

Earth's Future



RESEARCH ARTICLE

10.1029/2021EF002403

Special Section:

Modeling Multisector Dynamics
to Inform Adaptive Pathways

The Effects of Model Complexity on Model Output Uncertainty in Co-Evolved Coupled Natural-Human Systems

Chung-Yi Lin¹  and Yi-Chen Ethan Yang¹ 

¹Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA, USA

Key Points:

- Exploratory model output uncertainty is likely to increase with model complexity due to more uncertain model configurations
- Learning mechanisms in a human system (e.g., diversion) can offset the impact of the natural system (e.g., streamflow) on uncertainty
- The uncertainty caused by inappropriate model structural assumptions must be further examined

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. C. E. Yang,
yey217@lehigh.edu

Citation:

Lin, C. Y., & Yang, Y. C. E. (2022). The effects of model complexity on model output uncertainty in co-evolved coupled natural-human systems. *Earth's Future*, 10, e2021EF002403. <https://doi.org/10.1029/2021EF002403>

Received 3 SEP 2021

Accepted 22 MAY 2022

Author Contributions:

Conceptualization: Chung-Yi Lin

Data curation: Chung-Yi Lin

Formal analysis: Chung-Yi Lin

Methodology: Chung-Yi Lin

Software: Chung-Yi Lin

Supervision: Yi-Chen Ethan Yang

Visualization: Chung-Yi Lin

Writing – original draft: Chung-Yi Lin

Writing – review & editing: Yi-Chen Ethan Yang

Abstract Studies have recently focused on using coupled natural-human systems (CNHS) to inform policymaking. However, model uncertainty can increase with model complexity and affect the variance of the model outcomes. Therefore, this study explores an uncertainty analysis of coupled hydrological and human decision models to better evaluate CNHS modeling properties. Five coupled models are proposed with different model complexities for human behavior settings (i.e., model structure and the number of calibrated parameters): one static, two adaptive, and two learning adaptive. Learning adaptive models (the most complex) have both a learning component (capturing long-term trends) and an adaptive component (capturing short-term variations), while adaptive models omit the learning component. The static model is the simplest, without learning or adaptive components. Applying the law of total variance, the model output uncertainty is decomposed into three sources: (a) climate change scenario uncertainty, (b) climate internal variability, and (c) different model configurations with parameter sets or model structures that are equally capable of producing similar outcomes. Our exploratory analysis demonstrated that model uncertainty would likely increase with model complexity given uncertain input data (e.g., climate forcing) and different model configurations; the inclusion of a learning mechanism in the human system can potentially offset the impact of the natural system on uncertainty through coupling natural and human systems. We also discuss other uncertainty sources, such as assumptions about model structure due to incomplete knowledge and metrics for calibration target selection for future studies.

Plain Language Summary Recent studies have investigated connections between natural systems and related human impacts, such as those between the rainfall-runoff process and agricultural activities. This is primarily performed by using two-way connected, natural-human computer models. However, these coupled models require many model parameters, which results in larger prediction uncertainty, especially when the observed data is limited. This study quantifies the relationship between model output uncertainty and model complexity. We found that, depending on the type of model used, the uncertainty of model outputs will likely increase with model complexity; however, active two-way feedback between natural (e.g., streamflow) and human (e.g., water diversion) systems could offset the impact of the natural system on output uncertainty by increasing human variations. Note that such results were heavily influenced by corresponding model assumptions for human behaviors (e.g., how farmers will react to environmental feedback), which could be uncertain given incomplete knowledge. Therefore, continuously monitoring (i.e., acquiring data) and communicating uncertainty among modelers, policymakers, and other stakeholders is a necessary and worthy next step.

1. Introduction

In the anthropogenic era, most major basins experience human activity, requiring an understanding of interactions between human society and the natural environment, that is, co-evolution in coupled natural-human systems (CNHS; Bauch et al., 2016; Wada et al., 2017). With the recognition of the socio-physical nature of emerging water challenges, many methods and modeling techniques, including system dynamics modeling, agent-based modeling, Bayesian networks, and so on (Blair & Buytaert, 2016; Kelly et al., 2013), have been proposed to address complex water management issues, often in the style of the Harvard Water Program (Brown et al., 2015; Milly et al., 2008; Reuss, 2003). Among them, coupling an agent-based model (ABM) with a process-based model (e.g., hydrological model) is a promising method to investigate emerging phenomena and heterogeneous human behaviors in CNHS (Berglund, 2015; Giuliani et al., 2016; Yang et al., 2020). In this coupling framework, each agent (e.g., irrigation district or reservoir) serves as a decision-making unit with a set of rules and attributes governing its behavior (e.g., diversion request or release) and interacts with other agents in a shared physical environment (e.g., river basin). As a result, various ABMs have been developed with agents that have different

© 2022 The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](https://creativecommons.org/licenses/by-nc/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

complexities (e.g., model structure and number of calibrated parameters) to capture nonlinearity and non-stationarity in CNHS. For example, agent behaviors have been governed by a deterministic decentralized optimization model (Yang et al., 2009) or more sophisticated designs that consider adaptive (short-term) and learning (long-term) behaviors with respect to other agent decisions and the changing environment (Giuliani et al., 2016; Hyun et al., 2019).

Many case studies (Elshafei et al., 2014; Kandasamy et al., 2014; Marston & Konar, 2017; Song & Zhang, 2015; Xu et al., 2014; Yang et al., 2020) have suggested that an additional human complexity layer in the model will improve environmental planning and policy (Zellner, 2008). Yang et al. (2020) showed that the potential tipping point of farmer behavioral changes could be identified by the bottom-up nature of the coupled model. Hung and Yang (2021) argued that the learning adaptive agent design of a coupled model could assist water managers in developing soft policies. As the complexity of a method and the number of parameters increase, the challenge of handling model uncertainty and variability increases (McLean & McAuley, 2012; Srikrishnan & Keller, 2021). This challenge is aggravated in CHNS modeling since there is often interest in both natural (e.g., streamflow) and human (e.g., water diversions) system outputs, and the co-evolved natural and human dynamics in the complex system heavily rely on structural assumptions of the model (e.g., how humans react to environmental changes; Karthe et al., 2021; Messina et al., 2008). These model parameters and structural uncertainties significantly affect confidence in the inference and interpretation of model results (Allen & Gunderson, 2011; Allison et al., 2018; Kelly et al., 2013; Sun et al., 2016).

In the modeling literature, uncertainty is often ascribed to input data, model resolution level, model structure, and parameters (Saltelli et al., 2019); in addition, problem framing, perceptions of uncertainty, inappropriate assumptions, and epistemic uncertainty are also deeply embedded in model uncertainty (Beven, 2016; Di Baldassarre et al., 2016; Moallemi et al., 2020; Westerberg et al., 2017). However, uncertainty studies of coupled models in CNHS are still emerging (Elsawah et al., 2020). In this paper, we focused on parameter and structural uncertainties, that is, different model configurations that can produce similar/acceptable model outcomes, in CNHS modeling. A more complex model may suffer more severe uncertainty issues from a potentially wide range of behavioral model configurations (Arendt et al., 2012; Srikrishnan & Keller, 2021). In the hydrology field, this is also known as equifinality (Beven, 1993), and many studies have demonstrated the impact of equifinality in their analysis (Arsenault & Brissette, 2014; Beven, 2006; Ekblad & Herman, 2021; Khatami et al., 2019; Williams et al., 2020).

Generalized likelihood uncertainty estimation (GLUE; Beven & Binley, 1992) and Differential Evolution Adaptive Metropolis (DREAM; Vrugt et al., 2009) are two popular Bayesian methods to address different model configurations issues. GLUE weights model outputs based on data-fitting performance and then uses behavioral sampled models for ensemble forecasting. DREAM, on the other hand, is a more advanced formal Bayesian methods sampling model from the posterior. Instead of proposing new methods, this study compares model output uncertainties (both natural and human systems) under different model configurations (e.g., number of parameters and model structures). We quantify the model uncertainty brought by different model configurations using a variance-based uncertainty decomposition to evaluate uncertainty properties under different model structural assumptions in CNHS. We introduce the term “equifinal model representatives” (EMRs) to describe the selected model configurations that represent the range of model variability over the equifinal space (i.e., space of model configurations that produce similar model outcomes), and we refer to model output uncertainty as simply “model uncertainty” throughout the rest of the paper for simplicity.

This study aims to capture natural and human system model uncertainty given the available data with two hypotheses. First, model uncertainty in an exploratory analysis would likely increase with model complexity, given uncertain input data (e.g., climate forcing) and different model configurations. Second, the inclusion of a learning mechanism in the human system model can potentially offset the impact of the natural system's variability on model uncertainty. The Yakima River Basin (YRB) in the Northwest United States, an agriculture-dominated basin that heavily relies on irrigation, was selected as the study area to test these hypotheses. We presented four tasks: (a) quantify model uncertainty of CNHS in exploratory simulations, (b) decompose model uncertainty into different model configurations, and input data uncertainty (e.g., climate scenario uncertainty and internal climate variability) using the law of total variance, (c) compare model uncertainty and model complexity, and (d) observe the uncertainty offsetting relationship between natural and human systems in CNHS modeling.

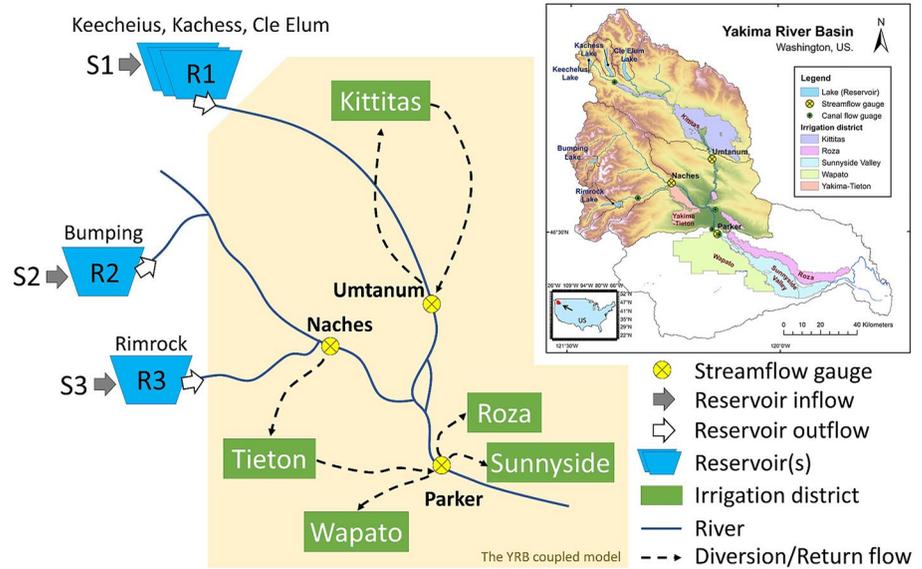


Figure 1. System diagram of Yakima River Basin (YRB). S1, S2, and S3 and R1, R2, and R3 are reservoir inflow and release models, respectively. YRB coupled natural-human model (yellow-colored area), which consists of three subbasins (yellow circles are their outlets) and five irrigation diversion districts (green boxes), is the focus of uncertainty analysis.

The article is structured as follows. Section 2 introduces the YRB study area and the climate change scenario design. The identification of EMRs, the uncertainty decomposition method, and the modeling schema are presented in Section 3. Section 4 compares the decomposition from coupled models with different model complexities. Discussion of additional uncertainty sources in CNHS modeling and limitations of the models is presented in Section 5, followed by the conclusions in Section 6.

