Spring 2022

Interventions Towards Sustainable Watershed Management as Demonstrated by Hydrologic Simulation

Shan Zuidema
University of New Hampshire, Durham

Interventions Towards Sustainable Watershed Management as Demonstrated by Hydrologic Simulation

By

Shantar Zuidema

Bachelor of Science, Environmental Earth Science, Eastern Connecticut State University, 2004

Master of Science, Hydrology, University of New Hampshire, 2011

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

In

Earth & Environmental Science

May 2022

This dissertation was examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Earth & Environmental Sciences by:

Dissertation Director, J. Matthew Davis, Associate Professor of Hydrogeology

Dissertation Director, Wilfred M. Wollheim, Associate Professor of Natural Resources and the Environment

Anne F. Lightbody, Associate Professor of Hydrology

Richard B. Lammers, Research Associate Professor, Institute for the Study of Earth, Oceans, and Space

Mark B. Green, Senior Research Associate, Department of Earth, Environmental, and Planetary Sciences, Case Western Reserve University

On April 11, 2022

Approval signatures are on file with the University of New Hampshire Graduate School.

TABLE OF CONTENTS

DEDICATION

For my daughters June and Hayley. Remember to think deeply, act decisively, and always carry love in your hearts.

ACKNOWLEDGEMENTS

My sincere gratitude goes to all members of my committee for their mentorship, rigor, and comradery through my journey to this point. I especially need to acknowledge my Co-Advisors Dr. J. Matthew Davis for seeing in me the potential to flourish in academy, and to provide space and pressure in equal measure; and to Dr. Wilfred Wollheim for his perspective and perseverance. I must also acknowledge my sincere appreciation for the mentorship of Dr. Richard Lammers. Dr. Anne Lightbody and Dr. Mark Green, though I may not have called on you as often as I may have wished, the knowledge that your eyes would fall on this body of work has shaped my decisions immensely.

Beyond my advisors and committee, I am deeply indebted to my colleagues in the Water Systems Analysis Group. Robert Stewart, Madeleine Mineau, Alex Prusevich, Stanley Glidden, and Danielle Grogan are in no small part responsible for my happiness and my success. Everyone in the lab, the rotating cast of graduate students especially, have enriched my time in the lab. I have benefited greatly from the mentorship and example from many throughout the Institute for the Study of Earth, Oceans, and Space and need to name Dr. Cameron Wake and Dr. Steve Frolking specifically for their support, conversation, and approach to research.

My research has been decidedly transdisciplinary, and I have many too many to thank throughout our multi-institutional research teams. My conceptualization of solutions in watershed management and views of appropriate research strategies has been broadened through my work with colleagues throughout the institutions involved in the Ecosystem & Society project through New Hampshire Established Program to Stimulate Competitive Research, at the Center for Resilient Communities at the University of Idaho, at Penn State University, at Purdue University, the University of Sheffield, and certainly many more institutions.

And finally, and most importantly, I dearly thank my family for being alongside me throughout this journey, and assuredly all our remaining journeys.

Funding for this research was provided by the National Science Foundation (NSF) Established Program to Stimulate Competitive Research award 1101245, NSF and United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Innovations at the Nexus of Food, Energy, and Water Systems program Project Grants 1639524 and 1855937, New Hampshire Agricultural Experiment Station (NHAES) through USDA NIFA Hatch Project 0225006, and the Department of Energy Program on Coupled Human and Earth Systems award DE-SC0016162.

PREFACE

> Given the complexity of the process of designing rules to regulate the use of common-pool resources, I argue that all policy proposals must be considered as experiments.
>
> – E. Ostrom, (1999)

The research presented in this dissertation represents contributions to transdisciplinary research projects, each examining critical issues at the intersection of human society and the natural environment. Three chapters each examine a unique environmental or resource problem, and an evaluation of possible management interventions that address the problem. Chapters I and II have been published in the peer reviewed literature, and citations for these works can be found in the bibliography under Zuidema et al. (2018) and (2020), respectively. Chapter III is being prepared as a manuscript as well. The three studies together provide a cross-section of issues facing water resources and aquatic habitat, and by examining them together, a potentially generalizable framework emerges for understanding key requirements for successful adaptations in sustainable watershed management.

The research presented here was inspired and motivated by the interdisciplinarity of the research teams. It is my sincere hope that the findings presented here can advance the role of hydrologic sciences and hydrologic modeling in the context of large societally relevant questions through the nascent discipline of sociohydrology. Apparent in the progression of the three studies is the formalization of the key understanding regarding the nature and design of adaptations developed through my interactions with a diverse set of scholars in hydrology, biogeochemistry, decision-science, environmental economics, and geography.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Increasing population, changing climate, and on-going legacies of environmental mismanagement motivate our need for deeper understanding of the process and limits of adaptation towards sustainable management of water resources. Movements towards open-science and transdisciplinary research have enabled deeper assessments of the co-evolution of human society and changing landscapes. Policies and decisions enabling environmental restoration or sustainable resource use have been actively pursued for decades. The social barriers that prevent adaptations to succeed are deep and entrenched, but equally important are the physical barriers. Successful adaptations in water resource management need to explicitly consider the joint interactions of intervention magnitude, or intensity, over the feasible extent of its operation. While a seemingly simple concept, many solutions to water resource management would be impossible to achieve without adequate consideration of these constraints.

In these three studies specific management practices were evaluated in the context of whole watershed responses and found to characterize this common constraint despite the diversity of applications. The three studies impose alternative management practices within a model of watershed-scale hydrologic processes, and the success of each practice was ultimately determined by the geographic constraints over which it could act, not by any deficiency in the policy's capacity to affect a sufficient intensity of change.

In Chapter 1, current rates of road salt loading and potential levels of aquatic habitat impairment are estimated for a New England watershed. The potential for reducing impairment through a combination of reduced salt application and buildout are investigated. Chapter 2 examines issues of aquifer sustainability in the Pacific Northwest and evaluates tradeoffs in modernizing

irrigation technology. As irrigation efficiency increased less water recharged the aquifer, which exacerbated aquifer drawdown. Drawdown was offset by enhanced aquifer recharge directly from the river. The study analyzes the constraints under which aquifer drawdown can be eliminated while minimizing any reduction in streamflow. Chapter 3 evaluates the efficacy of two programs that incentivize the restoration of wetlands within the Mississippi River basin to reduce nitrogen export as nitrate export to the Gulf of Mexico. A more thorough consideration of geographic and engineering constraints on restoration illustrates how complementary management practices would be necessary to meet nutrient reduction goals. Finally, Chapter 4 analyzes the three studies and develops the concepts of *intensity* and *extensity* in successful practices in watershed management. Chapter 4 also lays out the common methodology of model experimentation *in silico* used throughout these studies, and defends the epistemological framework chosen.

CHAPTER I:

CONTROLS OF CHLORIDE LOADING AND IMPAIRMENT AT THE RIVER NETWORK SCALE IN NEW ENGLAND

Increasing sodium chloride in watersheds alters aquatic biodiversity (Findlay and Kelly, 2011b; Cañedo-Argüelles et al., 2013, 2016). Both climate and development interact to control chloride contamination in temperate watersheds (Jackson and Jobbagy, 2005; Kaushal et al., 2005; Corsi et al., 2010, 2015). Aquatic biota are sensitive to salt content directly through biological maintenance of cellular osmotic balance (Cañedo-Argüelles et al., 2013), indirectly through mobilization of toxic metals (Findlay and Kelly, 2011b), or correlatively through a host of chemical and habitat alterations (de Zwart et al., 2006). The widespread increases of freshwater salinization (Murray, 1977; Ghassemi et al., 1995; Kaushal et al., 2005; Anning and Flynn, 2014), and growing evidence for altered aquatic community structure at salt concentrations lower than regulatory recommendations (Findlay and Kelly, 2011b; Morgan et al., 2012; Cañedo-Argüelles et al., 2013) motivate increased attention to regulating this ecotoxicological hazard (Cañedo-Argüelles et al., 2016).

Globally (Cañedo-Argüelles et al., 2013; Ghassemi et al., 1995), and in the United States (Anning and Flynn, 2014), irrigation is the primary cause of freshwater salinization. However, in New England where irrigated agriculture is minimal, the U.S. Environmental Protection Agency rated 22.8% of streams as poor or fair for salt content (USEPA, 2013). Since the mid-20th century, winter-time roadway deicing has become a primary driver of increasing freshwater salinity throughout northern temperate regions (Jackson and Jobbagy, 2005). Concentrations of chloride in streams covary with sodium chloride (halite) purchased for deicing purposes in northeastern (Godwin et al., 2003; Jackson and Jobbagy, 2005; Kaushal et al., 2005; Kelly et al., 2008; Trowbridge et al., 2010) and mid-western (Sander et al., 2007; Corsi et al., 2010, 2015) watersheds. Most roadway deicers are chloride salts, with halite being by far the predominant

source (Granato et al., 2015). We currently lack generalizable estimates of road salt application rates at regional scales.

