatTbl	STAS	Soil-category dataset
vegeParTbl	MODIFIED_IGBP_MODIS_NOAH	Vegetation category dataset
soilStress	NoahType	Function for soil moisture control on stomatal resistance

Deleted: process-based LHMs

Deleted: process-based

Deleted: process-based

Deleted: .

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Formatted: Space After: 0 pt

Deleted: soil

Formatted: Space After: 0 pt

Formatted Table

Deleted: vegetation

Formatted: Space After: 0 pt

Deleted: function

Formatted: Space After: 0 pt

Deleted: 26

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

stomResist	BallBerry	Function for stomatal resistance
num_method	itertive	Choice of numerical method
fDerivMeth	analytic	Method used to calculate flux derivatives
LAI_method	specified	Method used to determine LAI and SAI
f_Richards	mixdform	Form of Richard's equation
groundwatr	bigBuckt	Choice of groundwater parameterization
hc_profile	constant	Choice of hydraulic conductivity profile
bcUpprTdyn	nrg_flux	Type of upper boundary condition for thermodynamics
bcLowrTdyn	zeroFlux	Type of lower boundary condition for thermodynamics
bcUpprSoiH	liq_flux	Type of upper boundary condition for soil hydrology
bcLowrSoiH	drainage	Type of lower boundary condition for soil hydrology
veg_traits	Raupach_BLM1994	Parameterization for vegetation roughness length and displacement height
canopyEmis	difTrans	Parameterization for canopy emissivity
snowIncept	lightSnow	Parameterization for snow interception
windPrfile	logBelowCanopy	wind profile through the canopy
astability	louisinv	Stability function
canopySrad	BeersLaw	Canopy shortwave radiation method
alb_method	conDecay	Albedo representation
compaction	anderson	Compaction routine
snowLayers	CLM_2010	Method to combine and sub-divide snow layers
thCondSnow	jrdn1991	Thermal conductivity representation for snow
thCondSoil	funcSoilWet	Thermal conductivity representation for soil

Deleted: function
Formatted ... [64]
Deleted: choice
Formatted ... [65]
Deleted: method
Formatted ... [66]
Deleted: method
Formatted ... [67]
Deleted: form
Formatted ... [68]
Deleted: choice
Formatted ... [69]
Deleted: choice
Formatted ... [70]
Deleted: type
Formatted ... [71]
Deleted: type
Formatted ... [72]
Deleted: type
Formatted ... [73]
Deleted: type
Formatted ... [74]
Formatted ... [75]
Deleted: parameterization
Formatted ... [76]
Deleted: parameterization
Formatted ... [77]
Deleted: parameterization
Formatted ... [78]
Formatted ... [79]
Formatted Table ... [80]
Deleted: stability
Formatted ... [81]
Deleted: canopy
Formatted ... [82]
Deleted: albedo
Formatted ... [83]
Deleted: compaction
Formatted ... [84]
Deleted: method
Formatted ... [85]
Deleted: thermal
Formatted ... [86]
Deleted: thermal
Deleted: 27¶
Formatted ... [62]
Formatted ... [63]

spatial_gw	localColumn	Method for the spatial representation of groundwater
subRouting	timeDlay	Method for sub-grid routing

Table A2 Geo-attributes used in the large-sample emulator (LSE) training.

Attribute Name	Relevance	Description	Unit
mean_elev	Topography	Catchment mean elevation	m above sea level
mean_slope	Topography	Catchment mean slope	m/km
area_gauges2	Topography	Catchment area (GAC)	km^2
p_mean	Climate	Mean daily precipitation	mm/day
pet_mean	Climate	Mean daily PET (estimated)	mm/day
aridity	Climate	Aridity (PET/P ratio)	-
p_seasonality	Climate	Seasonality and timing	-
frac_snow	Climate	Fraction of precipitation as snow	-
high_prec_freq	Climate	Frequency of high precipitation	days/year
high_prec_dur	Climate	Average duration of high precipitation	days
low_prec_freq	Climate	Average frequency of low precipitation	days/year
low_prec_dur	Climate	Average duration of low precipitation	days
frac_forest	Landcover	Forest fraction	-
lai_max	Landcover	Maximum monthly LAI	-
lai_diff	Landcover	Difference between max and min LAI	-

Deleted: method
Formatted ... [89]
Deleted: method
Formatted ... [90]
Deleted: .
Formatted ... [91]
Formatted ... [93]
Formatted ... [92]
Formatted ... [94]
Deleted: catchment
Formatted ... [95]
Deleted: catchment
Formatted ... [96]
Deleted: catchment
Formatted ... [97]
Deleted: mean
Formatted ... [98]
Deleted: mean
Formatted ... [99]
Deleted: aridity
Formatted ... [100]
Formatted ... [101]
Deleted: seasonality
Formatted ... [102]
Deleted: fraction
Formatted ... [103]
Deleted: frequency
Formatted ... [104]
Deleted: average
Formatted ... [105]
Deleted: average
Formatted ... [106]
Deleted: average
Formatted ... [107]
Deleted: forest
Formatted ... [108]
Deleted: maximum
Formatted ... [109]
Deleted: difference
Deleted: 28
Formatted ... [87]
Formatted ... [88]

dom_land_cover	Landcover	Dominant land cover type	-
dom_land_cover_frac	Landcover	Fraction of the catchment area of dominant land cover	-
soil_depth_pelletier	Soil	Depth to bedrock	m
soil_depth_statsgo	Soil	Soil depth (maximum)	m
soil_porosity	Soil	Volumetric porosity	-
soil_conductivity	Soil	Saturated hydraulic conductivity	cm/h
max_water_content	Soil	Maximum water content	m
sand_frac	Soil	Sand fraction	%
silt_frac	Soil	Silt fraction	%
clay_frac	Soil	Clay fraction	%
carbonate_rocks_frac	Geology	Fraction of the catchment with carbonate rocks	-
geol_permeability	Geology	Subsurface permeability	m^2

Table A3 Summary of default and LSE-calibrated parameter values across all basins, including the percent change from default values and the range (min–max) of LSE-calibrated values. For parameters with basin-specific default values, the default median is used for comparison.