2. Study Area and Materials

2.1. Yakima River Basin

The YRB is selected as the study area and is shown in Figure 1. The YRB is a basin dominated by agriculture. Orchards (127,934 acres, 29.6%), small grains (67,434 acres, 15.6%), and corn (63,163 acres, 14.6%) (USDA, 2017 agriculture census) are the primary crops in the area. The precipitation in the YRB is concentrated in the mountains during the winter season (November–February). Agriculture heavily relies on irrigation; however, irrigation demands vary with the uncertain climate and the vicissitude of crop types across the growing season (late March to early October). Irrigation is facilitated by five major reservoirs (Keechelus, Kachess, Cle Elum, Bumping, and Rimrock), which are jointly operated by the U.S. Bureau of Reclamation (USBR). The irrigation operation (i.e., storage control) usually starts in July, and it continues through September to fulfill downstream demands (USBR, 2002). During this major storage control period, the streamflow at Parker gauge is tightly maintained by USBR to meet flow targets that range from 300 to 600 ft³/s (8.5–16.99 m³/s), based on available water (USBR, 2002). The streamflow in this period is a policy concern for USBR. Therefore, we adopt the median value of mean monthly streamflow from July to September at the Parker gauge over a given period (Q_M ; Equation 1) as the natural system indicator for the model uncertainty analysis:

$$Q_M = \text{median} \left(\left\{ \frac{1}{3} \sum_{m=7}^9 Q_{y,m}; y = 1, \dots, p \right\} \right) \quad (1)$$

where $Q_{y,m}$ is the monthly streamflow of month m in year y and p is the number of years. Note that the flow target standard is not static. The flow target standard of the Parker gauge is affected by district conservation programs. Based on historical data, we set the flow targets equal to 9.4 m³/s (1960–1984) and 14.4 m³/s (1985–2013) for the model calibration, where we use the flow deviation from the target as a driving force for the learning process of the water diversion model (Section 3.2).

In addition, we define the median value of annual diversion of five districts over a given period (D_M ; Equation 2) as the human system indicator that abstracts farmers' behaviors (e.g., crop selection and irrigation measures) for the model uncertainty analysis:

$$D_M = \text{median} \left(\left\{ \sum_{ag \in \text{districts}} D_{y,ag}; y = 1, \dots, p \right\} \right) \quad (2)$$

where $D_{y,ag}$ is the annual diversion of agent ag in year y . We also define the median value of shortage frequency (S_M ; Equation 3) as an auxiliary indicator that represents the potential crop losses that farmers tend to avoid:

$$S_M = \text{median} \left(\left\{ \sum_{r=1}^{20} S_{y,r}; y = 1, \dots, p \right\} \right) \quad (3)$$

where $S_{y,r}$ is a shortage index experienced in internal climate variability realization r . $S_{y,r}$ is equal to 1 if the total diversion request of five agents is not fulfilled in that year; otherwise, $S_{y,r}$ is equal to 0. Higher S_M implies a lower variation in Q_M , since the streamflow is likely to approximate minimum flow requirements.

The YRB coupled natural-human model (Figure 1) includes hydrological and water diversion models. The hydrological model contains three subbasins, Umtanum (328,818.70 ha), Naches (203,799.79 ha), and Parker (291,203.80 ha), and the diversion model includes five irrigation districts: Kittitas, Yakima-Tieton (Tieton), Roza, Wapato, and Sunnyside Valley (Sunnyside). Their diversion points are all above Parker gauge. The details of these five districts, such as water rights, average annual diversions, and district areas, are summarized in Table S1 in Supporting Information S1. We define five major irrigation districts in the basin as agents that make annual diversion requests. The inclusion of these five irrigation districts in the water diversion model abstracts possible causes of changes in diversions (land use, conservation programs, climate, and farmers' subjective decisions) into the empirical equations presented in Section 3.2.

Reservoirs are another human component in the YRB. The reservoir model uses the simulated inflows of S1, S2, and S3 (whose drainage areas are equal to 83,014.25, 11,601.47, and 28,016.20 ha, respectively); storage status; and pre-defined fixed operational rules to simulate the reservoir releases. These releases are then used as input data for the YRB coupled model for simplicity. Namely, the reservoir releases constrain the downstream water users' choices. For readers interested in finding optimal reservoir operating rules or interactions between reservoirs and diversion agents, we refer them to Giuliani et al. (2016) and Madani and Hooshyar (2014), respectively.

In sum, the Q_M and D_M uncertainties under an exploratory analysis, accumulate the input data uncertainty (e.g., reservoir releases and climate) and the errors of simplification (i.e., the conceptualization of physical mechanisms) in both the natural process (e.g., hydrological model) and human behaviors (e.g., water diversion model).

2.2. Data and Climate Change Scenarios

We use data from 1960 to 2013 as the diagnostic period to develop the model. The data are separated into calibration (1960–1999) and validation (2000–2013) periods. The weather data (i.e., precipitation and temperature) are collected from Livneh et al. (2015), and the reservoir storage and release, streamflow, and diversion data are downloaded from the USBR website. The precipitation data are bias-corrected through a simple annual water balance method to solve the temporal inconsistency issue mentioned in the limitation section in Livneh et al. (2015).

The weather inputs for the exploratory analysis (2021–2100) are generated by adjusting annually bootstrapped historical time series to different climate change scenarios (Figure 2). The climate change scenarios are defined as combinations of linearly interpolated precipitation ratios (PF_{ratio} ; average annual precipitation in future periods to the baseline; lines in Figure 2a) and temperature deltas (TF_{delta} ; average annual temperature in future periods minus the baseline; lines in Figure 2b) that are selected from five quantile values (q0.1, q0.3, q0.5, q0.7, and q0.9) over the range of 20 general circulation models (GCMs; Taylor et al., 2012; boxplots in Figure 2; Table S2 in Supporting Information S1) in the 2030s, 2050s, 2070s, and 2090s. Such a climate change scenario setting is similar to the climate stress test in Decision Scaling (Brown et al., 2012); however, the climate changes we tested are limited within the range of GCMs.

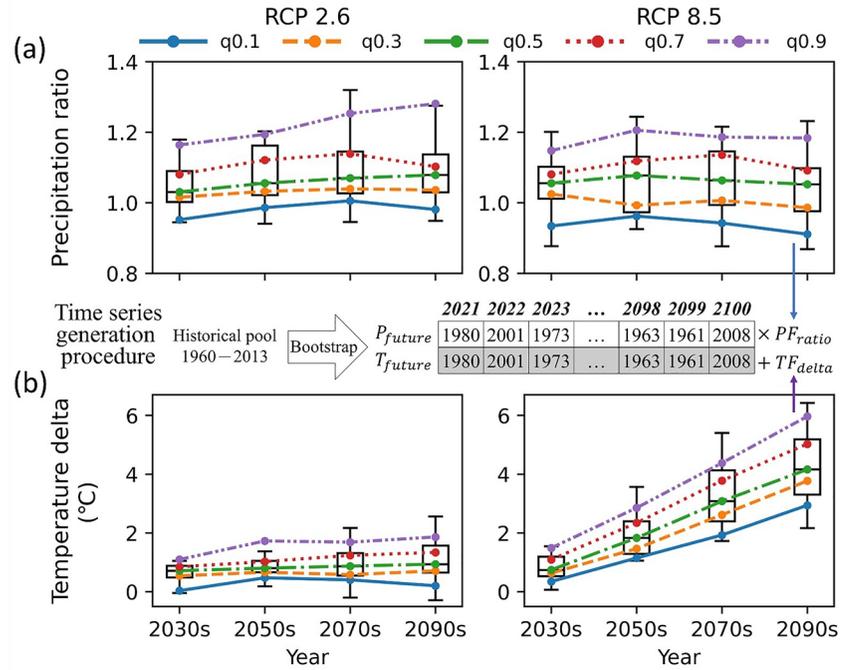


Figure 2. (a) Precipitation ratios and (b) temperature deltas sampled from five quantile values over the range of 20 GCMs (boxplots) under RCP 2.6 (left column) and RCP 8.5 (right column), respectively, in the 2030s, 2050s, 2070s, and 2090s. Future weather time series were generated by adjusting annually bootstrapped time series with climate change scenarios, combinations of linearly interpolated ratios, and deltas.

This study adopts two representative concentration pathways (RCPs), RCP 2.6 and RCP 8.5, which represent mild and severe future climate projections, respectively. By selecting two distinct climate change scenarios, we could potentially maximize the ability to separate climate change signals from noises. As a result, we generate 25 climate change combinations ($5 PF_{ratio} \times 5 TF_{delta}$) per RCP. For exploratory simulations, the climate change scenarios are statistically downscaled by the delta method (Walsh et al., 2018), to which we apply PF_{ratio} and TF_{delta} to 20 annually bootstrapped time series sampled from 1960 to 2013 in historical annual slices (daily precipitation and temperature time series of the entire year). This enables us to incorporate internal climate variability and output future weather time series (P_{future} and T_{future} ; Equation 4):

$$\begin{cases} P_{future,y} = P_{bootstrap,y} \times PF_{ratio,y} \\ T_{future,y} = T_{bootstrap,y} + TF_{delta,y} \end{cases}, y \in [2021, 2100], y \in \mathbb{Z} \quad (4)$$

where P and T are annually bootstrapped daily precipitation and temperature time series, respectively. Subscript y denotes the year. In sum, we generate 1,000 realizations (25 climate change combinations \times 2 RCPs \times 20 bootstrapped time series) for the exploratory analysis. Note that we only consider a subset of climate input data uncertainty in this study; other uncertainties, like single GCM outputs and different downscaling methods, are not included.

3. Methods

3.1. Hydrological Model

The semi-distributed hydrological model developed for the YRB is made of the modified hydrological module of the Generalized Watershed Loading Functions (GWLF; Haith & Shoemaker, 1987; Tung & Haith, 1995) and the Lohmann routing model (Lohmann et al., 1998; Wi et al., 2015). The hydrological module of the GWLF is a lumped rainfall-runoff model, which simulates daily streamflow by summing the surface quick flow, which is computed by the NRCS curve number method, and subsurface flow (Haith et al., 1996). The modified version

Table 1
Calibration Parameter Bounds for Hydrological and ABM Models

Model	Sub-model	Parameter name	Unit	Code	Bound
Hydrological model	GWLF	Curve number	–	<i>CN2</i>	[25, 100]
		Interception coefficient	–	<i>IS</i>	[0, 0.5]
		Recession coefficient	–	<i>Res</i>	[10 ⁻³ , 0.5]
		Deep seepage coefficient	–	<i>Sep</i>	[0, 0.5]
		Baseflow coefficient	–	<i>α</i>	[0, 1]
		Percolation coefficient	–	<i>β</i>	[0, 1]
		Available/soil water capacity	Cm	<i>U_r</i>	[1, 15]
		Degree-day coefficient for snowmelt	cm/°C	<i>D_f</i>	[0, 1]
		Land cover coefficient	–	<i>K_c</i>	[0.5, 1.5]
	Lohmann routing	Subbasin unit hydrograph shape parameter	–	<i>G_S</i>	[1, 100]
		Subbasin unit hydrograph rate parameter	–	<i>G_R</i>	[10 ⁻⁴ , 100]
		Wave velocity in the linearized Saint–Venant equation	m/s	<i>Ve</i>	[0.5, 100]
		Diffusivity in the linearized Saint–Venant equation	m ² /s	<i>Di</i>	[200, 5000]
ABM model		Return flow factor	–	<i>R_f</i>	[0, 0.5]
		Upper flow deviation threshold	m ³ /s	<i>L_U</i>	[25, 50]
		Lower flow deviation threshold	m ³ /s	<i>L_L</i>	[0, 15]
		Learning rate	–	<i>γ</i>	[0, 1]
		Standard deviation modifier	m ³ /s	<i>Sig</i>	[0, 2] ^a
		Prorated ratio	–	<i>R</i>	[0.4, 1]
		Slope of linear model	m ³ /s/cm	<i>a_L</i>	[0, 0.5]
		Intercept of linear model	m ³ /s	<i>b_L</i>	[–2, 2] ^a
		Quadratic coefficient of quadratic model	m ³ /s/cm ²	<i>a_Q</i>	[–27, 11.5] ^a
		Slope of quadratic model	m ³ /s/cm	<i>b_Q</i>	[–12, 42] ^a
	Intercept of quadratic model	m ³ /s	<i>c_Q</i>	[–24, 10] ^a	

^aUnion bound over five diversion agents. Each has customized bound estimated according to historical diversion data.

further includes a baseflow component (Luo et al., 2012) to depict low flow patterns. The Lohmann routing model is a simple linear transfer function model that considers routing both within the subbasin (concentration time of subbasin runoff to reach the outlet) and in rivers (from one upstream subbasin outlet to a lower subbasin outlet). The Lohmann routing model can be derived independently from the GWLF, and connecting its results with GWLF-simulated streamflow forms the semi-distributed hydrological model for the YRB.