Road salt application rates are largely controlled by the amount of snow and frozen precipitation received (Sander et al., 2007; Kelly et al., 2008). A significant fraction of applied deicer infiltrates to and accumulates in groundwater. As a result, streams maintain elevated chloride concentrations during summer when groundwater dominates streamflow (Godwin et al., 2003; Kelly et al., 2008, 2012; Daley et al., 2009; Kincaid and Findlay, 2009; Cooper et al., 2014). Baseflow chloride concentrations can exceed concentrations that biota can tolerate for extended periods (Findlay and Kelly, 2011b; Corsi et al., 2015) when animal activity is often greatest (Demars et al., 2011).

Empirical studies of road-salt driven salinization have previously focused on multi-year trends and have focused less on interannual meteorological variability, though such variability is evident (e.g. Kaushal et al., 2005, fig. 2). In many temperate regions, such as New England, significant changes are expected to both winter and summer climate (Hayhoe et al., 2006; Campbell et al., 2010; Wake et al., 2014a, b). Though the duration winters are expected to warm and shorten, it is not clear that the form of precipitation will necessarily lead to a reduction in required road salt usage (Arvidsson et al., 2012). Presently there have been few attempts to utilize macro-scale models to investigate changes in road salt usage and chloride impairment associated with interannual dynamics of climate and human development.

In this study, we examined the ability of a macro-scale model to represent the dynamics of chloride solute transport and characterize potential chloride impairment of a relatively large watershed across both space and time. We developed and applied the Non-point Anthropogenic Chloride Loading (NACL) and transport model within the Framework for Aquatic Modeling in

the Earth System (FrAMES) (Vörösmarty et al., 1998; Wollheim et al., 2008b; Wisser et al., 2010; Stewart et al., 2013). We focused on simulating chloride because the single anion is the primary driver of salinization throughout the study region, and is largely conservative in the hydrologic system (Kirchner et al., 2010) making it ideal for process-based representation. However, we relate chloride to specific conductance because the latter was the primary metric measured, and characterizes the total osmotic stress experienced by vulnerable aquatic species (Cañedo-Argüelles et al., 2013). Using data from an extensive network of electrical conductance sensors we inferred a regional loading and transport function to parameterize FrAMES-NACL. We hypothesized that road salt application rates derived from *inventory estimates* at northern U.S. locations (Godwin et al., 2003; Sander et al., 2007; Trowbridge et al., 2010) are the same as derived by calibration to stream chloride concentrations. Furthermore, we hypothesized that climate indices of summer dryness are as important to predict interannual variability in warm-season impairment as frozen precipitation the preceding winter in both small and large rivers. Understanding chloride response during meteorologically different years enhances our ability to manage chloride pollution of watersheds as annual precipitation regimes (Hayhoe et al., 2006) or land-cover (Samal et al., 2017) shift.

Materials and Methods

Study Watershed

The Merrimack (13,000 km^2) River watershed is representative of the northeastern USA watersheds (Figure 1) with increasing development towards the river outlet. Beginning in the White Mountain National Forest, the Merrimack flows 280 km south to the head of tide dam in Lawrence, MA. The watershed includes pristine and highly valued aquatic ecosystems that are important habitat for sport fishing (NH Fish and Game Department, 2015), and may be

threatened by increasing salinity. Though we focus on the Merrimack River watershed (MRW) because the watershed is illustrative of the processes relevant throughout the region, we utilize additional observational data from the neighboring Piscataqua River watershed (PRW) to inform model behavior.

The population in the MRW has been increasing steadily at 5-15% per decade since 1940 (NHOEP; U.S. Census Bureau, 2015), and is expected to double in the coming century (Bierwagen et al., 2010). The population residing in the MRW is approximately 1.7 million people. Impervious areas have increased from 4.6 to 5.1% between 2000 and 2010. For the MRW, annual precipitation averages 1100 mm, precipitation during winter months averages 260 mm, with 220 mm of winter precipitation falling in frozen forms. Projections of future climate suggest the potential for warmer winters (Contosta et al., 2017) consistent with recent trends (Burakowski et al., 2008) that may counteract increasing salt loading concomitant with continued development.

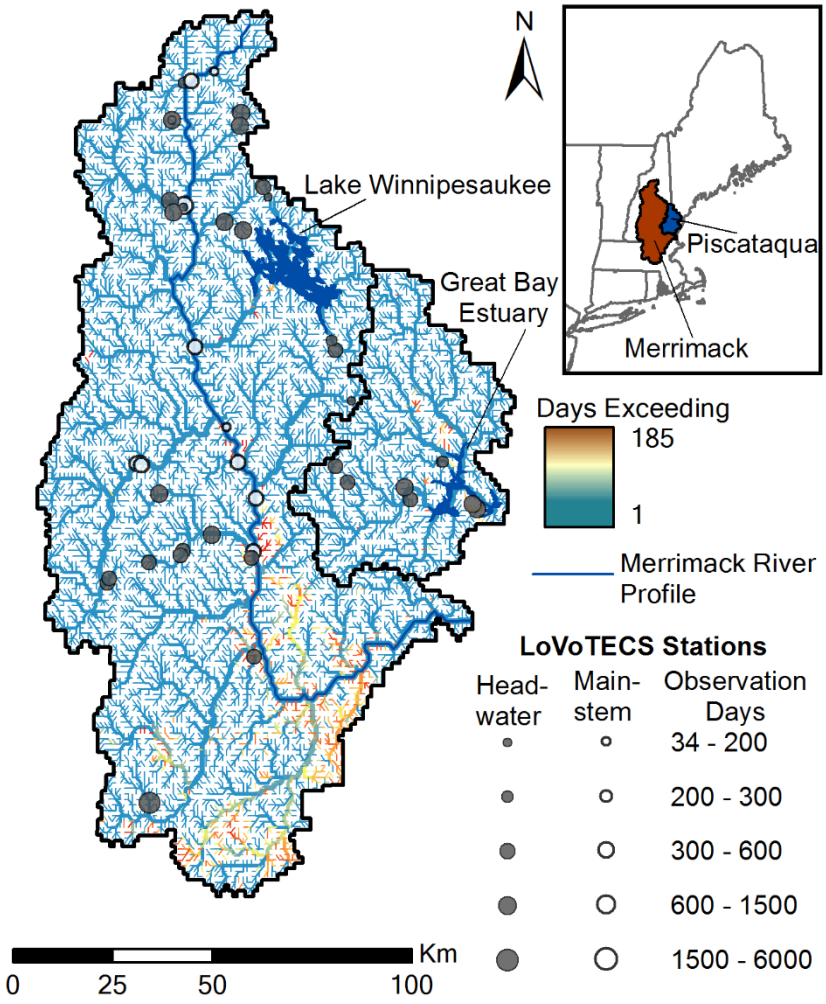


Figure 1: Study domain in the Merrimack and Piscataqua River watersheds in New Hampshire and neighboring Maine and Massachusetts. The map indicated the median number of days exceeding the 600-mS cm^{-1} threshold between 1998 and 2014. Lotic Volunteers Temperature, Electrical Conductivity, and Stage (LoVoTECS) stations in headwater (less than order 6) and mainstem (order 6 and greater) are depicted in closed gray and open white circles, respectively, and size is based on the number of days with qualified data. Lake Winnipesaukee and the Great Bay Estuary were excluded from analysis.

Conductivity Sensor Network

A network of in-situ, high frequency sensors characterized specific conductance in streams, rivers, and storm drainage infrastructure in diverse landcover throughout New Hampshire,

eastern Massachusetts, and western Maine (Contosta et al., 2017; Inserillo et al., 2017). The Lotic Volunteers Temperature, Electrical Conductivity and Stage (LoVoTECS) network collected data throughout the study region from 2011 to 2016. The network utilized long-term in-situ deployments of data-logging potentiometric electrical conductivity and pressure transducer sensors (and HOBO U24 specific conductivity and HOBO U20 stage, respectively) at locations throughout the study region (Figure 1). Sampled catchments exhibited 0 – 54% impervious cover and ranged in size from 1.4 km^2 to 7300 km^2. Sensor data are distributed through the Data Discovery Center (http://ddc.unh.edu). Periodic chemical analysis of stream water throughout the network showed a strong linear correspondence (Figure AI.1) between specific conductance (k_0) and riverine chloride concentration (C_R) across the network following Equation 1, similar to relationships reported by others (Novotny et al., 2008; Trowbridge et al., 2010).

$$k_0 = 3.96C_R + 31.2 \qquad \text{Eq. 1}$$

At low chloride concentrations, Equation 1 is greater than observed chloride concentrations evidenced by the separation between the linear best fit from locally-weighted scatter-plot smoothing (LOWESS) regression of the data around 20 $mg\ \mathrm{Cl}\ L^{-1}$ (Figure AI.1), which is attributed to the decreasing contribution of chloride to total electrolytic composition in pristine waters. The linear regression equation was used to convert between chloride and specific conductance in this study because it was consistent with the LOWESS at higher chloride concentrations that were the focus of the study. We found no significant differences from Equation 1 based on stream size or sub-region.

The LoVoTECS network included 47 stations with data in the Merrimack and Piscataqua River watersheds (Figure 1). These stations recorded data at 1 to 3 minute intervals, which were

resampled to hourly averages for comparison to chemistry data, and resampled to daily averages for model calibration and validation. For model calibration, we used data from 39 stations that had between 170 and 1240 days of data between 2011 and 2015.