Parameter name	Default	LSE median	% Change	LSE min	LSE max	Default type
k_soil	2.92e-06	1.29e-05	+342.3	7.41e-08	0.0098	Varying
theta_sat	0.38	0.57	+49.0	0.247	0.7	Varying
aquiferBaseflowExp	1.24	1.69	+36.2	1	4	Varying
aquiferBaseflowRate	0.0789	0.0426	- 46.0	0.0001	0.1	Varying
qSurfScale	5.47	3.47	-36.6	1	19.99	Varying

Deleted: dominant
Formatted: Space After: 0 pt
Deleted: fraction
Formatted: Space After: 0 pt
Deleted: depth
Formatted: Space After: 0 pt
Deleted: soil
Formatted: Space After: 0 pt
Deleted: volumetric
Formatted: Space After: 0 pt
Deleted: saturated
Formatted: Space After: 0 pt
Deleted: maximum
Formatted: Space After: 0 pt
Deleted: sand
Formatted: Space After: 0 pt
Deleted: silt
Formatted: Space After: 0 pt
Deleted: clay
Formatted: Space After: 0 pt
Deleted: fraction
Formatted: Space After: 0 pt
Formatted: Space After: 0 pt
Deleted: subsurface
Deleted: 29¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

summerLAI	3.48	4.8	+37.9	0.0107	19.95	Varying
frozenPrecipMultip	0.855	0.795	-7.0	0.5	2.499	Varying
routingGammaScale	51902	30970	-40.3	360.02	86399.92	Varying
routingGammaShape	2.886	1.55	-46.3	1	4.99	Varying
Fcapil	0.0143	0.0222	+55.6	0.009	0.1099	Varying
tempCritRain	273.16	271.74	-0.5	270.16	276.16	Constant
heightCanopyTop	4.67	5.46	+16.9	0.154	57.9	Varying
windReductionParam	0.28	0.23	-19.2	0.05	0.9995	Constant
vGn_n	1.51	1.34	-11.1	1.3	4	Varying

Table A4 Summarizing the key hyperparameters used in our calibration framework.

Method	Hyperparameter	Value / Setting	Notes
Genetic Algorithm (GA)	Population size	100	
	Crossover probability	0.9	SimulatedBinaryCrossover
	Crossover eta	15	SimulatedBinaryCrossover
	Mutation eta	20	PolynomialMutation
	Sampling method	FloatRandomSampling	Continuous variables
	Stopping criterion	Max generations	No early stopping used
Random Forest (RF)	n_estimators	100	Number of trees
	max_depth	40	Maximum depth of each tree

Deleted: 30

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Appendix B

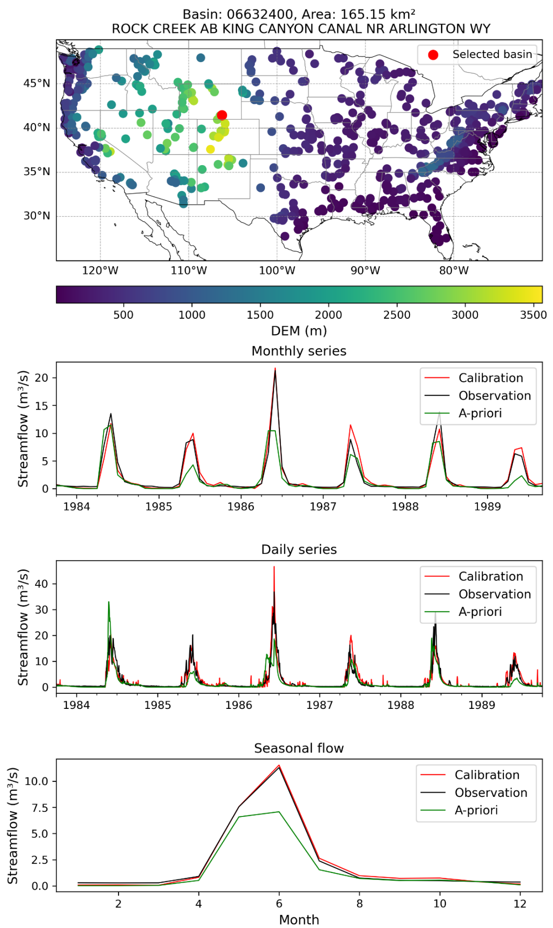
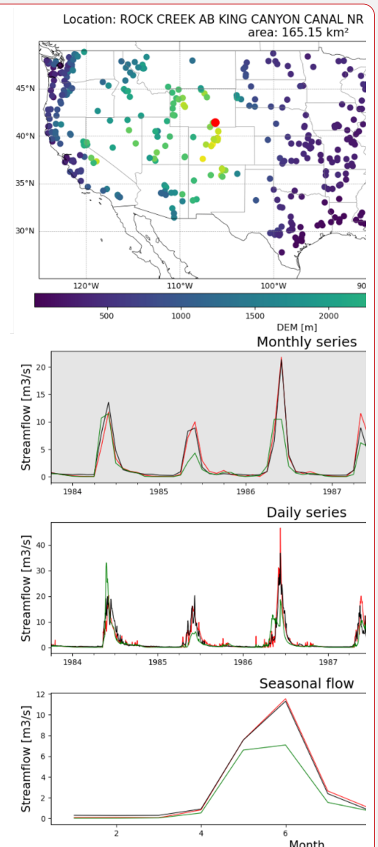


Figure B1 An example basin illustrates how simulated streamflow using LSE calibrated parameter values aligns significantly better with observations compared to simulations using *a priori* (i.e., default) parameters.



Deleted:

Deleted: .