The hydrological model in the YRB coupled model (Figure 1) contains 47 parameters that must be calibrated (27 GWLF + 20 Lohmann), where each subbasin has a unique parameterization. The S1, S2, and S3 subbasin models used to simulate reservoir inflows each have 11 parameters involved in the calibration process (9 GWLF + 2 Lohmann, *G_S* and *G_R*). In sum, the four hydrological models (S1, S2, and S3 subbasin models and the YRB coupled model) are independently calibrated. The calibrated parameters, abbreviations, and bounds are summarized in Table 1.

The calibration is performed by a genetic algorithm (GA) with the objective function that differs for each calibration (Table S3 in Supporting Information S1). For each subbasin model, the objective is to maximize the Kling-Gupta efficiency (KGE; Gupta et al., 2009; Equation 5), which comprehensively takes correlation, variability bias, and mean bias into account. For the integrated model, the objective is to maximize a penalized KGE in which $1 - \bar{D}$, the mean annual diversion shortage, is added to the KGE as a penalty factor (Table S3 in Supporting Information S1):

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \quad (5)$$

where r is the Pearson correlation coefficient, and μ and σ denote the mean and standard deviation of flows, respectively. The subscripts *obs* and *sim* refer to observed and simulated streamflow time series, respectively. In the GA setup, we use roulette wheel selection, uniform crossover (crossover probability = 0.5), and uniform mutation (mutation probability = 0.1), which are coded under the Distributed Evolutionary Algorithms in Python (DEAP; Fortin et al., 2012) framework. Theoretically, the global optimum can be found with the mutation mechanism involved if the GA runs long enough, but it is still likely to be trapped in the local optima given limited computational resources, especially with highly nonlinear problems. Hence, we do not aim to find the global optima. Instead, we consider multiple local optima with similar performance (i.e., above a certain threshold) as a form of model configuration uncertainty. The detailed GA calibration settings are provided in Table S3.

3.2. Human Model

3.2.1. Water Diversion Model

For the water diversion model in the YRB coupled model, we define each of the five irrigation districts as agents. These five agents made their annual irrigation requests on March 1st every year. In our design, the annual pattern of the historical diversion is described by three deterministic components and one stochastic component. The three deterministic components are learning, adaptive, and emergency response components.

The *learning component* enables agents to adjust their diversion request references ($Div_{req,ref}$) to achieve a goal (i.e., flow target at the Parker gauge) in the long run by learning from the flow deviation of the flow target through a set of empirical Equations 6–9:

$$Div_{req,ref,y} = Div_{req,ref,y-1} + V_{avg,y} \times \gamma \quad (6)$$

$$V_{avg,y} = \frac{1}{10} \sum_{i=1}^{10} V_{y-i} \quad (7)$$

where γ is a learning rate, subscript y denotes the year, and V_{avg} is the average strength value, with a 10-year rolling window indicating the magnitude and learning direction (e.g., increase $Div_{req,ref,y}$ if $V_{avg,y}$ is positive, or decrease $Div_{req,ref,y}$ if $V_{avg,y}$ is negative). This 10-year rolling window also creates learning momentum in the learning direction of $Div_{req,ref}$ where V_{avg} needs several (consecutive) counter events (e.g., $V = 1$ when $V_{avg} < 0$ or $V = -1$ when $V_{avg} > 0$) to reverse its sign (i.e., learning direction of $Div_{req,ref}$). The events are denoted as V and calculated by Equation 8:

$$V_y = \begin{cases} 1 & \text{if } |De_y| > L_u \text{ if } De_y > 0 \\ -1 & \text{if } |De_y| < L_l \text{ if } De_y < 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$De = Q_{789} - Q_{target} \quad (9)$$

where L_u and L_l are the upper and lower flow deviation thresholds, respectively. These two parameters control how sensitive an agent is to wet or dry hydrological conditions. De is the deviation of the average flow from July to September (Q_{789}) relative to the flow target (Q_{target}), as shown in Equation 9. We provide an example of this learning process in Figure S1 in Supporting Information S1. To maintain a rational value for $Div_{req,ref}$ we subjectively bound $Div_{req,ref}$ in a range of 1.2× the historical maximum and 0.8× the historical minimum from 1960 to 2013, because it is unlikely that diversion will be less than 80% of the historical minimum, given continuous economic growth and associated water use. Also, it is unlikely that the diversion can be greater than 120% of the historical maximum for the years of the simulation, even with technology improvement (e.g., irrigation efficiency), given the physical constraints of available water and canal capacities.

The *adaptive component* is designed to capture short-term variation in diversion requests, which are believed to be induced by different weather conditions. We consider two alternative functional forms governing agent adaptive diversion behavior: linear (Equation 10) or quadratic (Equation 11) functions.

$$Div_{req,mu,y} = Div_{req,ref,y} + a_L \times P_{11-6,y} + b_L \quad (10)$$

$$Div_{req,mu,y} = Div_{req,ref,y} + a_Q \times P_{11-6,y}^2 + b_Q \times P_{11-6,y} + c_Q \quad (11)$$

where a_L and b_L are parameters for a linear function, and subscript y denotes the year. a_Q , b_Q , and c_Q are parameters for a quadratic function. The linear function assumes farmers will strictly divert more water in wet years. The quadratic function can capture the more complex phenomenon. For example, farmers start to divert less water when precipitation reaches a certain level because irrigation is not needed to maintain soil moisture (Figure S2 in Supporting Information S1). $Div_{req,mu}$ is the mean of the annual diversion request. P_{11-6} is the total precipitation from November to June in the reservoir catchments. P_{11-6} abstracts the water storage conditions above the reservoirs (e.g., snowpack and reservoir storage; November–February) before a decision and a perfect precipitation forecast in the first half of the growing season (March–June). Precipitation in the latter half of the growing season (July–October) may not affect diversion requests because this decision is often made before the growing season; thus, it is not as informative in deciding $Div_{req,mu}$.

The term P_{11-6} is also used as an indicator to trigger the emergency response component. If $P_{11-6,y}$ is lower than a given threshold (i.e., drought year), then the *emergency response component* will take the place of the adaptive component and prorate $Div_{req,ref,y}$ according to a calibrated constant ratio (R) in Equation 12:

$$Div_{req,mu,y} = Div_{req,ref,y} \times R \quad (12)$$

Next, the *stochastic component* of agent behavior is computed by:

$$Div_{req,y} = Div_{req,mu,y} + Rn \times Sig \quad (13)$$

where Div_{req} is the annual diversion request, and subscript y denotes the year. We use a matrix to represent the covariance coefficient among historical agent diversion decisions to mimic the “social norm” effect (i.e., farmer diversion decisions will correlate to some degree with neighbor decisions; Bicchieri & Muldoon, 2014), and Rn is a random vector sampled from a multivariate normal distribution with this covariance matrix (i.e., social norm effect matrix). Sig is the calibrated standard deviation modifier, representing the modification of the randomness level. Finally, Div_{req} is disaggregated into a daily scale based on historical monthly reduction proportions to continue the simulation. Note that the return flow (Q_r) is computed from the actual diversion (Div_d) shown below:

$$Q_{r,d} = f_{rout}(R_f \times Div_d) \quad (14)$$

where $f_{rout}(\cdot)$ represents the routing process within the subbasin of the returned outlet, R_f is the return flow factor, and the subscript d is the day.

We simulate five different diversion agent types in the YRB coupled models. The five coupled models are, from simple to complex, (a) static model, M_s , (b) adaptive model with linear functions, $M_{A,L}$, (c) adaptive model with quadratic functions, $M_{A,Q}$, (d) learning adaptive model with linear functions, $M_{L,L}$, and (e) learning adaptive model with quadratic functions, $M_{L,Q}$. The complexities of the coupled models depend on how we simulate agent diversion behaviors. Agents in $M_{L,L}$ and $M_{L,Q}$ have all components, enabling them to learn to capture long-term trends and adapt to short-term shocks, with the difference in selected adaptive function mentioned above. The learning component is omitted in the adaptive models (i.e., $M_{A,L}$ and $M_{A,Q}$), which enables agents to capture short-term variations without learning (constant $Div_{req,ref}$). M_s is a deterministic model that only has the emergency response component, which is designed to mimic a traditional method that considers the diversion demand as a constant/deterministic input. The actual diversion variation from an agent only comes from physical constraints (e.g., minimum flow requirements) and prorated water rights during a drought year, represented by the emergency response component. Table 2 summarizes the diversion agent setup for these five YRB coupled models.

Similar to the hydrological model calibration, we maximize a penalized KGE, where the penalty here is $10 \times (1 - \bar{D})$ (Table S3 in Supporting Information S1). Note that we use the mean of 10 model simulations in

Table 2
Summary of Diversion Agent Type Setup of Five YRB Coupled Models

Component	Coupled model				
	M_s	$M_{A,L}$	$M_{A,Q}$	$M_{L,L}$	$M_{L,Q}$
Emergency response component	✓	✓	✓	✓	✓
Adaptive component	–	✓	✓	✓	✓
Learning component	–	–	–	✓	✓
Stochastic component	–	✓	✓	✓	✓
Number of parameters per agent	2	4	5	7	8
Calibrated parameters	b_L, R, R_f	a_L, b_L, R, Sig, R_f	$a_Q, b_Q, c_Q, R, Sig, R_f$	$\gamma, L_u, L_r, a_L, b_L, R, Sig, R_f$	$\gamma, L_u, L_r, a_Q, b_Q, c_Q, R, Sig, R_f$

an evaluation of the GA algorithm to address the model stochasticity issue. The detailed calibration settings and calibration parameter bounds are provided in Table S3 in Supporting Information S1 and Table 1, respectively. The ODD + D description (Müller et al., 2013) for the ABM is shown in Table S4 in Supporting Information S1.

3.2.2. Reservoir Model

Reservoirs are another human component in the YRB that we consider as input data. To build the reservoir release simulation model for exploratory analysis, we empirically set up monthly operational rules based on the monthly simulated inflows forecast (S1, S2, and S3), storage status, and control periods (flood or storage). Then, the simulated monthly releases are uniformly disaggregated into daily values. This has no effect on our results, which are presented on an annual scale. We group the adjacent Keechelus, Kachess, and Cle Elum reservoirs into one representative for simplicity, as shown in Figure 1. Consequently, we have three reservoir models (R1, R2, and R3). The pseudo-code of reservoir models is provided in Figure S3 in Supporting Information S1.