Model and Application

The Non-point Anthropogenic Chloride Loading (NACL) Model

We developed a module for Non-point Anthropogenic Chloride Loading (NACL) within FrAMES, a fully distributed, rasterized modeling platform used for studies of hydrologic and aquatic biogeochemical processes across scales (Wollheim et al., 2008b; Wisser et al., 2010; Stewart et al., 2011, 2013). FrAMES controls vertical water transfer and terrestrial runoff generation (Figure AI.2a) which is routed through a 1-D simulated river network (Vörösmarty et al., 1998; Wisser et al., 2010). The model uses a series of conceptual stores representing snowpack, soil storage, quick-flow storage, baseflow generating groundwater storage, and river storage across each 45-second ($\approx$1.4 km^2) pixel of the study domain. Chloride moves through soils and groundwater conservatively according to the simulated hydrological partitioning (Figure AI.2b). Chloride from precipitation and road salt is applied to the watershed at the soil or impervious surface. Chloride from domestic and agricultural sources is loaded directly to shallow groundwater. FrAMES-NACL transports chloride mass through the river network using a linear reservoir routing scheme. Within each reach, chloride is converted to specific conductance (k_0 [μS cm^{-1}]) using Equation 1. Additional changes to FrAMES to account for the dynamics of chloride including long-term storage in groundwater and direct snowmelt on impervious surfaces are discussed in the Appendix I. We parameterize chloride loading from precipitation, domestic waste, agricultural runoff, and road salt application using data and assumptions defined in Appendix I. Briefly, road salt is applied to a fraction of treated

impervious areas based on amount of received snowfall, at a rate specified by the deicer loading parameter (C_{DEI} [kg Cl $mm^{-1}m^{-2}$]). FrAMES-NACL is tested by comparing the relationship between chloride and impervious cover between simulations and observations (Kaushal et al., 2005; Daley et al., 2009). Table AI.1 provides descriptions of 10 parameters of the model relevant to this study. Geographic and meteorological data ingested by the FrAMES-NACL are described in Appendix I.

Assessing road salt loading

To address our first hypothesis, we use FrAMES-NACL to calibrate deicer application rate (C_{DEI}), and compare to empirical estimates of deicer application rates obtained from three studies (Godwin et al., 2003; Sander et al., 2007; Trowbridge et al., 2010) that inventoried road salt usage. Calibration used a Markov-Chain Monte Carlo (MCMC) approach (Appendix I) aimed to minimize model-observation misfit for discharge (Q [mm d^{-1}]) and chloride, after conversion to specific conductance (k_0 [µS cm^{-1}]). Observations represented daily mean discharge (Q) from USGS stream gages (n=28) and specific conductance (k_0) (n=39 LoVoTECS network stations). We compared station specific probabilities of non-exceedance (PONE) for both discharge and specific conductivity using the acceptance ratio (Appendix I), Nash-Sutcliffe efficiency, root mean squared error, and median residuals. We also assessed model performance using root mean square error on daily and temporally aggregated time-series for the suite of stations in our dataset.

Characterizing road salt impairment

We utilize FrAMES-NACL to investigate the influence of climatic drivers interacting with present day land use to generate riverine chloride impairment throughout the river network. We quantify the degree to which interannual climate variability explains variation in two metrics of

impairment in time and space. We select a threshold of 600 [$\mu S\ cm^{-1}$] to define chloride impairment (Appendix I) during the biological active (productive) season of mid-April through October, typical of ecotoxicological assessments of aquatic species (Ejsmond et al., 2010; Ippolito et al., 2012; Kraus et al., 2013). Impaired reach days (IRD) is calculated by counting the number of productive season days the riverine specific conductance predicted for a grid-cell exceeds 600 $\mu S\ cm^{-1}$ multiplying by the local river length in the grid cell and summing for the entire river network. We also use the outlet summer concentration (OSC) from USGS gaging station 01100000 on the Merrimack River at Lowell, Massachusetts as a metric that integrates all upstream processes. FrAMES-NACL predictions of stations that exhibited impairment were validated against LoVoTECS stations exhibiting impairment using the Peirce skill score. The Peirce skill score is the probability of detecting an exceedance (600 $\mu S\ cm^{-1}$) minus the probability of a false detection (Manzato, 2007). We controlled for stations with incomplete data records. Comparisons were limited to stations with at least 90 days of data during each productive season. To validate predictions of the OSC metric, predicted and observed August mean concentrations were compared at all fifth or greater order stations along the Merrimack River (n=8 stations).

The importance of seasonality on stream impairment due to road salt was characterized by correlating interannual variance in simulated impairment with a suite of watershed average meteorological indices. Indices tested for control on stream impairment included the total annual precipitation, total season precipitation (preceding winter [DJF], spring [MAM], summer [JJA], autumn [SON]), preceding winter snowfall, annual number of dry days (days with less than 1 mm precipitation), the annual number of days with heavy (>10 mm) precipitation, number of summer ($T_{avg} \geq 25°C$) days, and the number of dry (<1 mm) summer ($T_{avg} \geq 25°C$) days.

We expected the number of dry days each year to be the most predictive of chloride impairment throughout the river network, as it should be the main driver of baseflow conditions when chloride levels are high. We standardized each year's impairment metrics and climatic indices for years between 1998 and 2014 (n $= 17$) to z-scores after confirming normal distributions using a Shapiro-Wilks test at $\alpha = 0.05$. Using ordinary least-squares, we regress impairment metrics against climate indices testing for significance at $\alpha = 0.05$.

Results and Discussion

NACL model behavior

The time-series of discharge indicate reasonable timing and magnitude of flows across sites (Figure AI.4). At the optimal parameter values found by the MCMC, the median station residual of mean summer runoff was +9.8% of observations, and ranged from -79% to +55%. Probability of non-exceedance (PONE) of discharge correspond closely between observations and FrAMES-NACL (acceptance ratio R_A and NSE were 0.88 and 0.93 for calibration stations and 0.85 and 0.99 for validation stations, respectively Table AI.2, Figure AI.5a) with a very small negative bias of -0.04 $[m^3 s^{-1}]$ (Table AI.2), reflecting reliable representation of the distribution of storm magnitudes. FrAMES-NACL predicts flashier hydrology at all watershed scales, and as time-series are averaged over longer intervals, root mean square error (RMSE) of modeled runoff decreases from 3.3 to 2.3 $m^3 s^{-1}$ (Table AI.2).

FrAMES-NACL correctly identifies stations that are observed to exceed the 600 $\mu S\ cm^{-1}$ threshold (Figure AI.5b), indicating that the model can be used to predict the spatial distribution of exceedance values. The acceptance ratio (R_A) calculated on probabilities of non-exceedance is 0.53 for calibration stations, and 0.38 for downstream stations reserved for validation (Table AI.2). The calibrated NSE on specific conductivity for PONE is 0.75, and lower for validation

stations (Table AI.2). Because low values of specific conductance are over-predicted by Equation 1, specific conductivity at stations with low human development, including many of the stations on the largest river segments, is biased high indicated by a positive median residual $\hat{r}_{0.5}$ (Table AI.2) and decreasing the NSE for these stations. Despite the low NSE, absolute performance described by the root mean square error ($32\ \mu S\ cm^{-1}$, Table AI.2) at downstream stations are much lower than for upstream stations ($180\ \mu S\ cm^{-1}$) indicating that upstream errors cancel out.

FrAMES-NACL correctly captures 1) temporally stable baseflow chloride concentrations 2) small increases in baseflow concentration late in the summer for moderately to highly developed catchments, and 3) winter-time peaks from deicer application only in highly developed catchments (Figure 2). As with discharge times-series, resampling the specific conductivity time-series to week or longer timescales (most consistent with our analysis) improves the RMSE of model performance. Whereas annual modeled specific conductivity exhibited a high bias (Table AI.2), mean summer specific conductance was biased slightly low (median residual -7.4%). During summer storms, the model's dilution response is often flashier than observed, consistent with the flashy response observed in discharge. Since we focus on summer periods when exceedance values are most biologically relevant, the low bias means our estimates of impairment are conservative.