Deleted: illustrating

Formatted: Font: Bold, Font color: Black

Formatted ... [110]

Formatted ... [111]

Deleted: -

Formatted ... [112]

Deleted: 31

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

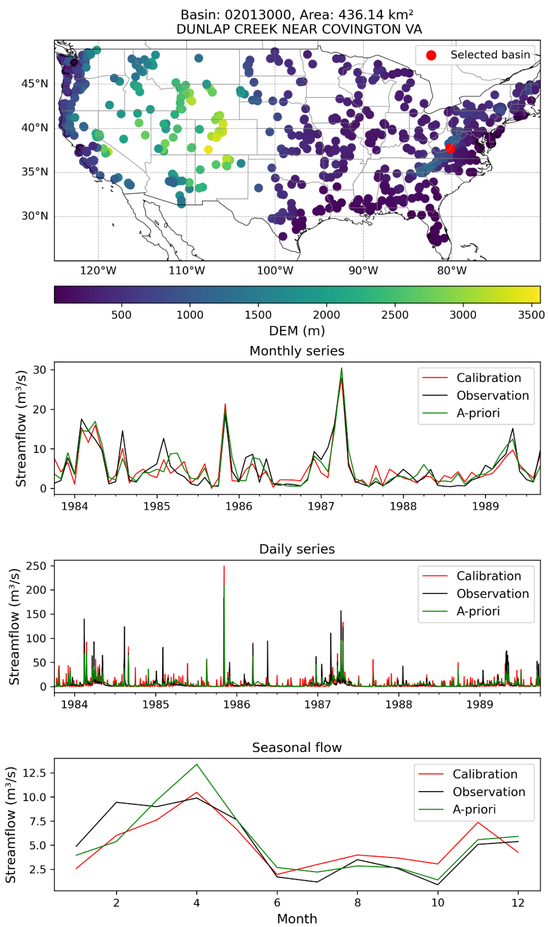


Figure B2 Same with Fig. B1 but showing an example basin where the LSE calibration shows lesser improvement regarding the seasonal streamflow compared to *a priori* (i.e., default) parameters, while for the daily and monthly series, the improvement is still notable.

Deleted: *<object>*

Deleted: .

Formatted: Font: Bold, Font color: Black

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Deleted: -

Formatted: Font: Italic, Font color: Black

Formatted: Font: Italic, Font color: Black

Formatted: Font color: Black

Formatted: Font: Bold

Deleted: 32

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

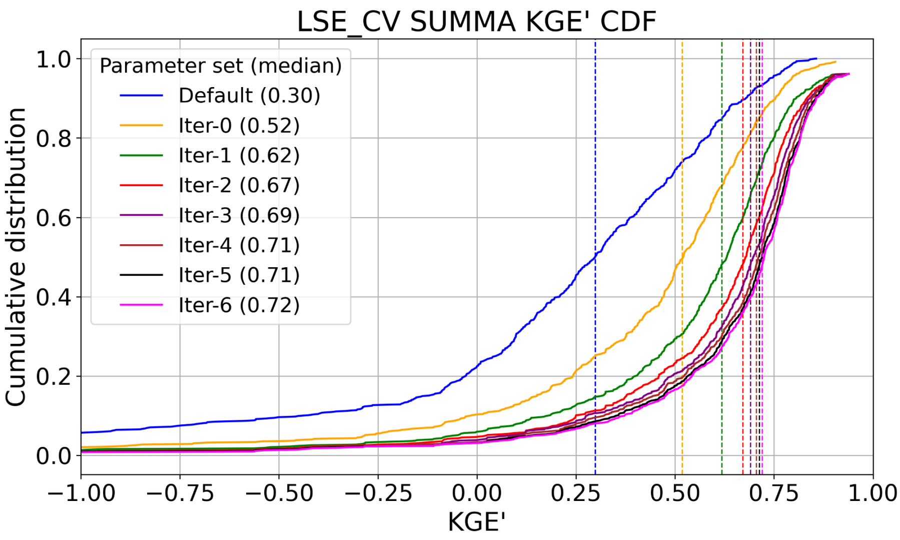


Figure B3 Comparison of the best model KGE' CDFs for LSE_CV over six iterations, selected post facto, and including sample median KGE' values.

Code and dataset availability

The SUMMA model is available at https://github.com/CH-Earth/summa/ and the mizuRoute model is available from https://github.com/ESCOMP/mizuRoute. The original CAMELS dataset is available at https://gdex.ucar.edu/dataset/camels.html. EM-Earth forcings are available at https://doi.org/10.20383/102.0547. ERA5-Land data is available at https://doi.org/10.24381/cds.e2161bac. The LSE-based SUMMA optimization codes are available at https://github.com/NCAR/opt_landhydro and relevant data sets are available at https://doi.org/10.5281/zenodo.16422768.

Deleted: <object>

Deleted: .

Deleted: post-facto

Deleted: KGE'

Formatted: Font: Bold, Font color: Black

Formatted: Font: Bold, Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Formatted: Font color: Black

Deleted: KGE'

Deleted:

Formatted: Font color: Black

Deleted: data

Deleted: https://github.com/CH-Earth/summa/

Deleted: https://github.com/ESCOMP/mizuRoute.

Deleted: is available at https://doi.org/10.20383/102.0547.

Deleted: The LSE-based optimization codes and associated datasets will be shared on an open-access platform after manuscript publication.

Deleted: 33¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Acknowledgments

We acknowledge invaluable guidance, support and feedback from Chanel Mueller (US Army Corps of Engineers) and Chris Frans (US Bureau of Reclamation), who helped to motivate and steer this effort toward US water agency relevance and applications.