3.3. Model Uncertainty Decomposition

Model uncertainty is attributed to various uncertainty sources. We adopt a variance-based uncertainty decomposition to isolate model uncertainty caused by different model configurations and input data. Figure 3 shows this concept. Using GA, or any other calibration methods (e.g., dynamically dimensioned search algorithm; Tolson & Shoemaker, 2007), to calibrate nonlinear models often result in multiple model configurations with similar outcomes (because most methods cannot guarantee the global optimum in a nonlinear solution space). We call

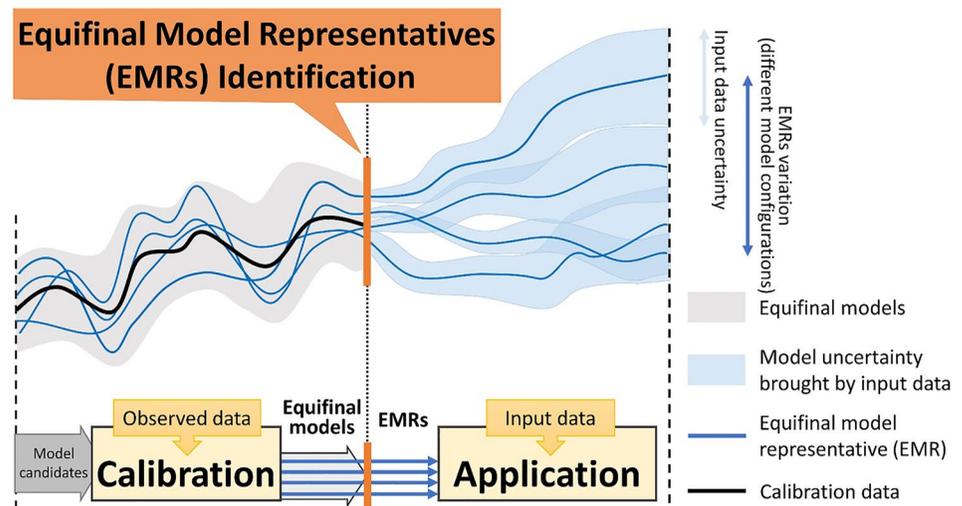


Figure 3. Conceptual diagram of EMRs and model uncertainty decomposition. Blue lines indicate model uncertainty caused by different model configurations. Light blue zones represent model uncertainty introduced by uncertain input data (e.g., climate uncertainty).

such models as “equifinal models” (i.e., different model configurations) in this paper. An infinite number of equifinal models (gray area) could be found within a continuous (e.g., real number) or unbounded equifinal space. For example, infinite parameter combinations are within the given calibration bounds (Table 1). It is impossible to evaluate or identify this infinite number of equifinal models. Therefore, we define an EMRs identification step (orange box in Figure 3) to identify a finite number of EMRs (blue lines in Figure 3) to represent the range of model variability over the equifinal space for the uncertainty evaluation. This study separately analyzes the uncertainty of five coupled model structures with different diversion agent types.

In application, model uncertainty is composed of the different model configurations from EMRs (e.g., variation of blue lines in Figure 3) and the input data (e.g., blue zone in Figure 3). To decompose the model uncertainty, we apply the law of total variance. In the first step, we isolate the uncertainty of the different model configurations (e.g., different parameter sets producing similar outcomes) by:

$$\text{Var}(I) = \text{E}[\text{Var}(I|EMR)] + \text{Var}(\text{E}[I|EMR]) \quad (15)$$

The total model uncertainty, $\text{Var}(I)$, of an output indicator (I ; e.g., Q_M and D_M) is decomposed into two parts: uncertainty caused by the input data, $\text{E}[\text{Var}(I|EMR)]$ and uncertainty caused by the variation of EMRs, $\text{Var}(\text{E}[I|EMR])$. In an exploratory application, input data uncertainty is often related to climate uncertainty. In our study, the climate uncertainty could be further decomposed into climate change scenario uncertainty (e.g., different combinations of PF_{ratio} and TF_{delta}) and internal climate variability (e.g., different realizations under a given climate scenario). We can apply the law of total variance again to $\text{E}[\text{Var}(I|EMR)]$ to have

$$\text{E}[\text{Var}(I|EMR)] = \text{E}[\text{Var}(I|EMR, ICR)] + \text{E}[\text{Var}(\text{E}[I|EMR, ICR])|EMR] \quad (16)$$

where $\text{E}[\text{Var}(I|EMR, ICR)]$ is the model uncertainty resulting from uncertain climate change scenarios. $\text{E}[\text{Var}(\text{E}[I|EMR, ICR])|EMR]$ represents the model uncertainty caused by internal climate variabilities, where ICR is a realization under a given climate change scenario. The derivations of Equations 15 and 16 are provided in Text S1 in Supporting Information S1.

To explain the impact of different model structures on the uncertainty in the co-evolving CNHS, we adopt the following variance formula:

$$\text{Var}(Q_{down}) = \text{Var}(Q_{up} - Div) = \text{Var}(Q_{up}) + \text{Var}(Div) - 2 \times \text{Cov}(Q_{up}, Div) \quad (17)$$

Equation 17 indicates that if the upstream flow (Q_{up} ; streamflow before diversion) and diversion (Div) are positively correlated, then the variance of downstream flow (Q_{down} ; streamflow after diversion) will be lower than the sum of $\text{Var}(Q_{up})$ and $\text{Var}(Div)$. That is, even if uncertainty in both the human and natural systems increases, that is, $\text{Var}(Q_{up})$ and $\text{Var}(Div)$ both increase, it is possible for uncertainty in outputs, $\text{Var}(Q_{down})$ to decrease if Q_{up} and Div are sufficiently positively correlated. This implies that learning and adaptive mechanisms, in which agents change Div in response to Q_{up} , can potentially offset streamflow uncertainty over time. Namely, the agent's action co-evolves with the changing environment after they learn from environmental feedback (Woodard et al., 2019).

The entire numerical experiment is visualized in Figure 4. It consists of (a) calibration and EMRs identification and (b) exploratory simulations and the model uncertainty analysis. In the first part, the calibration is separated for the hydrological model and ABM. We first calibrate the hydrological model using weather data, observed reservoir releases, and irrigation diversions. After reaching the GA termination criteria (e.g., maximum generation), we identify two HydroEMRs following the procedure shown in the orange box in Figure 4. The chosen maximum generation as termination criteria ensures consistency in the number of evaluations for EMRs identification. During the EMRs identification step, we collect all simulations from the GA calibration and then select the top 1% models, that is, those with better fits, as feasible models. Then, we normalize those calibrated parameters into [0,1] according to their calibration bounds (Table 1) and input them into the K-means algorithm (Pedregosa et al., 2011). The number of clusters/EMRs is empirically determined from the explained variance elbow plot (e.g., two clusters for the hydrological model; Bholowalia & Kumar, 2014; Figure S4 in Supporting Information S1). After that, models with the best fit are selected within each cluster as EMRs.

Next, we use these two HydroEMRs to drive ABM calibration. Following the same procedure as calibrating HydroEMRs, four ABMEMRs (elbow plot in Figure S5 in Supporting Information S1) are found for each Hydro-

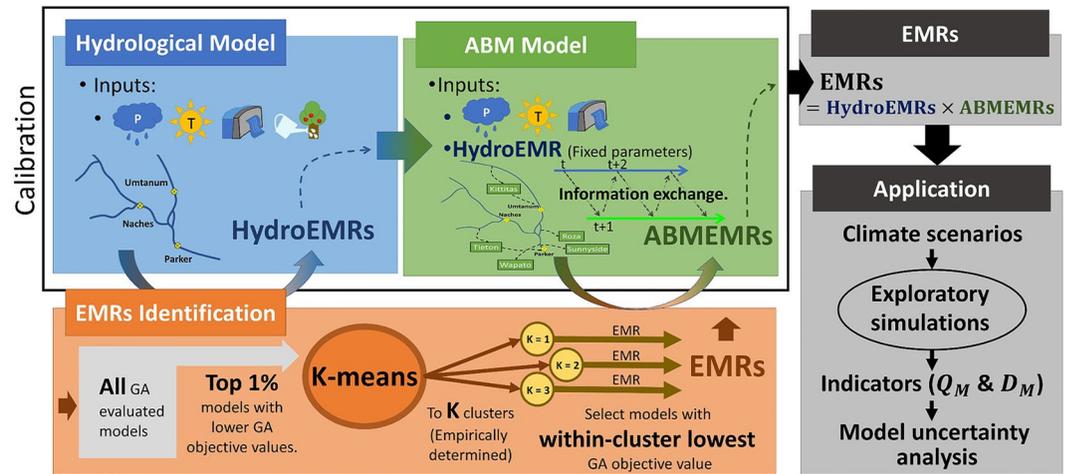


Figure 4. Modeling schema for the numerical experiment. Schema begins by calibrating the hydrological model and identifying HydroEMRs through the K-means algorithm; then, HydroEMRs drive ABM calibration and ABMEMRs identification. Identified EMRs are then applied to exploratory application and uncertainty analysis.

EMRs. Note that the information (e.g., streamflow and diversion requests) is exchanged annually between the hydrological and water diversion models during the evaluation. We refer readers to Hyun et al. (2019) for more details about this coupling technique. Then, we repeat this calibration process for five coupled models with the different diversion agent types introduced in Section 3.2. Therefore, at the end of the first part, eight EMRs (2 HydroEMRs \times 4 ABMEMRs) are identified per coupled model and then applied to the second part, exploratory application. In the second part, we conduct exploratory simulations with 1,000 input realizations and then calculate indicators (Q_M and D_M) for each EMRs of the five coupled models. Then, we repeat this exploratory Experiment 30 times to address the stochastic nature of our model. Consequently, we use 1,200,000 model evaluations (1,000 climate input realizations \times 5 types of diversion agents \times 8 EMRs \times 30 iterations) for the model uncertainty analysis.

4. Results

4.1. Calibration and Validation of EMRs

This section describes the calibration and validation results of EMRs. Table 3 shows that both HydroEMR1 and HydroEMR2 have similar mean KGE values (0.672 and 0.675) in the calibration, and the differences are a result of different compromises in the subbasin KGE performance. For example, HydroEMR1 has a higher KGE at the Naches outlet, while the Umtanum and Parker outlets perform better in HydroEMR2. The monthly time series plot is provided in Figure S6 in Supporting Information S1, and calibration statistics are shown in (Table S6 and Figure S7 in Supporting Information S1).