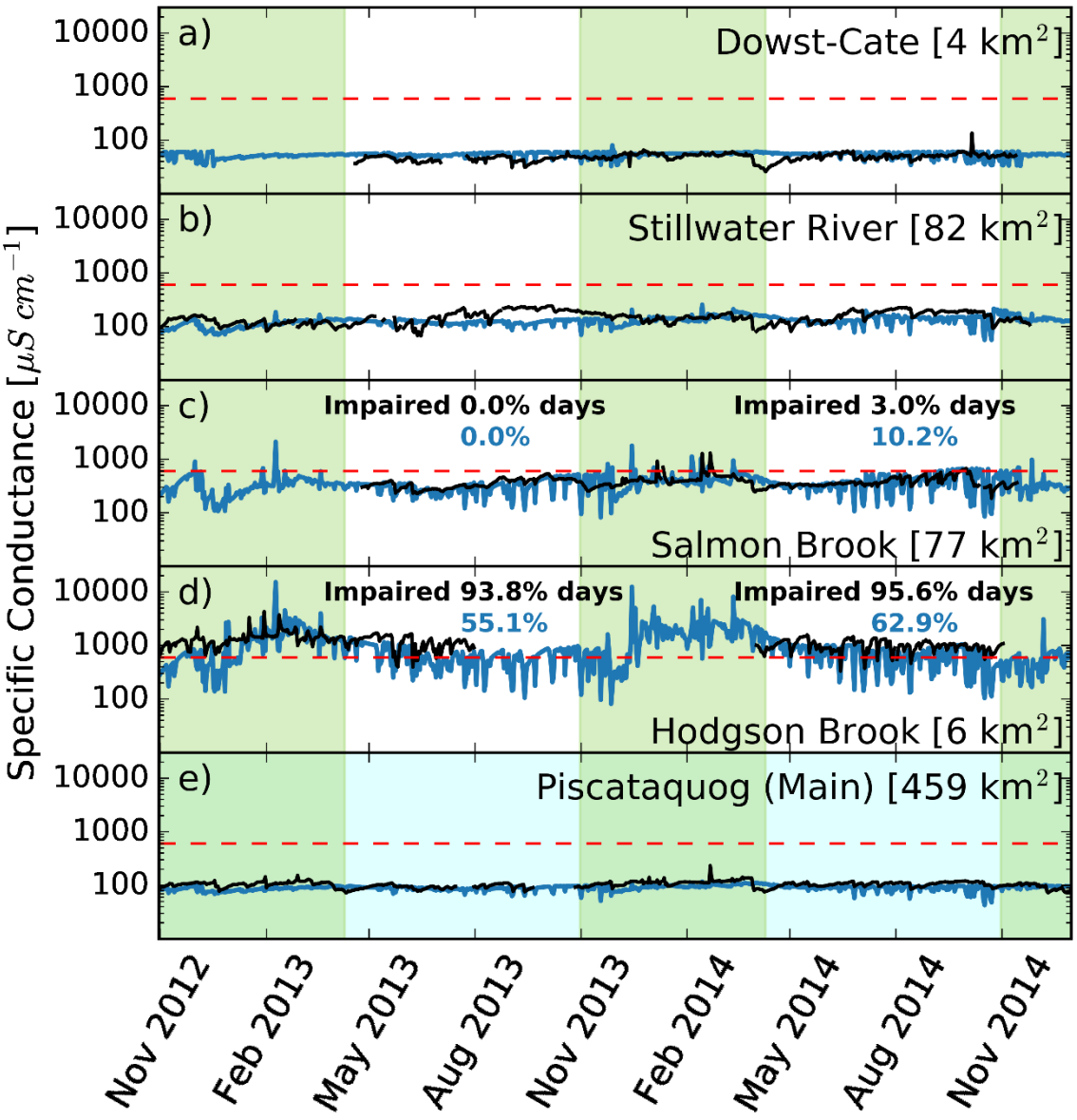


Figure 2: Time series of specific conductivity at five stations comparing the Framework for Aquatic Modeling in the Earth Systems (FrAMES) Nonpoint Anthropogenic Chloride Loading (NACL) data (blue) with observational data (black). Values in parentheses denote catchment area. Light bands indicate the productive season (mid-April through October). Dashed red lines depict the 600-mS cm−1 impairment threshold. The percentage of reporting days exceeding the impairment threshold is provided for each year where applicable.

Mean annual chloride concentration in rivers correlates with upstream development density

(Nimiroski and Waldron, 2002; Kaushal et al., 2005; Novotny et al., 2008; Daley et al., 2009;

Trowbridge et al., 2010; Kelting et al., 2012; Corsi et al., 2015). FrAMES-NACL predicted increases in mean annual chloride as a function of impervious surfaces similar to Daley *et al*. (2009) for the study region (Figure 3a). FrAMES predictions are also similar to those reported by Kaushal *et al.* (2005) over the range of imperviousness in their study.

Road salt loading parameterization

FrAMES-NACL predicted an average road salt loading rate of 6.5 g Cl $mm^{-1}m^{-2}$ with a 95% credible interval in (5.3,11 g Cl $mm^{-1}m^{-2}$) (Figure AI.6). The majority (95%) of C_{DEI} posterior distribution is within the uncertainty of the local inventory estimate from Trowbridge et al. (2010) (T10), and completely bounded by the range of all three inventory studies. Moreover, the appropriateness of estimated road salt inputs (36 ± 24 [$Mg\ Cl\ km^{-2}$]) from FrAMES-NACL is further supported by the consistency in the ratio of road salt to total inputs of chloride between this study (89%) and estimates from southeastern New York (91%, (Kelly et al., 2008)). A lower ratio of road salt to total chloride in Minnesota (61%) (Novotny et al., 2009) reflects significantly lower agricultural input in the MRW. The similarity in ranges of model uncertainty and inventory uncertainty suggest the model is capturing road salt loading to within expectations. Therefore, our hypothesis that deicer application rates inferred from stream chemistry and from inventories of applicators are equivalent is supported. Methodology presented here may be an appropriate alternative to inventorying deicer application rates, or to corroborate, or scale estimates from such studies.

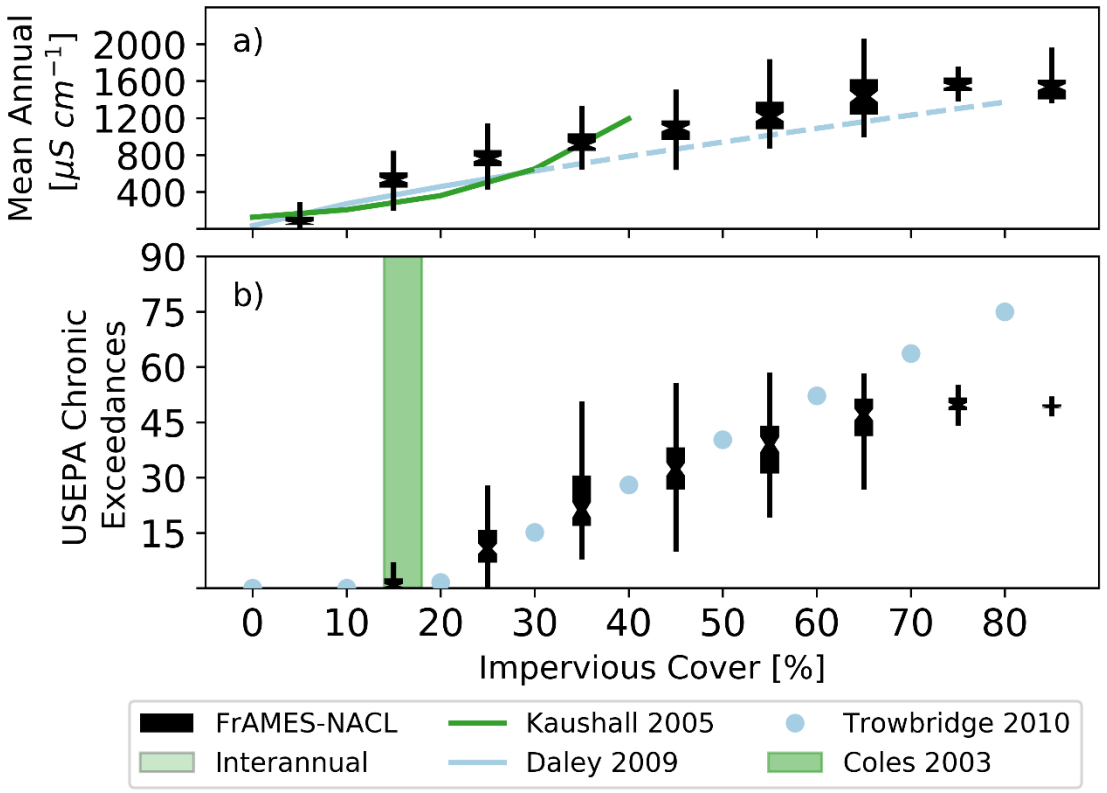


Figure 3: Boxplots binned by imperviousness of the Framework for Aquatic Modeling in the Earth Systems (FrAMES) - Nonpoint Anthropogenic Chloride Loading (NACL) model for headwater (a) flow-weighted mean annual specific conductance, and (b) number of USEPA chronic threshold exceedances for chloride. Relationships between imperviousness and mean annual chloride concentrations (Kaushall et al., 2005; Daley et al., 2009) are given in terms of specific conductivity following Eq. [1]. USEPA chronic exceedances of the chloride standard from Trowbridge et al. (2010) uses mean annual concentration from Daley et al. (2009). The band of most biologic stress identified by Coles et al. (2004) is shown in (b).

Despite the consistencies between FrAMES-NACL estimates of loading and the pool of inventory studies, loading in the Merrimack was distinct from any one inventory. Best estimates of loading rate from three inventories yield 3.8, 11, and 12 $g\ \mathrm{Cl}\ mm^{-1}m^{-2}$ from T10, Sander et al. [2007] (S07), and Godwin et al. [2003] (G03), respectively. C_{DEI} therefore has a mode 73% higher than the local inventory of T10, and 40 to 45% lower than values from other inventories (G03,S07).