Financial support

This study is supported by the research grants to NCAR from the NASA Subseasonal-to-Seasonal Hydrometeorological Prediction Program (Award #80NSSC23K0502), the US Army Corps of Engineers (Agreement #HQUSACE17IIS), the US Bureau of Reclamation Research and Development Office (Award # R24AC00121), and the NOAA Climate Observations and Modeling Program (Award #NA23OAR4310448). We acknowledge high-performance computing support provided by NCAR's Computational and Information Systems Laboratory, sponsored by the US National Science Foundation.

Author contributions

MF and AW wrote and revised the manuscript, and AW, MF, and GT designed the study, and contributed experiments and analyses. GT developed the initial version of the calibration codes, workflows, and forcing dataset, with strategy and design guidance from AW (see Tang et al., 2025), and assisted MF in adapting these workflows and techniques to SUMMA. NM also helped to incorporate channel routing with mizuRoute. MF ran all simulations, analyzed and refined the approaches, and prepared all of the figures with initial assistance from GT. GT and NM provided final manuscript edits and feedback. AW leads the research projects at NCAR that are undertaking this body of work.

Competing interests

The authors declare that they have no conflict of interest.

References

Adams, B. M., Bohnhoff, W. J., Canfield, R. A., Dalbey, K. R., Ebeida, M. S., Eddy, J. P., Eldred, M. S., Geraci, G., Hooper, R. W., Hough, P. D., Hu, K. T., Jakeman, J. D., Carson, K., Khalil, M., Maupin, K. A., Monschke, J. A., Prudencio, E. E., Ridgway, E. M., Rushdi, A. A., Seidl, D. T., Stephens, J. A., Swiler, L. P., Tran, A., Vigil, D. M., von Winckel, G. J., Wildey, T. M., and Winokur, J. G. (with Menhorn, F., and Zeng, X.): Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis (Version 6.18 Developers Manual), Sandia National Laboratories, Albuquerque, NM, http://snl-dakota.github.io, 2023

Deleted: valuable
Deleted: USACE
Deleted: which
Deleted: application
Deleted: United States
Deleted: Climate Preparedness and Resilience Program (the 'Robust Hydrology' projects), and the
Deleted: .
Deleted: and
Deleted: contributing
Deleted: core
Deleted: 2024
Deleted: sponsoring
Formatted: Space After: 0 pt
Deleted: &
Deleted: &
Deleted: .). (2023).
Deleted: : A
Deleted:).
Deleted: . Available online:
Deleted:
Formatted: Font color: Auto
Deleted: .
Deleted: 34¶
Formatted: Font color: Black
Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sc., 21(9), 5293–5313, https://doi.org/10.5194/hess-21-5293-2017, 2017.

Arsenault, R., Martel, J.-L., Brunet, F., Brissette, F., and Mai, J.: Continuous streamflow prediction in ungauged basins: long short-term memory neural networks clearly outperform traditional hydrological models, Hydrol. Earth Syst. Sc., 27, 139–157, https://doi.org/10.5194/hess-27-139-2023, 2023.

Baker, E., Harper, A. B., Williamson, D., and Challenor, P.: Emulation of high-resolution land surface models using sparse Gaussian processes with application to JULES, Geosci. Model Dev., 15, 1913–1929, https://doi.org/10.5194/gmd-15-1913-2022, 2022.

Bennett, A., Tran, H., De la Fuente, L., Triplett, A., Ma, Y., Melchior, P., Maxwell, R. M., and Condon, L. E.: Spatio-temporal machine learning for regional to continental-scale terrestrial hydrology, J. Adv. Model. Earth Sy., 16(1), https://doi.org/10.1029/2023MS004095, 2024.

Broman, D. P., and Wood, A. W.: Better representation of low elevation snowpack to improve operational forecasts, Final Report No. ST-2019-178-01 to the Science and Technology Program, Research and Development Office, US Bureau of Reclamation, Denver, USA, https://www.usbr.gov/research/projects/download_product.cfm?id=3099, 2021.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept, Water Resour. Res., 51(4), 2498–2514, https://doi.org/10.1002/2015WR017198, 2015a.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Gochis, D. J., Rasmussen, R. M., Tarboton, D. G., Mahat, V., Flerchinger, G. N., and Marks, D. G.: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies, Water Resour. Res., 51(4), 2515–2542, https://doi.org/10.1002/2015WR017200, 2015b.

Clark, M. P., Zolfaghari, R., Green, K. R., Trim, S., Knoben, W. J. M., Bennett, A., Nijssen, B., Ireson, A., and Spiteri, R. J.: The numerical implementation of land models: problem formulation and laugh tests, J Hydrometeorol., 22(12), 3143–3161, https://doi.org/10.1175/JHM-D-20-0175.1, 2021.

Dagon, K., Sanderson, B. M., Fisher, R. A., and Lawrence, D. M.: A machine learning approach to emulation and biophysical parameter estimation with the Community Land Model, version 5, Adv. Stat. Clim. Meteorol. Oceanogr., 6(1), 223–244, https://doi.org/10.5194/ascmo-6-223-2020, 2020.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evolut. Computat., 6(2), 182–197, https://doi.org/10.1109/4235.996017, 2002.

Farahani A., M., Wood, A., Tang, G., Mizukami, N.: Calibrating a large-domain land/hydrology process model in the age of AI: the SUMMA CAMELS emulator experiments, Zenodo [data set], https://doi.org/10.5281/zenodo.16422768, 2025.

Farahani A., M., Wood, A., Tang, G., Mizukami, N.: Calibrating a large-domain land/hydrology process model in the age of AI: the SUMMA CAMELS emulator experiments, Github [code], https://github.com/NCAR/opt_landhydro, 2025.