We show the ABMEMR calibration (1960–1999) and validation (2000–2013) for annual diversion results (separated by vertical dotted lines) for five coupled models (rows), from simple to complex (top to down), in Figure 5. The red lines are the observed data, and the gray areas are 95% confidence intervals. Each EMR result is one of four ABMEMRs (line styles) that are driven by HydroEMR1 (blue lines) or HydroEMR2 (orange lines) and averaged over 10 simulations. For example, EMR2.1 is the result of the coupled model consisting of HydroEMR2 and ABMEMR1. We show the detailed calibration statistics in Table S6 and Figure S8 in Supporting Information S1. In general, the EMRs of M_S show the lowest variations because it is the simplest two-parameter deterministic model type. Two adaptive models ($M_{A,L}$

Period	Model	Mean KGE	Subbasin	KGE
Calibration (1960–1999)	HydroEMR1	0.672	Umtanum	0.653
			Naches	0.750
			Parker	0.611
	HydroEMR2	0.675	Umtanum	0.686
			Naches	0.712
			Parker	0.628
Validation (2000–2013)	HydroEMR1	0.591	Umtanum	0.536
			Naches	0.752
			Parker	0.487
	HydroEMR2	0.599	Umtanum	0.569
			Naches	0.712
			Parker	0.517

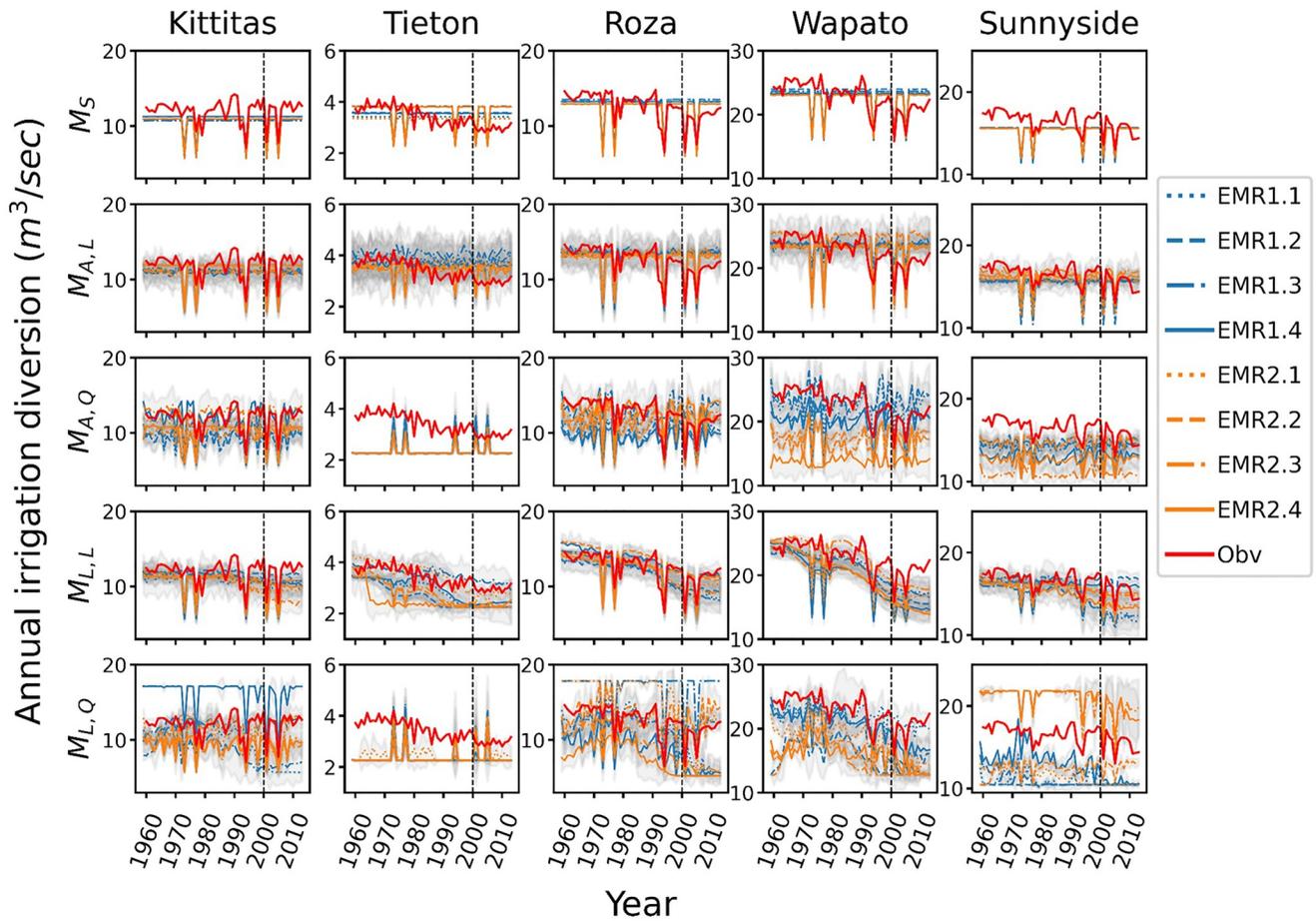


Figure 5. Calibration and validation results (separated by vertical lines) of five diversion irrigation districts (columns) for five coupled models (rows). Each EMR result is one of the four ABMEMRs driven by HydroEMR1 (blue lines) or HydroEMR2 (orange lines) that were averaged over 10 simulations. Red lines are observed data, and gray areas are 95% confidence intervals.

and $M_{A,Q}$) recreate short-term variations (e.g., yearly fluctuations), while two learning adaptive models ($M_{L,L}$ and $M_{L,Q}$) capture both the long-term decrease and short-term variations. Figure 5 also shows that coupled models with quadratic functions have greater variation than the linear function, which implies that coupled models with quadratic functions have larger equifinal space. In Figure 5, column 2, Tieton's diversion pattern is not well-captured, especially in $M_{A,Q}$ and $M_{L,Q}$; however, it has limited influence on the overall results due to the small diversion amount. The “spikes” in Tieton's pattern are the diversion output governed by the emergency response component and the prorated ratio during drought years instead of the adaptive component. Tieton's adaptive component output low diversions to achieve a better calibration objective value of the model. Although such diversion patterns are not realistic, we keep them in our uncertainty analysis because these EMRs satisfied our equifinal model definition (i.e., similar objective value in calibration).

4.2. Comparing Model Complexity and Uncertainty

4.2.1. Uncertainty Decomposition of Natural System Output (Q_M)

We run the identified EMRs under 1,000 future climate realizations 30 times each to calculate the Q_M uncertainty (i.e., $\text{Var}(Q_M)$). In Figure 6, we decompose $\text{Var}(Q_M)$ into three parts that are attributed to three uncertainty sources: climate change scenario (blue areas; combinations of PF_{ratio} and TF_{delta}), internal climate variability (orange areas; bootstrapped realizations), and different model configurations (green areas; variation of EMRs) for five coupled models (columns), with mean values from 30 runs to address stochasticity in the ABM. The diversity of model configurations is contributed by variations of HydroEMRs (light green areas; $\text{Config}_{HydroEMR}$; $\text{Var}(E[|IEMR,HydroEMR|])$) and

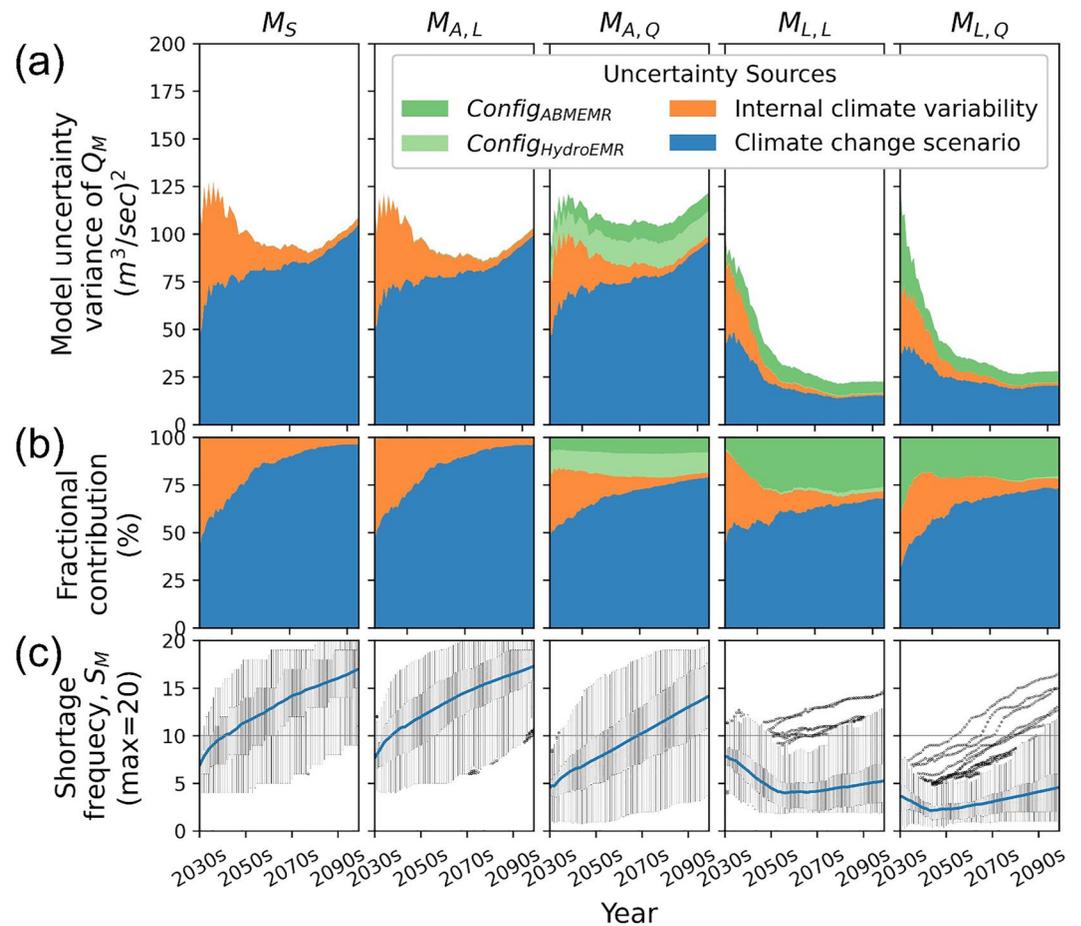


Figure 6. (a) Model uncertainty decomposition of Q_M for five coupled models (columns). (b) Fractional contribution percentage of decomposed variances. (c) Boxplot time series of S_M with mean values (blue lines) and outliers (black dots). Three uncertainty sources were climate scenario uncertainty, internal climate variability, and different model configurations.

ABMEMRs (dark green areas; $Config_{ABMEMR}$; $E[\text{Var}(\mathcal{I}EMR, \text{HydroEMR})]$). Figure 6a shows the actual variance value of Q_M ; Figure 6b indicates the fractional contribution percentage of three uncertainty sources; and Figure 6c is the S_M boxplot time series representing the shortage frequency of 20 internal climate variability realizations over climate change scenarios and EMRs.

For M_S , the uncertainty caused by climate change scenarios increases over the simulation period, while the uncertainty driven by internal climate variability gradually decreases. The model configuration uncertainty is almost negligible (<0.3) in M_S . A similar pattern is observed in $M_{A,L}$, where the model configuration uncertainty (<1) is only slightly larger than M_S . For the more complex model $M_{A,Q}$, we observe a significantly larger contribution from different model configurations. However, this trend is not present in models with the learning component, which we discuss later.