Differences in loading may be attributed to location or to time varying application rates. Mean winter temperature in Twin Cities S07 is approximately $3°C$ cooler than the MRW explaining the higher loading from that study (by reduced effectiveness of sodium chloride as a deicer); however, climatologic explanations do not explain greater loading in the Mohawk River G03 than the MRW because mean winter temperatures are similar. The study catchments of T10 are located within the MRW, but are located towards the southern extent of the domain at lower elevations, making a climatological explanation possible since average winter-temperatures average approximately $6°C$ warmer in southern portions of the watershed. In addition, temporal variation in loading rates (Jackson and Jobbagy, 2005) provide a realistic explanation for greater estimated road salt loading between our study and that by T10. Stream concentrations used for calibration reflect average loading over the recent past integrated by the transit time distributions through the studied catchments. Road salt loading may be higher on average over that time period than the single winter of the study of T10.

Chloride impairment in the Merrimack River Watershed

Validation of chloride impairment metrics

FrAMES-NACL accurately identified stations in our observation network exceeding the impairment threshold ($600\ [\mu S\ cm^{-1}]$). The Peirce skill score was a perfect 1.0 ± 0.02 for identifying stations with exceedances of either 1 day or 30 days above the threshold for stations that have at least 90 days of data in a productive season. However, FrAMES-NACL under-predicts the number of days above $600\ [\mu S\ cm^{-1}]$ by about 55% for stations that do exhibit exceedances. Under-prediction of exceedance duration is consistent with the more dynamic nature of conductivity in the model (e.g. Figure 2 and AI.4b). Observations indicate that stations cluster near zero and 100 days of impairment (for stations with complete data coverage through

the productive season) whereas FrAMES-NACL predicts a more even distribution of intermediate days impaired. Thus, FrAMES-NACL provides an underestimate of number of days impaired, while accurately estimating locations of impairment, allowing us to make qualified inferences at the network scale.

FrAMES-NACL captured both the longitudinal trend in observed chloride between 2012 and 2014, and at the furthest downstream station (at Lowell MA) between 1999 and 2004 (Figure AI.7), which validate the Outlet Summer Concentration (OSC) metric. All observational data on Figure AI.7 were not used in model calibration. For the basin profile stations, the relative absolute error was $15.4 \pm 5.3\%$ (Figure AI.7b). Moreover, the mean of summertime USGS grab samples from 1999 to 2004 at Lowell (195 ± 44 [$\mu S\ cm^{-1}$], $n = 23$) was not significantly ($\alpha \leq 0.05$) different from simulated values on sampled days (213 ± 64 [$\mu S\ cm^{-1}$]) (Figure AI.7c).

Recent chloride impairment

FrAMES-NACL predicts that up to 11% of the Merrimack River network exceed the 600 $\mu S\ cm^{-1}$ chloride threshold at least one day in the productive season (Figure 1). Figure 1 shows that impaired reaches are most prevalent in the southern Merrimack watershed and consist predominately of headwater (1st order) reaches elsewhere. The largest river exhibiting any impairment is 4th order river draining 211 km^2; however, the largest stream with a median value of impaired days each year above zero over the 17-year record is 2nd order draining 14 km^2. The period mean of IRD, the product of river length exceeding 600 $\mu S\ cm^{-1}$ times duration of impairment, was $1.3 \times 10^5 \pm 4.0 \times 10^4$ [$km\ d$] ($\pm\sigma$ interannual variability). The period mean for OSC (summer-time conductivity at basin mouth) was 216 ± 27 [$\mu S\ cm^{-1}$], which is lower than the 600 [$\mu S\ cm^{-1}$] impairment criteria threshold but higher than pristine conditions.

FrAMES-NACL shows that a threshold of imperviousness leading to chloride impairment exists at less than 20% impervious cover. The annual number of exceedances of the USEPA chronic water quality threshold for chloride (230 $mg\ \mathrm{CL}\ L^{-1}$ for 4 days) predicted by FrAMES-NACL are similar to that predicted from the relationship between chloride exceedances and mean annual chloride concentration identified by Trowbridge et al. (2010, fig. 4) (Figure 3b), when using mean annual chloride from Daley et al. (2009) as the predictor (Figure 3a). Both FrAMES-NACL and the function of Trowbridge et al. (2010) predict negligible USEPA chronic exceedances below 20% impervious cover. Exceedances per year increase linearly through about 60% impervious cover (Figure 3b). Previous findings for streams in New England found changing aquatic community structure at mean impervious cover ranging from 14-18% (Figure 3b) (Coles et al., 2004). Impairment of aquatic taxa observed at impervious cover lower than 20% is consistent with stress from prolonged exposure to chloride at lower concentrations than 230 $mg\ \mathrm{CL}\ L^{-1}$ (Cañedo-Argüelles et al., 2013; Findlay and Kelly, 2011b).

Interannual variability in riverine chloride impairment

Between 1998 and 2014, both metrics of chloride impairment (IRD, OSC) suggest that climate drives variability in impairment (Figure 4). The two impairment metrics correlate with different individual meteorological indices. Even with the damping from groundwater storage the predicted interannual standard deviation of riverine impairment was 12% for IRD and 18% for OSC (Figure 4). IRD was positively correlated with snowfall (WINS) ($r^2 = 0.34, p = 0.015$) and OSC inversely correlated with summer precipitation (SUMP) ($r^2 = 0.50, p = 0.001$) (Figure 4a-d). We hypothesized number of dry days (NDD) should correlate with increased impairment. Though greater NDD led to greater impairment, the relationships were not statistically significant, ($p = 0.075$, and 0.068 for IRD and OSC, respectively) providing only

weak support for our hypothesis (Figure 4c,f). Inverse correlations of OSC with SUMP support summertime dryness as an important factor in controlling productive season salt impairment in larger rivers (Figure 4d).

The positive correlation between IRD and WINS (Figure 4b) results from salt application being driven by frozen precipitation during the winter. Most road salt infiltrates groundwater (f_{hci} is low) and despite the large groundwater exchange pool, shallow groundwater is responsive enough to propagate the effects from the previous winter. Since small streams dominate total network length (Leopold, 1964) the IRD is driven by small stream responses particularly in urban headwaters.

Greater WINS has an inverse effect on OSC, a metric for large river, because it increases dilution from pristine catchments during snowmelt; however, the relationship is not statistically significant ($p = 0.66$; Figure 4e). The weak correlation between WINS (i.e. recent loading) and downstream concentrations (OSC) illustrates the importance of the scale dependent processes of river network dilution on defining concentrations experienced by aquatic organisms.

As SUMP increases, IRD and OSC decline, though only the response of OSC is much stronger (Figure 4a,d). WINS and SUMP have similar effects on the OSC metric because both dilute chloride emanating from the relatively few headwater catchments that are chloride sources. Dilution from high SUMP has a limited effect on the IRD metric because higher than average precipitation cannot substantially alter chloride concentration in a large groundwater storage pool. Instead, SUMP can only increase the time that a contaminated headwater catchment is diluted by storm runoff, thereby lowering IRD in some catchments. Despite the overly dynamic response of FrAMES-NACL (Section 5.1), this influence of SUMP on IRD was not significant (Figure 4a).

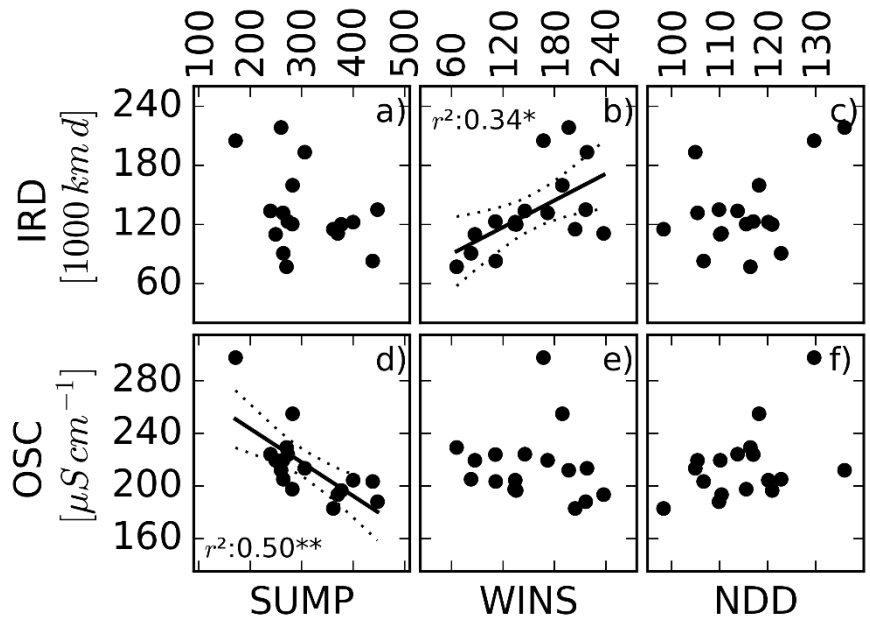


Figure 4: Interannual correlations between impaired reach days (IRD) and outlet summer concentration (OSC) impairment metrics and select meteorological indices, including total summer precipitation (SUMP), winter frozen precipitation (WINS), and number of dry days (NDD). Bold lines represent best fit for relationships, with dotted lines representing 95% confidence intervals on the fit. *,** Significant at the 0.05 and 0.01 probability levels, respectively.