Deleted: &…nd Clark, M. P. (2017)….: The CAMELS ... [115]
Deleted: *System Sciences,*
Deleted: …. https://doi.org/10.5194/hess-21-5293-2017 ... [118]
Deleted: .
Formatted ... [116]
Formatted ... [117]
Deleted: &…nd Mai, J. (2023)….: Continuous streamfl ... [119]
Deleted: *System Sciences,*
Deleted: (1),… 139–157. https://doi.org/10.5194/hess-27 ... [122]
Formatted ... [120]
Formatted ... [121]
Deleted: &…nd Challenor, P. (2022)….: Emulation of h ... [123]
Deleted: et al. (2024)….axwell, R. M., and Condon, L. E ... [124]
Deleted: *Systems,*
Formatted ... [125]
Formatted ... [126]
Deleted: e2023MS004095….HYPERLINK ... [127]
Formatted ... [128]
Deleted: .
Deleted: &…nd Wood, A. W. (2019)….: Better represen ... [129]
Deleted: .
Deleted: &…nd Rasmussen, R. M. (2015a)….: A unifie ... [130]
Formatted ... [131]
Deleted: *Resources Research,*
Formatted ... [132]
Deleted: …. https://doi.org/10.1002/2015WR017198 ... [133]
Deleted: .
Deleted: &…nd Marks, D. G. (2015b)….: A unified app ... [134]
Formatted ... [135]
Deleted: *Resources Research,*
Formatted ... [136]
Deleted: .
Formatted ... [137]
Deleted: .
Deleted: &…nd Spiteri, R. J. (2021)….: The numerical ... [138]
Deleted: .
Formatted ... [139]
Formatted ... [140]
Deleted: .
Deleted: &…nd Lawrence, D. M. (2020)….: A machine ... [141]
Formatted ... [142]
Deleted: …. https://doi.org/10.5194/ascmo-6-223-2020 ... [143]
Deleted: .
Deleted: &…nd Meyarivan, T. (2002)….: A fast and eli ... [144]
Formatted ... [145]
Deleted: *Transactions on Evolutionary Computation,*
Formatted ... [146]
Deleted: .
Formatted ... [147]
Deleted: .
Deleted: 35¶
Formatted ... [113]
Formatted ... [114]

Feng, D., Fang, K., and Shen, C.: Enhancing streamflow forecast and extracting insights using long-short term memory networks with data integration at continental scales, Water Resour. Res., 56(5), https://doi.org/10.1029/2019WR026793, 2020.

Feng, D., Liu, J., Lawson, K., and Shen, C.: Differentiable, learnable, regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy, Water Resour. Res., 58, https://doi.org/10.1029/2022WR032404, 2022.

Frame, J. M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., Qualls, L. M., Gupta, H. V., and Nearing, G. S.: Deep learning rainfall–runoff predictions of extreme events, Hydro. Earth Syst. Sc., 26(11), 3377–3392, https://doi.org/10.5194/hess-26-3377-2022, 2022.

Gharari, S., Clark, M. P., Mizukami, N., Knoben, W. J. M., Wong, J. S., and Pietroniro, A.: Flexible vector-based spatial configurations in land models, Hydro. Earth Sys. Sc., 24(12), 5953–5971, https://doi.org/10.5194/hess-24-5953-2020, 2020.

Gong, W., Duan, Q., Li, J., Wang, C., Di, Z., Ye, A., Miao, C., and Dai, Y.: Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical models, Water Resour. Res., 52(4), 1984–2008, https://doi.org/10.1002/2015WR018230, 2016.

Herrera, P.A., Marazuela, M.A., and Hofmann, T.: Parameter estimation and uncertainty analysis in hydrological modeling, WIREs Water, 9(1), e1569, https://doi.org/10.1002/wat2.1569, 2022.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of predictions in ungauged basins (PUB)—a review, Hydrolog. Sci. J., 58(6), 1198–1255, https://doi.org/10.1080/02626667.2013.803183, 2013.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., and Nearing, G.: Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets, Hydrol. Earth Syst. Sc., 23, 5089–5110, https://doi.org/10.5194/hess-23-5089-2019, 2019.

Kratzert, F., Gauch, M., Klotz, D., and Nearing, G.: HESS Opinions: Never train a Long Short-Term Memory (LSTM) network on a single basin, Hydrol. Earth Syst. Sc., 28, 4187–4201, https://doi.org/10.5194/hess-28-4187-2024, 2024.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M. P., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model version 5: description of new features, benchmarking, and impact of forcing uncertainty, J. Adv. Model. Earth Sys., 11(12), 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Deleted: &…nd Shen, C. (2020)….: Enhancing streamfl ... [151]
Formatted ... [150]
Deleted: *Resources Research,*
Deleted: e2019WR026793….HYPERLINK ... [154]
Deleted: .
Formatted ... [152]
Formatted ... [153]
Formatted ... [155]
Deleted: &…nd Shen, C. (2022)….: Differentiable, lear ... [156]
Deleted: *Resources Research,*…esour. Res., 58, e2022W ... [158]
Formatted ... [157]
Formatted ... [159]
Deleted: &…nd Nearing, G. S. (2022)….: Deep learning ... [160]
Deleted: *System Sciences,*
Deleted: .
Formatted ... [161]
Formatted ... [162]
Formatted ... [163]
Deleted: .
Deleted: &…nd Pietroniro, A. (2020)….: Flexible vecto ... [164]
Formatted ... [165]
Deleted: *System Sciences,*
Formatted ... [166]
Deleted: …. https://doi.org/10.5194/hess-24-5953-2020 ... [167]
Deleted: .
Deleted: &…nd Dai, Y. (2016)….: Multiobjective adapt ... [168]
Formatted ... [169]
Deleted: *Resources Research,*
Deleted: …. https://doi.org/10.1002/2015WR018230 ... [171]
Formatted ... [170]
Deleted: .
Formatted ... [172]
Deleted: &…nd Cudennec, C. (2013)….: A decade of ... [173]
Formatted ... [174]
Deleted: .
Formatted ... [175]
Deleted: .
Formatted ... [176]
Deleted: &
Deleted: .S. (2019). Benchmarking a Catchment-Aware ... [178]
Formatted ... [177]
Formatted ... [179]
Deleted: -368
Deleted: &…nd Nearing, G. (2024)….: HESS Opinions ... [180]
Formatted ... [181]
Deleted: *System Sciences,*
Formatted ... [182]
Deleted: (10),… 4187–4201. https://doi.org/10.5194/hes ... [183]
Deleted: &…nd Zeng, X. (2019)….: The Community La ... [184]
Deleted: *Systems,*
Deleted: …. https://doi.org/10.1029/2018MS001583 ... [187]
Formatted ... [185]
Formatted ... [186]
Deleted: .