The increasing amount of blue areas in M_S , $M_{A,L}$, and $M_{A,Q}$ indicates the limited ability of the agents in these coupled models to address long-term trends in climate change scenarios since the ABMs are not designed to “learn” from the long-term environmental changes. The adaptive component could absorb short-term shocks caused by internal climate variability, so $M_{A,L}$ and $M_{A,Q}$ have smaller orange areas than M_S ($M_{A,Q} < M_{A,L} < M_S$). However, the shrinking magnitude of uncertainty from climate variability toward the end of the simulation is caused by increasing shortage frequency, S_M (Figure 6c). Diversion shortage implies that streamflow is at the minimum flow requirement, as no more water can be diverted to meet the demand. Therefore, increasing S_M across climate change scenarios and EMRs results in lower internal climate variability in Q_M . In addition, $M_{A,Q}$ shows the smallest uncertainty caused by input data (i.e., blue and orange areas), but it has the largest model

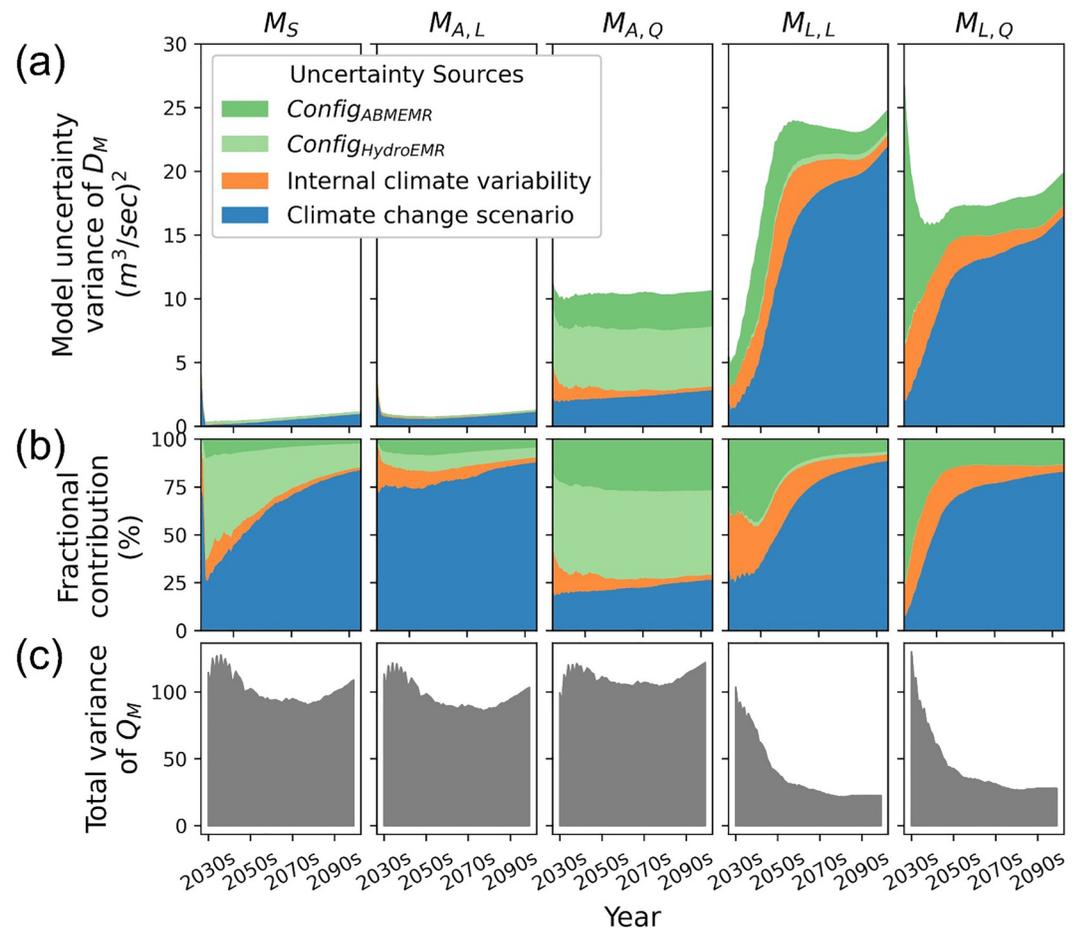


Figure 7. (a) Model uncertainty decomposition of D_M for five coupled models (columns). (b) Fractional contribution percentage of decomposed variances. (c) Total model uncertainty of Q_M . Three uncertainty sources are climate scenario uncertainty, internal climate variability, and different model configurations.

configuration uncertainty. Therefore, ignoring possible model configurations could artificially reduce uncertainty (i.e., no green area), which may affect the interpretation of model results and bias the concluding information.

In Figure 6, coupled models with learning components ($M_{L,L}$ and $M_{L,Q}$) have much smaller Q_M uncertainty because agents in these models can learn to adjust their diversion requests for both long-term and short-term changes; this offsets the Q_M uncertainty contributed by input data (i.e., P_{future} and T_{future}). Furthermore, we observe that the model configuration differences of $M_{L,L}$ and $M_{L,Q}$ is primarily caused by ABMEMRs variations ($Config_{ABMEMR}$), while the model configuration differences in $M_{A,Q}$ are contributed by both HydroEMRs and ABMEMRs variations (Figure 6a). The negligible $Config_{HydroEMR}$ of the two learning adaptive models can again be attributed to learning abilities (i.e., achieving flow target by updating $Div_{req,ref}$). When we use learning adaptive models to simulate irrigation requests, the driver of ABMEMR (i.e., HydroEMR1 or HydroEMR2) does not matter because agents could adjust their diversion behaviors according to feedback from different hydrological environments. Therefore, we observe smaller $Config_{HydroEMR}$ results in the exploratory simulations. Furthermore, the learning abilities of $M_{L,L}$ and $M_{L,Q}$ result in a lower frequency of encountering shortages upon diversion (Figure 6c).

4.2.2. Uncertainty Decomposition of Human System Output (D_M)

After analyzing the model uncertainty of the nature system indicator (Q_M), we next consider how flexible agent behaviors need to be to reach the level of Q_M uncertainty revealed in Figure 6. Therefore, we examine similar variance decomposition plots for D_M (human-system indicator) in Figure 7, with the total model uncertainty of Q_M (Figure 7c).

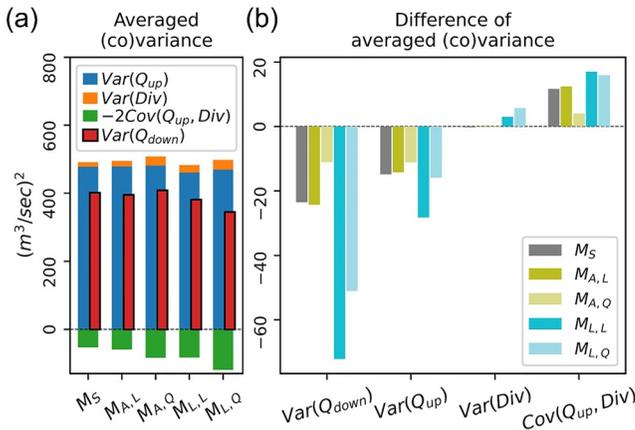


Figure 8. Factors showing co-evolution, particularly with the learning mechanism, leading to overall lower model uncertainty. Q_{down} , Q_{up} , and Div are streamflow before diversion, streamflow after diversion, and diversion at Parker gauge, respectively. (a) Averaged (co)variances of five coupled models from 2030 to 2100. (b) Differences between averaged (co)variance from 2066 to 2100 and from 2030 to 2065.

For M_S , the D_M uncertainty gradually increases over time due to the climate change scenario uncertainty (Figure 7a). As we discussed earlier, diversion shortage dominates in M_S , where the quantity of available water under different input data (i.e., P_{future} and T_{future}) determines the model uncertainty for diversion behavior. Namely, if there is no diversion shortage, the variance of D_M will be a flat line, with only the contribution of different model configurations because the diversion request is a fixed value in M_S .

The greater internal climate variability fractional contribution (orange area in Figure 7b) in $M_{A,L}$ and $M_{A,Q}$, compared to M_S , is contributed by agent adaptability (e.g., different diversion requests based on water supply conditions). Furthermore, variations caused by HydroEMR1 and HydroEMR2 play an important role in driving D_M model configuration variations in M_S , $M_{A,L}$, and $M_{A,Q}$, demonstrated by greater $Config_{HydroEMR}$ than $Config_{ABMEMR}$ in Figure 7b. This is because M_S , $M_{A,L}$, and $M_{A,Q}$ cannot learn. Therefore, the identified EMRs driven by different HydroEMRs show distinct patterns in calibration (Figure 5), leading to larger contributions from different model configurations in D_M uncertainty.

The two learning adaptive models, $M_{L,L}$ and $M_{L,Q}$, show distinct patterns from the other three models. The D_M uncertainty caused by climate change scenarios increases significantly over time, which is the opposite behavior as the Q_M shown in Figure 7c. This further indicates that agents can learn from

environmental feedback (e.g., flow deviations) and then mitigate some environmental uncertainty through their flexibility in adjusting long-term diversion behaviors. These results visualize an opposing trend between Q_M and D_M uncertainties in CNHS. In the next section, we will mathematically explain this learning behavior.

4.3. Model Uncertainty and Co-Evolution in CNHS

In the previous analysis, we show the uncertainty decomposition for both natural (Q_M) and human (D_M) system outputs. We claim that the learning mechanism limits $Var(Q_M)$ with more complex model designs. Although the two learning models exhibit greater uncertainties in human system behavior, the increasing magnitude of D_M uncertainty is less than the decreasing magnitude of Q_M uncertainty. We hypothesize that co-evolution, particularly in the learning mechanism of CNHS, leads to such results. To test this hypothesis, we adopt Equation 17 and show the results at the Parker gauge in Figure 8.

Figure 8 shows the value of each term in Equation 17 using the Parker gauge as an example. Q_{down} , Q_{up} , and Div are the streamflow before diversion, the streamflow after diversion, and the diversion at Parker gauge, respectively. The averaged (co)variances of five coupled models (x-axis) from 2030 to 2100 are shown in Figure 8a. Figure 8b shows differences between the averaged (co)variance from 2066 to 2100 and from 2030 to 2065.

Equation 17 shows that $Var(Q_{down})$ can only be less than $Var(Q_{up})$ only if $Cov(Q_{up}, Div)$ is not 0 and $2 \times Cov(Q_{up}, Div)$ is greater than $Var(Div)$. Following this logic, Figure 8a mathematically demonstrates that our designed agents act (i.e., divert) differently (i.e., co-evolve; green bars) according to the environment and indicates a decreasing pattern in $Var(Q_{down})$ as the model becomes more complex (red bars). For example, M_S diverts different amounts of water based on the available water (Q_{up}). $M_{A,L}$ and $M_{A,Q}$ have larger covariance values because of their adaptability to short-term shocks. $M_{L,Q}$ and $M_{L,L}$ have the highest covariance value because they have both adaptive and learning capabilities. The $Cov(Q_{up}, Div)$ pattern of models with different complexities is highly correlated with the pattern in $Var(Div)$, verifying the design of our diversion agent given how closely an agent co-evolves with the environment.

The decreasing trend in Q_M uncertainty (e.g., $M_{L,Q}$ and $M_{L,L}$ in Figure 6a) only occurs when the difference in the summation of $Var(Q_{up})$ and $Var(Div)$ is lower than the difference in $2 \times Cov(Q_{up}, Div)$ between the former and latter simulation periods (Figure 8b). The negative difference of M_S and $M_{A,L}$ in $Var(Q_{down})$ is primarily contributed by the decrease in $Var(Q_{up})$ and the shortage-induced $Cov(Q_{up}, Div)$. $M_{A,Q}$ has a similar explanation: the difference in $Var(Div)$ is almost zero (i.e., the adaptability remains the same over the entire simulation). The most interesting results are in the two learning models, $M_{L,Q}$ and $M_{L,L}$, where increasing $Var(Div)$ results in a

significant increase in $Cov(Q_{up}, Div)$ that leads to a larger magnitude of decreasing $Var(Q_{down})$. This demonstrates our explanation for Q_M variance decomposition results in the previous section and supports our hypothesis: learning behaviors in a human system can potentially offset the impact of a natural system's variability on output uncertainty (i.e., $Var(Q_M)$). However, we do not extend these results to claim that the learning adaptive model is a better design. This study quantifies model uncertainty properties for different designs used to model CNHS. It is important to verify the existence of learning behavior or any other structural designs in a chosen case study; otherwise, uncertainty may be introduced from inappropriate structural assumptions. For this particular case study, we may argue that $M_{L,L}$ is a more appropriate model design as it has higher objective values in calibration (Table S6 in Supporting Information S1).