Studies of riverine chloride contamination tend to focus on the trends of annual or seasonal chloride concentration (Anning and Flynn, 2014; Corsi et al., 2015). One study reported 20% interannual variability in mean annual chloride concentrations from subcatchments surrounding Baltimore Maryland (Kaushal et al., 2005 fig. 2), similar to this study (Figure 4). Following significant floods in southeastern New Hampshire during 2006 and 2007, observed chloride concentrations across a broad range of discharge were distinctly lower, suggesting flushing of legacy chloride from the fifth order watershed (Daley et al., 2009). The mechanisms that control the dilution of road salt at impacted reaches should be considered as an important secondary predictor of potential habitat degradation (e.g. Hale et al. 2014). Mechanisms that can influence dilution potential from clean headwaters include drinking and irrigation water abstractions in headwaters, and storage behind recreational dams. If climate patterns change (Hayhoe et al.,

2006), reduced headwater dilution capacity from increasing forest evapotranspiration or drought would exacerbate the effects of chloride contamination.

Limitations of Model Structure

Our immobile zone parameterization for long-tail groundwater transport appears reasonable. Optimized values for the immobile zone exchange mechanism ($\alpha_{MIMT} \approx {}^{1}/_{400\ d}$), suggest a groundwater transit time >1 to 2 years consistent with typical catchment transit times (McDonnell et al., 2010), including catchments of the Merrimack River (Benettin et al., 2015), and shallow groundwater flow-paths (Morgenstern et al., 2010). Some observations of groundwater transit time (e.g. Morgenstern et al., 2010; Stewart et al., 2010) are longer than calibrated here, and other immobile zones would be required to represent the broad range of groundwater transport timescales (Haggerty and Gorelick, 1995).

Greater dilution during rain events (Figure 2) result in underestimates of specific conductance in developed catchments. The greater model responsiveness compared to observations follows primarily from assuming timescales of transport and hydrodynamic response to be equal in the soil and surface flow-paths, though transport should lag hydrodynamic catchment response (Beven, 1982; Kirchner et al., 2000).

Neglected processes may account for FrAMES-NACL's stronger dilution response than observations. Non-conservative chloride transport and processing including organochlorine formation, microbial processing, plant uptake, and storage in sediment micropores would dampen catchment chloride response (Bastviken et al., 2007; Kincaid and Findlay, 2009; Redon et al., 2011, 2012; Shaw et al., 2012; Öberg and Bastviken, 2012), however, these processes are most significant in pristine systems, and diminish in importance as total loading (e.g. via road

salt) increases (Svensson et al., 2010). Prolonged chloride storage on rough road surfaces is mobilized during precipitation through the summer, and first flush specific conductance can locally exceed 50 mg Cl L^{-1} (230 $\mu S\ cm^{-1}$) (Ostendorf et al., 2001), considerably greater than that represented by NACL (typically < 1 mg Cl L^{-1}). Representing a reservoir for chloride storage and release from infrastructure (Ostendorf et al., 2006) is an appropriate improvement for future work. A final possibility is that storm events mobilize more groundwater with elevated chloride than represented by FrAMES-NACL (McDonnell, 1990, 2013; Kirchner et al., 2000). The shallow groundwater pool is represented with an exponential RTD; a gamma distribution (with $\alpha < 1$), characteristic of many environmental systems (Hrachowitz et al., 2010), may better represent rapid mobilization from groundwater with high chloride. Models accounting for these processes can be tested against observational data to evaluate evidence for these mechanisms.

Conclusions from contemporary analyses

We find road salt loading inferred from stream chloride concentrations at regional watershed scales to be consistent with inventoried estimates. A combination of inventory approaches and catchment scale chloride mass balance modeling informed by stream chemistry data is likely to provide more realistic estimates of chloride balance through time. Loading estimates, derived here and from existing inventories, are considerably greater than recommended deicer application rates, offering an opportunity for improved management that can reduce impairment. Based on sensitivity of FrAMES-NACL to interannual variability of winter snowfall, which drives loading via road salt applications, adoption of loading recommendations may reduce chloride impairment within several seasons throughout much of the moderately developed MRW. The variability of the large river chloride is predominately controlled by dilution from

water available from pristine catchments upstream, so management strategies on large rivers need to recognize the role of the entire river network.

Future Scenarios of Chloride Impairment

Changes in snow management, climate, and population influence the rate of loading and flushing of road salt from catchments. Projections of the stress of chloride to aquatic organisms must simultaneously consider the interactions of these changes with their associated timescales. We evaluate a suite of land-cover change and management drivers of aquatic habitat change within the context of key uncertainties in groundwater storage, legacy road salt application, and future climate. We specifically ask whether transitioning to recommended rates of road salt application (Environment Canada, 2004; Salt Institute, 2007) would improve aquatic habitat within the context of growing population and commensurate increase in impervious surfaces. As an additional hypothesis, we ask whether maintaining status quo rates of salt application would exacerbate the decline in aquatic habitat even if all new development followed an in-filling paradigm on existing developed areas. Examining these opposing drivers in greater detail can inform how management of impervious surfaces for winter travel may affect the viability of lotic and lentic ecosystems in developed regions.

Methods for future scenario experiments

Future scenarios focused on the Upper Merrimack River Watershed (UMRW) defined as the river upstream and including the confluence of the Merrimack and Nashua Rivers in Nashua, New Hampshire USA. The two rivers drain 10,545 km^2 in New Hampshire and Massachusetts. The study region is described in Samal et al. (2017). Moderate development throughout the watershed is concentrated along the largest rivers and increases near the outlet. At present

approximately 3.1% of the watershed is covered by impervious surfaces that potentially receive deicing treatment.

To evaluate the roles of build-out pattern and deicer application rate in the context of future aquatic impairment, a suite of scenarios considered these factors and potential uncertainties associated with future climate projections (Table 1). Future simulations followed the analysis of Samal et al. (2017) using land-cover scenarios developed for New Hampshire by Thorn et al. (2017). We analyzed pair-wise differences between estimates of impairment for the last two decades of the 21st century between combinations of the scenarios analyzed. We pooled estimates of impairment for simulations derived from each downscaled climate model and compared the distributions of impairment across build-out scenarios and deicer application rates. The results were strongly dependent on winter weather embodied by the two carbon emission pathways (IPCC, 2000) with a high emissions track (A1FI) leading to significantly warmer winters at the end of the century, and a hard reduction in emissions (B1) leading to modestly warmer winters than present-day. Therefore, significant differences in road salt application were projected for each emissions pathway, which are largely imposed as an external force on the UMRW. Therefore, analyses for each pathway were performed separately. Uncertainties internal to the analysis are defined by the Global Climate Model (GCM) generating the coarse climate projections that were downscaled to produce data for this analysis; either GFDL (Dunne et al., 2012) or CCSM (Stern, 2012). Climate data were downscaled using Localized Constructed Analogs algorithm (Pierce et al., 2014). Three land-cover and population change scenarios were simulated following Thorn et al. (2017); Linear Trends assumed constant growth and build-out patterns as observed in the region over the recent past, whereas the Community and Backyard scenarios assumed greater population increases with urban infilling and sprawling

development patterns, respectively. The Community scenario assumed all new population growth and impervious area was concentrated in existing developed area and the Backyard scenario assumed zoning regulations favored larger building lots leading to greater impervious surface cover (Thorn et al., 2017). We compared the distribution of impairment between each sample and current estimates of impairment using the Tukey range test (Tukey, 1949) testing for differences between group means at an alpha value of 0.05.

Table 1: Dimensions of variates defining scenarios to characterize future chloride impairment.

Emissions Pathway	Land Cover	Road Salt	Climate uncertainty
High (A1FI) Low (B1)	Linear Community Backyard	Status Quo Recommended (2015 transition)	GFDL CCSM

Results of future scenario experiments

The two downscaled climate inputs resulted in divergent trajectories of deicer application. Under the Linear trends build out scenario, with status quo deicer application rates (6.5 g Cl mm^{-1} m^{-2}), deicer inputs averaged approximately 200 Gg Cl y^{-1} with GFDL climate input following the low emissions pathway (B1), whereas deicer application averaged 480 Gg y^{-1} using the CCSM data as input due to the greater frozen precipitation projected by that climate model for New England (Figure 5). Deicer flux predicted using downscaled inputs from GFDL and CCSM under the high emissions pathway were more similar and much lower (averaging 140 and 160 Gg y^{-1} after 2050 for GFDL and CCSM, respectively) as the models predict far less frozen precipitation late in the 21st century (Figure 6). These ranges of deicer application fluxes were

incorporated into the ranges of FrAMES-NACL projections of impairment by deicer application rate, and build-out scenario.