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global reconstruction of naturalized river flows at 2.94 million reaches, Water Resour. Res., 55, 6499–6516, https://doi.org/10.1029/2019WR025287, 2019.

Mai, J., Shen, H., Tolson, B. A., Gaborit, É., Arsenault, R., Craig, J. R., Fortin, V., Fry, L. M., Gauch, M., Klotz, D., Kratzert, F., O'Brien, N., Princz, D. G., Rasiya Koya, S., Roy, T., Seglenieks, F., Shrestha, N. K., Temgoua, A. G. T., Vionnet, V., and Waddell, J. W.: The Great Lakes runoff intercomparison project phase 4: the Great Lakes (GRIP-GL), Hydrol. Earth Sys. Sc., 26, 3537–3572, https://doi.org/10.5194/hess-26-3537-2022, 2022.

Mitchell, M.: An introduction to genetic algorithms, The MIT Press, https://doi.org/10.7551/mitpress/3927.001.0001, 1996.

Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R. J., Markstrom, S. L., Hay, L. E., Woods, R., Arnold, J. R., and Brekke, L. D.: MizuRoute version 1: A river network routing tool for continental domain water resources applications, Geosci. Model Devel., 9(7), 2223–2238, https://doi.org/10.5194/gmd-9-2223-2016, 2016.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resour. Res., 53(9), 8020–8040, https://doi.org/10.1002/2017WR020401, 2017.

Mizukami, N., Clark, M. P., Gharari, S., Kluzek, E., Pan, M., Lin, P., Beck, H. E., and Yamazaki, D.: A vector-based river routing model for Earth System Models: Parallelization and global applications, J. Adv. Model. Earth Sys., 13, https://doi.org/10.1029/2020MS002434, 2021.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13(9), 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., Hassidim, A., Klotz, D., Kratzert, F., and Metzger, A.: Global prediction of extreme floods in ungauged watersheds, Nature, 627(5), 559–563, https://doi.org/10.1038/s41586-024-07145-1, 2024.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sc., 19(1), 209–223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Newman, A. J., Mizukami, N., Clark, M. P., Wood, A. W., Nijssen, B., and Nearing, G.: Benchmarking of a physically based hydrologic model, J. Hydrometeor., 18, 2215–2225, https://doi.org/10.1175/JHM-D-16-0284.1, 2017.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46, W05523, https://doi.org/10.1029/2008WR007327, 2010.

Formatted ... [190]
Deleted: &…nd Waddell, J. W. (2022)….: The Great La ... [191]
Deleted: *System Sciences,*
Formatted ... [192]
Formatted ... [193]
Deleted: (12),… 3537–3572. https://doi.org/10.5194/hes ... [194]
Deleted: . (1998)….: An introduction to genetic algorith ... [195]
Deleted: &…nd Brekke, L. D. (2016). mizuRoute…: Mi ... [196]
Deleted: *Development,*
Deleted: …. https://doi.org/10.5194/gmd-9-2223-2016 ... [199]
Deleted: .
Formatted ... [197]
Formatted ... [198]
Deleted: &…nd Samaniego, L. (2017)….: Towards sean ... [200]
Deleted: *Resources Research,*
Deleted: .
Formatted ... [201]
Formatted ... [202]
Formatted ... [203]
Deleted: .
Deleted: et al. (
Moved up [1]: 2021).
Deleted: . Journal of Advances in Modeling… J. Adv. M ... [204]
Formatted ... [205]
Deleted: &…nd Thépaut, J.-N. (2021)….: ERA5-Land: ... [206]
Formatted ... [207]
Deleted: *System Science*
Formatted ... [208]
Deleted: …. https://doi.org/10.5194/essd-13-4349-2021 ... [209]
Deleted: .
Deleted: &…nd Metzger, A. (2024)….: Global predictic ... [210]
Formatted ... [211]
Deleted: &…nd Duan, Q. (2015)….: Development of a l ... [212]
Formatted ... [213]
Deleted: *System Sciences,*
Formatted ... [214]
Deleted: …. https://doi.org/10.5194/hess-19-209-2015 ... [215]
Deleted: .
Deleted: N. …izukami, M.P.…., Clark, AW…. P., Woo ... [216]
Deleted: &…nd Attinger, S. (2010)….: Multiscale paran ... [217]
Deleted: *Resources Research,*
Deleted: …. https://doi.org/10.1029/2008WR007327 ... [220]
Formatted ... [218]
Formatted ... [219]
Deleted: .
Deleted: 37¶
Formatted ... [188]
Formatted ... [189]

Shen, C., Appling, A. P., Gentine, P., Gentine, P., Bandai, T., Gupta, H., Alexandre, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L., Liu, X., Ren, W., Zheng, Y., Harman, C. J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck, H., E., Bindas, T., Dwivedi, D., Fang, K., Höge, M., Rackauckas, C., Mohanty, B., Roy, T., Xu, C., and Lawson, K.: Differentiable modelling to unify machine learning and physical models for geosciences, Nat Rev Earth Environ, 4, 552–567, https://doi.org/10.1038/s43017-023-00450-9, 2023.