5. Discussion

5.1. Other Sources of Uncertainty in CNHS

Our results, which show that learning adaptive models have less observed variance in the nature system (e.g., Q_M), align with the philosophy of “adaptive management” strategies in facing increasing climate change uncertainty (Giordano & Shah, 2014; Karthe et al., 2021). However, it is also important to consider the causes of decreasing nature system uncertainty. First, our results show that, if possible model configurations are ignored, Q_M uncertainty is only driven by input data whose uncertainty appears to be smaller in the adaptive models than in the static model. This might lead to a false conclusion that the adaptive model is more reliable (i.e., less uncertain). Second, although utilizing learning adaptive models shows a decreasing trend in Q_M uncertainty, they could be impacted by uncertainty from incomplete knowledge of proposed model structural assumptions, such as a learning structure (Karthe et al., 2021), which would not be directly revealed in the modeling results. Such introduced assumption-based uncertainty could be a barrier to interpreting modeling results and implementing adaptive management strategies (Allen & Gunderson, 2011; Lee, 2001). Therefore, continuously monitoring (i.e., acquiring data) and communicating uncertainty among modelers, policymakers, and other stakeholders is essential.

While this study decomposes specific uncertainties (e.g., input data and different model configurations), other uncertainty sources are not analyzed, such as indicator selection (Khatami et al., 2019), model resolution level (Saltelli et al., 2019), and coupling structure (e.g., data exchange frequency and exchanged information). Some benchmark study cases may help the CNHS modeling community eliminate indicator selection factors when evaluating model uncertainty across different model designs (Vallario et al., 2021). However, the indicator selection procedure requires its own systematic study. For example, we choose KGE in this study, but other widely used indicators, like Nash-Sutcliffe efficiency (Nash & Sutcliffe, 1970) or R-squared, might lead to different uncertainty analysis results. Investigations of the uncertainty caused by the model resolution level and coupling structures will benefit the use of CNHS models in cross-scale (temporal and spatial) studies (Aburto et al., 2012; Fleischmann et al., 2019), especially when coupling to models with various simulation timesteps and frequency of information exchange. However, this requires each coupled-model component to be run over the full spectrum of temporal and spatial resolutions to set up comparable numerical experiments. This is particularly challenging when coupling to large-scale models or standalone software, which are often restricted to coarse or limited temporal or spatial resolutions.

5.2. Limitations

We identify several limitations in our study that could be improved in future studies. First, the reservoir settings are highly simplified, where reservoir releases are considered as input data. Interactions between reservoirs and water demands are ignored. Namely, the CNHS of the YRB is only partially tested. Second, we assume that calibrated agent behaviors (e.g., γ for learnability; a_L and b_L for linear adaptability; and a_Q , b_Q , and c_Q for quadratic adaptability) could be directly applied to exploratory analyses. In addition, agent learnability and adaptability might not remain the same in the future (Aburto et al., 2012) for more accessible and accurate forecast information or changing environmental regulations. Third, we subjectively bound $Div_{req,ref}$ based on historical records. This might affect uncertainty analysis results since canal capacity is possible to increase over our assumption (e.g., 120% of the historical maximum). Fourth, our numerical experimental design ignores some uncertainty sources. For example, we only consider a subset of input data uncertainty (i.e., climate change scenarios and

internal climate variability), we ignore the downscaling uncertainty, and we follow the conventional calibration and validation procedure under the stationary assumption of the nature system, which a more comprehensive study can be done to evaluate the uncertainty brought by the stationary assumption. Also, based on the equifinal model definition, our identified EMRs could be approximations of a single local optimum rather than multiple local optima. We encourage the adoption of pre-calibration methods (Tarawneh et al., 2016), alternate calibration algorithms (Tolson & Shoemaker, 2007), and carefully selected feasibility model criteria (i.e., equifinal models) for model configuration uncertainty quantification in future projects focusing on case studies.

6. Conclusions

A deeper understanding of CNHS modeling enables us to better use modeling results to inform policymaking. This study explores modeling uncertainties. We quantify and then decompose the uncertainty of a coupled natural-human model (i.e., a semi-distributed hydrological model coupled with an agent-based water diversion model) into three sources by the law of total variance: (a) climate scenario uncertainty, (b) climate internal variability, and (c) model configuration uncertainty. The YRB in the Northwest US is adopted as our study area, where irrigation districts are defined as agents. We analyze how co-evolution influences the relationship between model output uncertainty and model complexity through five coupled models with different water diversion agent types: (a) a static model, (b) an adaptive model with linear functions, (c) an adaptive model with quadratic functions, (d) a learning adaptive model with linear functions, and (e) a learning adaptive model with quadratic functions.

Our hypotheses—(a) model uncertainty in an exploratory analysis will likely increase with model complexity, given uncertain input data (e.g., climate forcing), and different model configurations, and (b) the inclusion of a learning mechanism in a human system can potentially offset the impact of natural system's variability on output uncertainty—are accepted, according to our results. The two learning adaptive models show a decreasing trend in the natural system output because agents learn and adapt to environmental changes via co-evolution between the two subsystems. The learnability and adaptability are revealed by increased variability of the human system outputs. Although the learning adaptive models generate smaller nature system output uncertainty, the modeler should be aware of how assumptions of model structure (e.g., Is the learning assumption appropriate in the exploratory analysis of a given case study?) affect the results. Finally, additional uncertainty sources should be investigated in future work, such as indicator selection, model resolution level, and coupling structure with a more sophisticated agent setup and scale-up experiments.

Data Availability Statement

The weather data used in the study are available at Livneh et al. (2015), and data on reservoir storage and release, streamflow, and diversion of the YRB are available at the Hydromet platform (<https://www.usbr.gov/pn/hydromet/yakima/>). The GCMs data can be downloaded from the National Center for Atmospheric Research platform (<https://climatedataguide.ucar.edu/climate-model-evaluation/cmip-climate-model-intercomparison-project-overview>; National Center for Atmospheric Research Staff (Eds)). Also, the authors would like to thank Mr. Edward C. Young from the USBR for providing GIS shapefiles of streamflow monitoring sites and stream networks. The authors make our code available at <https://doi.org/10.5281/zenodo.6564777> (Lin & Ethan Yang, 2022).

References

- Aburto, M. O., de los Angeles Carvajal, M., Barr, B., Barbier, E. B., Boesch, D. F., Boyd, J., et al. (2012). *Ecosystem-based management for the oceans*. Island Press.
- Allen, C. R., & Gunderson, L. H. (2011). Pathology and failure in the design and implementation of adaptive management. *Journal of Environmental Management*, 92(5), 1379–1384. <https://doi.org/10.1016/j.jenvman.2010.10.063>
- Allison, A. E., Dickson, M. E., Fisher, K. T., & Thrush, S. F. (2018). Dilemmas of modelling and decision-making in environmental research. *Environmental Modelling & Software*, 99, 147–155. <https://doi.org/10.1016/j.envsoft.2017.09.015>
- Arendt, P. D., Apley, D. W., & Chen, W. (2012). Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. *Journal of Mechanical Design*, 134(10), 100908. <https://doi.org/10.1115/1.4007390>
- Arsenault, R., & Brissette, F. P. (2014). Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches. *Water Resources Research*, 50(7), 6135–6153. <https://doi.org/10.1002/2013wr014898>
- Bauch, C. T., Sigdel, R., Pharaon, J., & Anand, M. (2016). Early warning signals of regime shifts in coupled human-environment systems. *Proceedings of the National Academy of Sciences of the United States of America*, 113(51), 14560–14567. <https://doi.org/10.1073/pnas.1604978113>

Acknowledgments

The work was supported by the US National Science Foundation (NSF): CBET #1941727. The authors would like to thank the editors and three anonymous reviewers for their suggestions to improve the quality of the manuscript.