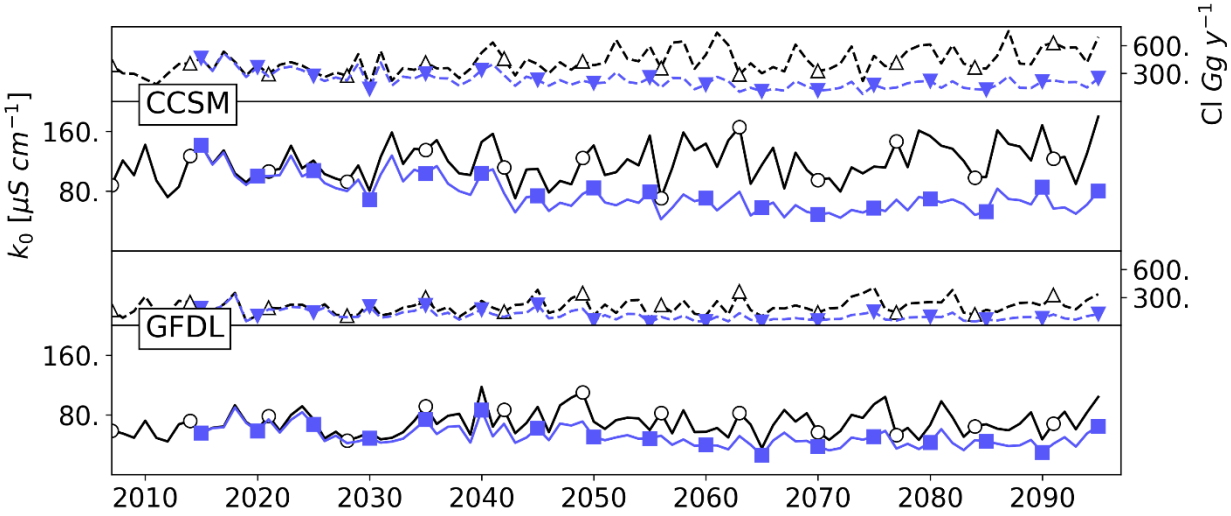


Figure 5: Time-series of annual deicer for the high emissions pathway, and current trends buildout assumptions. Status quo loading in black, recommended loading in blue.

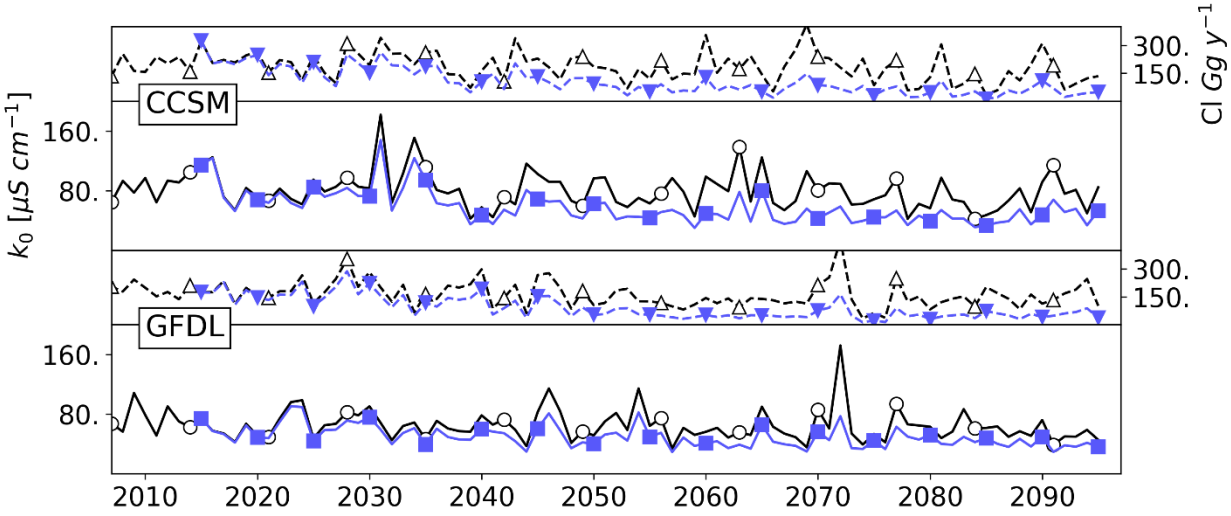


Figure 6: Time-series of annual deicer for the moderate emissions pathway, and current trends buildout assumptions. Status quo loading in black, recommended loading in blue.

Impervious surface area change through buildout is another primary driver of changing chloride load from road deicer. The Backyard build out scenario predicted greater impervious area over the domain (increasing from 330 km^2 to 700 km^2 or from 3.1% to 6.6% of the basin). This differs considerably from the Community trajectory where impervious areas decrease to 300 km^2 or 2.8% of the basin as farmland replaced developed land late in the century. For reference, the comparable amount of impervious area predicted for the Linear trends scenario is 452 km^2 or 4.2% of the catchment area.

Three dimensions were critical for defining the potential changes for chloride impairment of streams (Figures 7 and 8). First, chloride impairment was higher under the moderate emissions pathway (Figure 7) due to maintenance of cooler winter temperatures throughout the 21st century. However, the lower chloride impairment resulting from a higher emissions pathway (Figure 8) was partially offset by a greater impairment associated with increasing water temperature. The impairment metric selected for temperature follows Stewart et al. (2013) and indicates a lower abundance of cool-water fish species. It should be noted that the two impairment metrics measure slightly different aspects of aquatic habitat quality, and a direct comparison between impairment duration-length between the two metrics was unwarranted; a total length of impairment is not presented. The impairment metric selected reflects potential declines in fish abundance and close to a level that affected zooplankton (Figure 9). Recent findings suggest that prolonged chloride concentrations just below the threshold selected here result in 10% reduction in crustacean and rotifer abundance in lakes, along with declines in taxa richness (Hébert et al., 2022).

Examination of chloride impairment under the low emissions pathway (Figure 7) suggested that a transition to recommended deicer application rate would be sufficient to maintain today's level

of reach impairment in the UMRW with any buildout strategy. With present-day deicer application rates, linear trends in buildout (Linear), or dispersed buildout (Backyard), chloride impairment was expected to increase from present day levels. With present-day application rates, impairment under the infilled buildout (Community) was indistinguishable from either current levels or the increase associated with Linear trends. Maintaining recent application rates and the Backyard buildout pattern led to more than 400% increase in chloride impairment. To contextualize these findings, following the publication of the Special Report on Emissions Scenarios by the Intergovernmental Panel on Climate Change (IPCC, 2000), it appears that global carbon emissions were most consistent with A1FI (Allison et al., 2009), the highest rate of emissions considered.

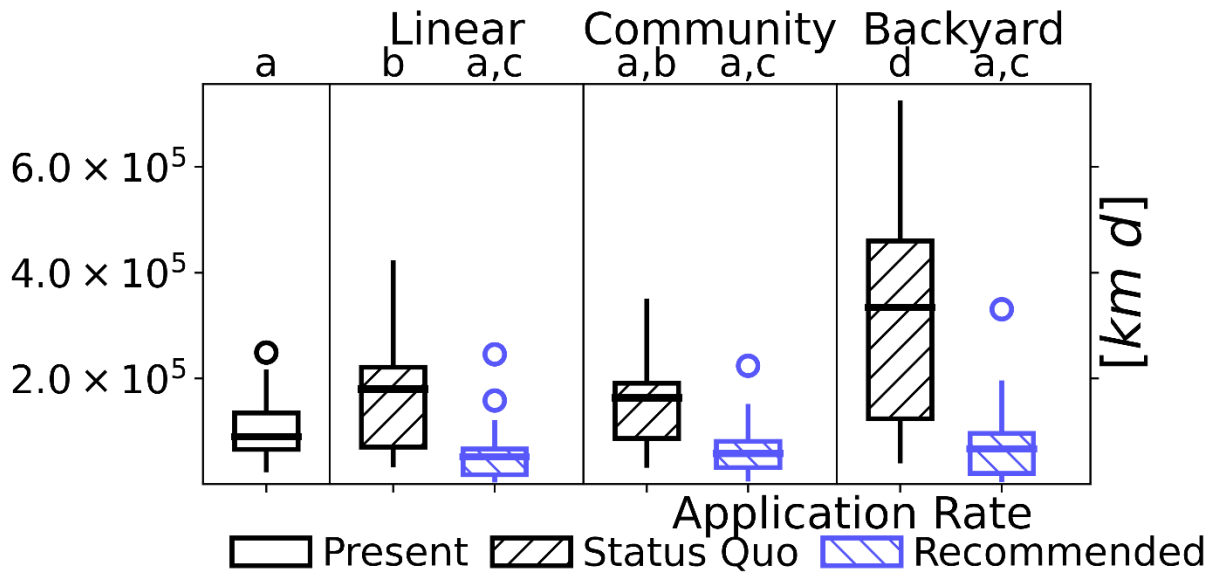


Figure 7: Chloride impairment for scenarios that followed the moderate emissions pathway (B1) measured as the reach length times the average numbers of days per year from 2080 to 2100 that stream chloride concentration exceeded 140 mg L^{-1}.