Song, Y., Bindas, T., Shen, C., Ji, H., Knoben, W. J. M., Lonzarich, L., Lonzarich, L., Clark, M. P., Liu, J., van Werkhoven, K., Lamont, S., Denno, M., Pan, M., Yang, Y., Rapp, J., Kumar, M., Rahmani, F., Thébault, C., Adkins, R., Halgren, J., Patel, T., Patel, A., Sawadekar, K. A., and Lawson, K.: High-resolution national-scale water modeling is enhanced by multiscale differentiable physics-informed machine learning, Water Resour. Res., 61, https://doi.org/10.1029/2024WR038928, 2025.

Tang, G., Clark, M. P., and Papalexiou, S. M.: EM-Earth: The ensemble meteorological dataset for planet Earth, Bull. .Amer. Meteorol. Soc., 103(5), E996–E1018, https://doi.org/10.1175/BAMS-D-21-0106.1, 2022.

Tang, G., Wood, A. W., and Swenson, S.: On using AI-based large-sample emulators for land/hydrology model calibration and regionalization, Water Resour. Res., 61, https://doi.org/10.1029/2024WR039525, 2025

Tsai, WP., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., Liu, J., and Shen, C.: From calibration to parameter learning: harnessing the scaling effects of big data in geoscientific modeling, Nat. Commun, 12, 5988, https://doi.org/10.1038/s41467-021-26107-z, 2021.

Tolson, B. A., and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resour. Res., 43, W01413, https://doi.org/10.1029/2005WR004723, 2007.

Van Beusekom, A. E., Hay, L. E., Bennett, A. R., Choi, Y.-D., Clark, M. P., Goodall, J. L., Li, Z., Maghami, I., Nijssen, B., and Wood, A. W.: Hydrologic model sensitivity to temporal aggregation of meteorological forcing data: a case study for the contiguous United States, J. Hydrometeorol., 23(1), 167–183, https://doi.org/10.1175/JHM-D-21-0111.1, 2022.

Vano, J. A., Nijssen, B., and Lettenmaier, D. P.: Seasonal hydrologic responses to climate change in the Pacific Northwest, Water Resour. Res., 51, 1959–1976, https://doi.org/10.1002/2014WR015909, 2015.

Wagener, T., and Wheater, H. S.: Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, J. hydrol., 320, 132–154, https://doi.org/10.1016/j.jhydrol.2005.07.015, 2006.

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment classification and hydrologic similarity, Geography Compass, 1: 901-931, https://doi.org/10.1111/j.1749-8198.2007.00039.x, 2007.

Watson-Parris, D., Williams, A., Deaconu, L., and Stier, P.: Model calibration using ESEm v1.1.0 – an open, scalable Earth system emulator, Geosci. Model Dev., 14(12), 7659–7672, https://doi.org/10.5194/gmd-14-7659-2021, 2021.

Wood, A. W., Sturtevant, J., Barrett, J. L., and Llwewellyn, D.: Improving the robustness of southwestern US water supply forecasting, Final Report NO. ST-2018-8117-01 to the Science and Technology Program, Research and Development Office, US Bureau of Reclamation, Denver, USA, https://www.usbr.gov/research/projects/download_product.cfm?id=3029, 2021.

Deleted: . et al.…, Gentine, P., Bandai, T., Gupta, H., Alexandre, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li L., Liu, X., Ren, W., Zheng, Y., Harman, C. J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck ... [222]
Deleted:). https://doi.org/10.1038/s43017-023-00450-9
Deleted: , Tadd Bindas1, Chaopeng Shen2, Haoyu Ji2, W ... [223]
Deleted: &…nd Papalexiou, S. M. (2022).…: EM-Earth ... [224]
Deleted: …. https://doi.org/10.1175/BAMS-D-21-0106. ... [225]
Deleted: .
Formatted: Font: Not Italic
Deleted: &…nd Swenson, S. (2024).…: On using AI-ba ... [226]
Deleted: W.-P…P., Feng, D., Pan, M., Beck, H., Lawson ... [227]
Deleted: . https://doi.org/10.1038/s41467-021-26338-0
Formatted: Font: Not Italic
Deleted: &…nd Shoemaker, C. A. (2007).…: Dynamica ... [228]
Deleted: *Resources Research,*
Deleted: …. https://doi.org/10.1029/2005WR004723 ... [229]
Deleted: .
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Deleted: &…nd Wood, A. W. (2022).…: Hydrologic mo ... [230]
Formatted: Font: Not Italic
Deleted: …. https://doi.org/10.1175/JHM-D-21-0121 ... [231]
Deleted: B. …ijssen, B., and D. P. …ettenmaier (2015), ... [232]
Formatted: Space After: 0 pt
Deleted: …, and Woods, R. (2007),…: Catchment Class ... [233]
Formatted: Font color: Auto
Deleted: &…nd Stier, P. (2021).…: Model calibration u ... [234]
Formatted: Font: Not Italic
Deleted: *Development,*
Formatted: Font: Not Italic
Deleted: …. ... [235]
Deleted: .
Formatted: Font color: Auto
Deleted: &…nd Llwewellyn, D. (2021).…: Improving th ... [236]
Formatted: Font color: Auto
Deleted: .
Deleted: 38¶
Formatted: Font color: Black
Formatted ... [221]

Wood, A. W., and Mizukami, N.: SUMMA CAMELS dataset, HydroShare, http://www.hydroshare.org/resource/0513cf5e792a4dc4acd0ca77a8146036, 2022.

Deleted: &

Deleted: . (2022).

Deleted: .

Deleted: . Accessed 1 January 2021. Available online:

Formatted: Font: Not Italic

Deleted:

Deleted: .¶
Yamazaki, D., Ikeshima, D., Sosa, J., Emanuel, J., Lin, Y., Tawatari, R., Sayama, T., & Kanae, S. (2019). MERIT Hydro: A high-resolution global hydrography map based on the MERIT DEM. *Water Resources Research, 55*(6), 5053–5073. https://doi.org/10.1029/2019WR024873.