- Berglund, E. Z. (2015). Using agent-based modeling for water resources planning and management. *Journal of Water Resources Planning and Management*, 141(11), 04015025. [https://doi.org/10.1061/\(asce\)wr.1943-5452.0000544](https://doi.org/10.1061/(asce)wr.1943-5452.0000544)
- Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16(1), 41–51. [https://doi.org/10.1016/0309-1708\(93\)90028-E](https://doi.org/10.1016/0309-1708(93)90028-E)
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of hydrology*, 320(1–2), 18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>
- Beven, K. (2016). Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, 61(9), 1652–1665. <https://doi.org/10.1080/02626667.2015.1031761>
- Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3), 279–298. <https://doi.org/10.1002/hyp.3360060305>
- Bholowalia, P., & Kumar, A. (2014). EBK-means: A clustering technique based on elbow method and k-means in WSN. *International Journal of Computer Applications*, 105(9).
- Bicchieri, C., & Muldoon, R. (2014). Social norms. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/archives/spr2014/entries/social-norms/>
- Blair, P., & Buytaert, W. (2016). Socio-hydrological modelling: A review asking "why, what and how?" *Hydrology and Earth System Sciences*, 20(1), 443–478. <https://doi.org/10.5194/hess-20-443-2016>
- Brown, C. M., Ghile, Y., Lavery, M., & Li, K. (2012). Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. *Water Resources Research*, 48(9). <https://doi.org/10.1029/2011wr01212>
- Brown, C. M., Lund, J. R., Cai, X., Reed, P. M., Zagana, E. A., Ostfeld, A., et al. (2015). The future of water resources systems analysis: Toward a scientific framework for sustainable water management. *Water Resources Research*, 51(8), 6110–6124. <https://doi.org/10.1002/2015WR017114>
- Di Baldassarre, G., Brandimarte, L., & Beven, K. (2016). The seventh facet of uncertainty: Wrong assumptions, unknowns and surprises in the dynamics of human–water systems. *Hydrological Sciences Journal*, 61(9), 1748–1758. <https://doi.org/10.1080/02626667.2015.1091460>
- Ekblad, L., & Herman, J. D. (2021). Toward data-driven generation and evaluation of model structure for integrated representations of human behavior in water resources systems. *Water Resources Research*, 57(2), e2020WR028148. <https://doi.org/10.1029/2020wr028148>
- Elsawah, S., Filatova, T., Jakeman, A. J., Kettner, A. J., Zellner, M. L., Athanasiadis, I. N., et al. (2020). Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modelling*, 2, 16226. <https://doi.org/10.18174/sesmo.2020a16226>
- Elshafei, Y., Sivapalan, M., Tonts, M., & Hipsey, M. R. (2014). A prototype framework for models of socio-hydrology: Identification of key feedback loops and parameterisation approach. *Hydrology and Earth System Sciences*, 18(6), 2141–2166. <https://doi.org/10.5194/hess-18-2141-2014>
- Fleischmann, A., Paiva, R., & Collischonn, W. (2019). Can regional to continental river hydrodynamic models be locally relevant? A cross-scale comparison. *Journal of Hydrology*, 3, 100027. <https://doi.org/10.1016/j.jhydro.2019.100027>
- Fortin, F. A., De Rainville, F. M., Gardner, M. A. G., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, 13(1), 2171–2175.
- Giordano, M., & Shah, T. (2014). From IWRM back to integrated water resources management. *International Journal of Water Resources Development*, 30(3), 364–376. <https://doi.org/10.1080/07900627.2013.851521>
- Giuliani, M., Li, Y., Castelletti, A., & Gandolfi, C. (2016). A coupled human-natural systems analysis of irrigated agriculture under changing climate. *Water Resources Research*, 52(9), 6928–6947. <https://doi.org/10.1002/2016WR019363>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Haith, D. A., Mandel, R., & Wu, R. S. (1996). *Generalized watershed loading functions version 2.0 user's manual*. Cornell University.
- Haith, D. A., & Shoemaker, L. L. (1987). Generalized watershed loading functions for stream flow nutrients. *Journal of the American Water Resources Association*, 23(3), 471–478. <https://doi.org/10.1111/j.1752-1688.1987.tb00825.x>
- Hung, F., & Yang, Y. E. (2021). Assessing adaptive irrigation impacts on water scarcity in nonstationary environments—A multi-agent reinforcement learning approach. *Water Resources Research*, 57(9), e2020WR029262. <https://doi.org/10.1029/2020wr029262>
- Hyun, J. Y., Huang, S. Y., Yang, Y. C. E., Tidwell, V., & Macknick, J. (2019). Using a coupled agent-based modeling approach to analyze the role of risk perception in water management decisions. *Hydrology and Earth System Sciences*, 23(5), 2261–2278. <https://doi.org/10.5194/hess-23-2261-2019>
- Kandasamy, J., Sounthararajah, D., Sivabalan, P., Chanan, A., Vigneswaran, S., & Sivapalan, M. (2014). Socio-hydrologic drivers of the pendulum swing between agricultural development and environmental health: A case study from Murrumbidgee River Basin, Australia. *Hydrology and Earth System Sciences*, 18(3), 1027–1041. <https://doi.org/10.5194/hess-18-1027-2014>
- Karthe, D., Bogardi, J. J., & Borchardt, D. (2021). Water resources management: Integrated and adaptive decision making. In *Handbook of water resources management: Discourses, concepts and examples* (pp. 365–381). Springer. https://doi.org/10.1007/978-3-030-60147-8_12
- Kelly, R. A., Jakeman, A. J., Barreteau, O., Borsuk, M. E., ElSawah, S., Hamilton, S. H., et al. (2013). Selecting among five common modeling approaches for integrated environmental assessment and management. *Environmental Modelling & Software*, 47, 159–181. <https://doi.org/10.1016/j.envsoft.2013.05.005>
- Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and flux mapping: A new approach to model evaluation and process representation under uncertainty. *Water Resources Research*, 55(11), 8922–8941. <https://doi.org/10.1029/2018wr023750>
- Lee, K. N. (2001). Appraising adaptive management. In *Biological diversity* (pp. 3–26). CRC Press. <https://doi.org/10.1201/9781420042597.sec1>
- Lin, C.-Y., & Ethan Yang, Y. C. (2022). Code: The effects of model complexity on model output uncertainty in co-evolved coupled natural–human systems (v0.0.1). [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.6564777>
- Livneh, B., Bohn, T. J., Pierce, D. W., Muñoz-Arriola, F., Nijssen, B., Vose, R., et al. (2015). *A spatially comprehensive, meteorological data set for Mexico, the U.S., and southern Canada (NCEI Accession 0129374)*. NOAA National Centers for Environmental Information. Dataset. <https://doi.org/10.7289/v5x34vf6>. Accessed June 4, 2021.
- Livneh, B., & National Center for Atmospheric Research Staff (Eds). *Last modified 12 Dec 2019. "The Climate Data Guide: Livneh gridded precipitation and other meteorological variables for continental US, Mexico and southern Canada."* Retrieved from <https://climatedataguide.ucar.edu/climate-data/livneh-gridded-precipitation-and-other-meteorological-variables-continental-us-mexico>
- Lohmann, D., Raschke, E., Nijssen, B., & Lettenmaier, D. P. (1998). Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. *Hydrological Sciences Journal*, 43(1), 131–141. <https://doi.org/10.1080/02626669809492107>
- Luo, Y., Arnold, J., Allen, P., & Chen, X. (2012). Baseflow simulation using SWAT model in an inland river basin in Tianshan Mountains, Northwest China. *Hydrology and Earth System Sciences*, 16(4), 1259–1267. <https://doi.org/10.5194/hess-16-1259-2012>
- Madani, K., & Hooshyar, M. (2014). A game theory–reinforcement learning (GT–RL) method to develop optimal operation policies for multi-operator reservoir systems. *Journal of Hydrology*, 519, 732–742. <https://doi.org/10.1016/j.jhydrol.2014.07.061>

- Marston, L., & Konar, M. (2017). Drought impacts to water footprints and virtual water transfers of the Central Valley of California. *Water Resources Research*, 53(7), 5756–5773.
- McLean, K. A., & McAuley, K. B. (2012). Mathematical modelling of chemical processes—Obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. *Canadian Journal of Chemical Engineering*, 90(2), 351–366. <https://doi.org/10.1002/cjce.20660>
- Messina, J. P., Evans, T. P., Manson, S. M., Shortridge, A. M., Deadman, P. J., & Verburg, P. H. (2008). Complex systems models and the management of error and uncertainty. *Journal of Land Use Science*, 3(1), 11–25. <https://doi.org/10.1080/17474230802047989>
- Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863), 573–574. <https://doi.org/10.1126/science.1151915>
- Moallemi, E. A., Kwakkel, J., de Haan, F. J., & Bryan, B. A. (2020). Exploratory modeling for analyzing coupled human-natural systems under uncertainty. *Global Environmental Change*, 65, 102186. <https://doi.org/10.1016/j.gloenvcha.2020.102186>
- Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., et al. (2013). Describing human decisions in agent-based models—ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software*, 48, 37–48.
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- National Center for Atmospheric Research Staff (Eds). Last modified 06 Jul 2016. "The Climate Data Guide: CMIP (Climate Model Intercomparison Project) overview." Retrieved from <https://climatedataguide.ucar.edu/climate-model-evaluation/cmip-climate-model-intercomparison-project-overview>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning Research*, 12, 2825–2830.
- Reuss, M. (2003). Is it time to resurrect the Harvard water program? *Journal of Water Resources Planning and Management*, 129(5), 357–360. [https://doi.org/10.1061/\(asce\)0733-9496\(2003\)129:5\(357\)](https://doi.org/10.1061/(asce)0733-9496(2003)129:5(357))
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., et al. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, 114, 29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
- Song, W., & Zhang, Y. (2015). Expansion of agricultural oasis in the Heihe River Basin of China: Patterns, reasons and policy implications. *Physics and Chemistry of the Earth, Parts A/B/C*, 89, 46–55. <https://doi.org/10.1016/j.pce.2015.08.006>
- Srikrishnan, V., & Keller, K. (2021). Small increases in agent-based model complexity can result in large increases in required calibration data. *Environmental Modelling & Software*, 138, 104978. <https://doi.org/10.1016/j.envsoft.2021.104978>
- Sun, Z., Lorscheid, I., Millington, J. D., Lauf, S., Magliocca, N. R., Groeneveld, J., et al. (2016). Simple or complicated agent-based models? A complicated issue. *Environmental Modelling & Software*, 86, 56–67. <https://doi.org/10.1016/j.envsoft.2016.09.006>
- Tarawneh, E., Bridge, J., & Macdonald, N. (2016). A pre-calibration approach to select optimum inputs for hydrological models in data-scarce regions. *Hydrology and Earth System Sciences*, 20(10), 4391–4407. <https://doi.org/10.5194/hess-20-4391-2016>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498. <https://doi.org/10.1175/bams-d-11-00094.1>
- Tolson, B. A., & Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1). <https://doi.org/10.1029/2005wr004723>
- Tung, C. P., & Haith, D. A. (1995). Global-warming effects on New York streamflows. *Journal of Water Resources Planning and Management*, 121(2), 216–225. [https://doi.org/10.1061/\(asce\)0733-9496\(1995\)121:2\(216\)](https://doi.org/10.1061/(asce)0733-9496(1995)121:2(216))
- USBR. (2002). *Interim comprehensive basin operating plan for the Yakima Project Washington*.
- USDA. (2017). National Agricultural Statistics Service, 2017 Census of Agriculture. Retrieved from www.nass.usda.gov/AgCensus
- Vallario, B., Arnold, J., Weyant, J., Scheetz, A., Lipschultz, F., Arrigo, J., et al. (2021). Coastal integrated HydroTerrestrial modeling: A multi-agency invited workshop. U.S. Global Change Research Program. Retrieved from https://downloads.globalchange.gov/coasts-IG/workshop-report/C-IHTM_Workshop_Report_Nov2020.pdf
- Vrugt, J. A., terBraak, C. J. F., Gupta, H. V., & Robinson, B. A. (2009). Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, 23(7), 1011–1026. <https://doi.org/10.1007/s00477-008-0274-y>
- Wada, Y., Bierkens, M. F., Roo, A. D., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., et al. (2017). Human–water interface in hydrological modelling: Current status and future directions. *Hydrology and Earth System Sciences*, 21(8), 4169–4193. <https://doi.org/10.5194/hess-21-4169-2017>
- Walsh, J. E., Bhatt, U. S., Littell, J. S., Leonawicz, M., Lindgren, M., Kurkowski, T. A., et al. (2018). Downscaling of climate model output for Alaskan stakeholders. *Environmental Modelling & Software*, 110, 38–51. <https://doi.org/10.1016/j.envsoft.2018.03.021>
- Westerberg, I. K., Di Baldassarre, G., Beven, K. J., Coxon, G., & Krueger, T. (2017). Perceptual models of uncertainty for socio-hydrological systems: A flood risk change example. *Hydrological Sciences Journal*, 62(11), 1705–1713. <https://doi.org/10.1080/02626667.2017.1356926>
- Williams, T. G., Guikema, S. D., Brown, D. G., & Agrawal, A. (2020). Assessing model equifinality for robust policy analysis in complex socio-environmental systems. *Environmental Modelling & Software*, 134(8). <https://doi.org/10.1016/j.envsoft.2020.104831>
- Wi, S., Yang, Y. C. E., Steinschneider, S., Khalil, A., & Brown, C. M. (2015). Calibration approaches for distributed hydrologic models in poorly gaged basins: Implication for streamflow projections under climate change. *Hydrology and Earth System Sciences*, 19(2), 857–876. <https://doi.org/10.5194/hess-19-857-2015>
- Woodard, D. L., Davis, S. J., & Randerson, J. T. (2019). Economic carbon cycle feedbacks may offset additional warming from natural feedbacks. *Proceedings of the National Academy of Sciences*, 116(3), 759–764. <https://doi.org/10.1073/pnas.1805187115>
- Xu, W., Lowe, S. E., & Adams, R. M. (2014). Climate change, water rights, and water supply: The case of irrigated agriculture in Idaho. *Water Resources Research*, 50(12), 9675–9695. <https://doi.org/10.1002/2013wr014696>
- Yang, Y. C. E., Cai, X., & Stipanović, D. M. (2009). A decentralized optimization algorithm for multiagent system-based watershed management. *Water Resources Research*, 45(8), 8430. <https://doi.org/10.1029/2008WR007634>
- Yang, Y. C. E., Son, K., Hung, F., & Tidwell, V. (2020). Impact of climate change on adaptive management decisions in the face of water scarcity. *Journal of Hydrology*, 588, 125015. <https://doi.org/10.1016/j.jhydrol.2020.125015>
- Zellner, M. L. (2008). Embracing complexity and uncertainty: The potential of agent-based modeling for environmental planning and policy. *Planning Theory & Practice*, 9(4), 437–457. <https://doi.org/10.1080/14649350802481470>