Formatted: Font color: Auto

Deleted: 39¶

Formatted: Font color: Black

Formatted: Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 1: [1] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [2] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 1: [3] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font color: Black

Page 1: [4] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 7: [5] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 7: [5] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 7: [5] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 7: [6] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 7: [7] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Page 7: [8] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font color: Black

Page 7: [8] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font color: Black

Page 7: [9] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [10] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [11] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [12] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [13] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [14] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [15] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [16] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [17] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [18] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [19] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [20] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [21] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [22] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [23] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [24] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [25] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [26] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [27] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [28] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [29] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [30] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [31] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [32] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [33] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [34] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [35] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [36] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [37] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [38] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [39] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [40] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [41] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [42] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [43] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [44] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [45] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [46] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [47] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [48] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 7: [49] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Centered, Space After: 0 pt

Page 7: [50] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Left, Space After: 0 pt

Page 14: [51] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 14: [51] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 14: [52] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 14: [52] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 14: [52] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 14: [53] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 14: [53] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 14: [53] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 14: [53] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 1: [54] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 16: [55] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 16: [55] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 16: [56] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 16: [56] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 16: [56] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 16: [57] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 16: [57] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 1: [58] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 18: [59] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Space After: 10 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Page 18: [60] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 10 pt, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Page 1: [61] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 1: [62] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Black

Page 1: [63] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 27: [64] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [65] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [66] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [67] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [68] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [69] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [70] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [71] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [72] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [73] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [74] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [75] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [76] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [77] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [78] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 27: [79] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [80] Formatted Table	Mozhgan Farahani	7/25/25 2:32:00 PM

Formatted Table

Page 27: [81] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [82] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [83] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [84] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [85] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 27: [86] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 1: [87] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 1: [88] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 28: [89] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [90] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [91] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 28: [92] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border)

Page 28: [93] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 28: [93] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 28: [93] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 28: [94] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [95] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [96] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [97] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [98] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [99] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [100] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [101] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [102] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [103] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [104] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [105] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [106] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [107] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [108] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 28: [109] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 31: [110] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 31: [110] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Bold, Font color: Black

Page 31: [111] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 31: [111] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 31: [112] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Italic, Font color: Black

Page 31: [112] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Italic, Font color: Black

Page 1: [113] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Black

Page 1: [114] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 35: [115] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [115] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [115] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [115] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [116] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [117] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [118] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [118] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [119] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [119] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [119] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [119] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [120] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font: Not Italic

Page 35: [121] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font: Not Italic

Page 35: [122] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [122] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [123] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [123] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [123] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [124] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [124] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 35: [125] Formatted Mozhgan Farahani 7/25/25 2:32:00 PM

Font: Not Italic

Page 35: [126] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 35: [127] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [127] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [128] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Auto

Page 35: [129] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [129] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [129] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [129] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [129] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [130] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [130] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [130] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 35: [131] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 35: [132] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [133] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [133] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [134] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [134] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [134] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [135] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [136] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [137] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 35: [138] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [138] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [138] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [138] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [139] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [140] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 35: [141] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [141] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [141] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [142] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [143] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [143] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [144] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [144] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [144] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 35: [145] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [146] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 35: [147] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Auto

Page 1: [148] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Black

Page 1: [149] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 36: [150] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Space After: 0 pt

Page 36: [151] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [151] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [151] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [152] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [153] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [154] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [154] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [155] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font color: Auto

Page 36: [156] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [156] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [156] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [157] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [158] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [158] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [159] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 36: [160] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [160] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [160] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [161] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [162] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [163] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 36: [164] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [164] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [164] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [165] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [166] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [167] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [167] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [168] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [168] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [168] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [169] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [170] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 36: [171] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 36: [171] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [172] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [173] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [174] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [175] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 36: [176] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Not Highlight

Page 36: [177] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Not Highlight

Page 36: [178] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [179] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Not Highlight

Page 36: [179] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Not Highlight

Page 36: [180] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [180] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [180] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [180] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [181] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [182] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [183] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [183] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [184] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [184] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [184] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [184] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [184] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [185] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [186] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 36: [187] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 36: [187] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 1: [188] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Black

Page 1: [189] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 37: [190] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Space After: 0 pt

Page 37: [191] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [191] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [191] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [191] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [191] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [192] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [193] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [194] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [194] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [195] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [195] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [196] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [196] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [196] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [197] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [198] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [199] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [199] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [200] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [200] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [200] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [201] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [202] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [203] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 37: [204] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [204] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [204] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [205] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font color: Auto

Page 37: [206] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [206] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [206] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [206] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [207] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [208] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [209] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [209] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [210] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [210] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [210] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [211] Formatted	Mozhgan Farahani	7/25/25 2:32:00 PM

Font: Not Italic

Page 37: [212] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [212] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 37: [212] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [212] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [213] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [214] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [215] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [215] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [216] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [217] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [217] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [217] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [218] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [219] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Font: Not Italic

Page 37: [220] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 37: [220] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 1: [221] Formatted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Normal, Border: Top: (No border), Bottom: (No border), Left: (No border), Right: (No border), Between : (No border), Tab stops: 3.13", Centered + 6.27", Right

Page 38: [222] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 38: [222] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 38: [222] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 38: [223] Deleted **Mozhgan Farahani** **7/25/25 2:32:00 PM**

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [223] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [224] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [224] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [224] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [224] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [224] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [225] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [225] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [226] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [226] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [226] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [227] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [227] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [227] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [227] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [227] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [228] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [228] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [228] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [229] Deleted	**Mozhgan Farahani**	**7/25/25 2:32:00 PM**

Page 38: [229] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [230] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [230] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [230] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [230] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [231] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [231] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [232] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [232] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [232] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [232] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [233] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [233] Deleted Mozhgan Farahani 7/25/25 2:32:00 PM

Page 38: [233] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [233] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [234] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [234] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [234] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [234] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [235] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [235] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [236] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [236] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [236] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [236] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM

Page 38: [236] Deleted	Mozhgan Farahani	7/25/25 2:32:00 PM