Chloride impairment projected following the high emissions pathway was generally lower than present day (Figure 8). Under the current deicer application rate, impairment was lower when

assuming linear increase in population and buildout (Linear). Impairment for the Community or Backyard buildout scenarios was not different from current impairment or the lower impairment level of Linear. If recommended deicer application rates were adopted; however, chloride impairment was projected to decline for each buildout scenario. Counter-intuitively, the highest rate of impairment for the high emissions pathway, and recommended deicer application rates was for the Community buildout scenario, despite this scenario representing the most aggressive suite of policy adoptions for environmental protection. This scenario has the highest concentration of impervious surfaces, which were sufficient to yield greater impervious despite less intense application of deicer per application area. Still, impairment was projected to decline to approximately half of present-day levels. Again, this would be paired with a substantial warming of stream water, expected to have significant adverse effects on aquatic habitat (Samal et al., 2017, Figure 10).

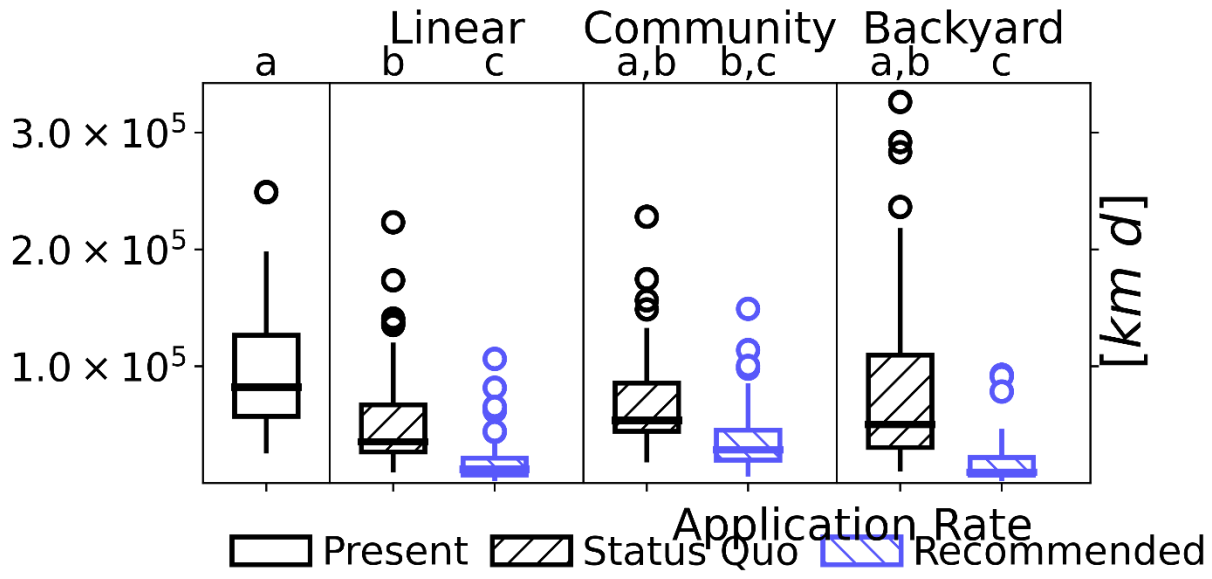


Figure 8: Chloride impairment for scenarios that followed the high emissions pathway (A1FI) measured as the reach length times the average numbers of days per year from 2080 to 2100 that stream chloride concentration exceeded 140 mg L^{-1}.

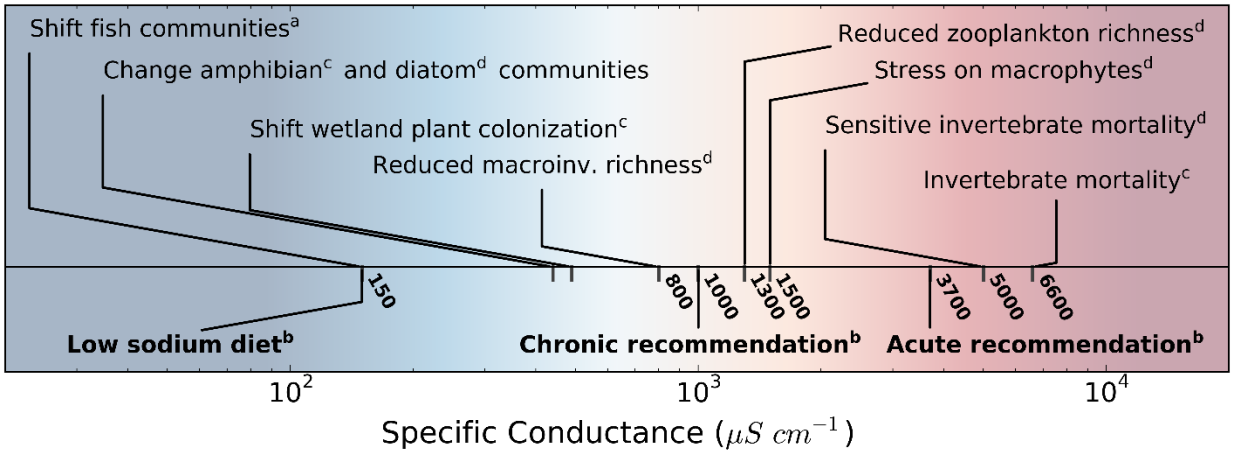


Figure 9: Observed ecosystem effects associated with chloride concentrations (expressed as specific conductance according to Equation 1), or regulatory recommendations for chloride concentration in the environment. [a] (Morgan et al., 2012) [b] (USEPA, 2012) [c] (Findlay and Kelly, 2011a) [d] (Cañedo-Argüelles et al., 2013).

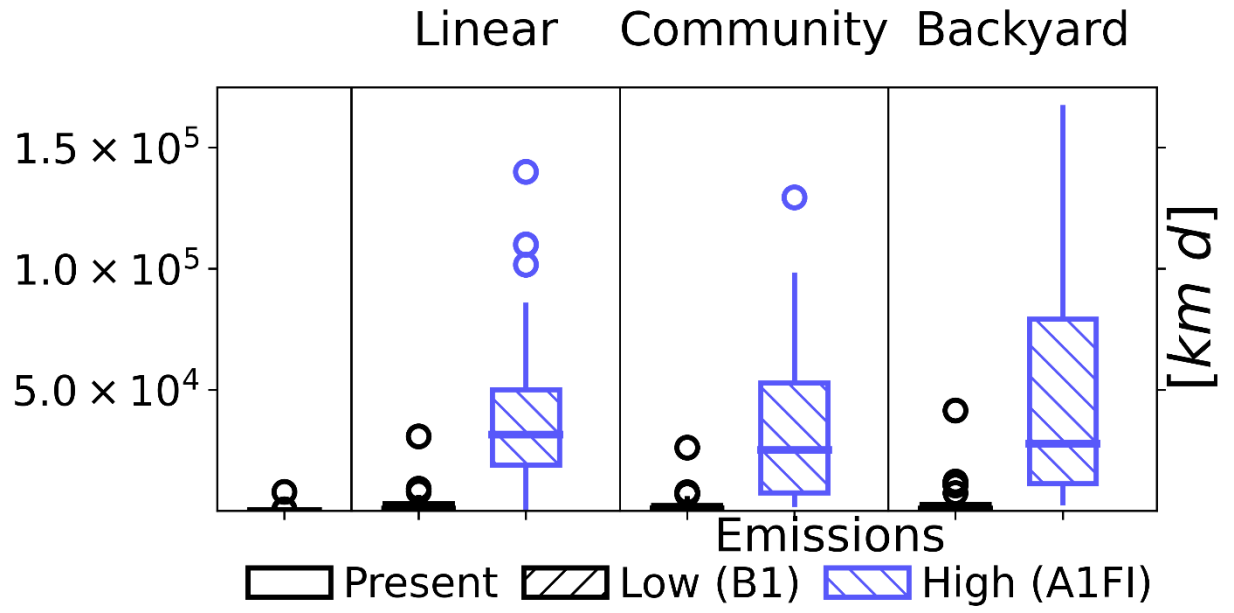


Figure 10: Temperature impairment between high and low emissions pathways based on a seven-day average water temperature exceeding 29°C protective of cool water fish (Eaton and Scheller, 1996).

Discussion of future scenarios experiments

Regardless of emissions pathway, transitioning to recommended rates of deicer application was less likely to be associated with an increase in stream chloride impairment. Maintaining current deicing rates could lead to increasing impairment in some scenarios. Therefore, the scenarios performed here suggest that recommended application rates are a more relevant intervention for chloride impairment than buildout paradigm. However, if deicer application rates remained constant, impairment was consistent with current levels if development followed an infilling buildout paradigm (Community scenario) under either emissions pathway. The emissions pathways used to project future climate in this analysis are not associated with any probability, so no assessment of likelihood of chloride projections represented either pathway (Figures 3 and 4) was conducted. Outside of chloride impairment, numerous water quality and environmental service benefits suggest management policies that target the Community scenario (Borsuk et al., 2019).

CHAPTER II

INTERPLAY OF CHANGING IRRIGATION TECHNOLOGIES AND WATER REUSE: EXAMPLE FROM THE UPPER SNAKE RIVER BASIN, IDAHO, USA

Introduction

Access to irrigation water is critical to determining the future resiliency of many agricultural systems (Foley et al., 2011), and challenges of providing irrigation water require close scrutiny of its efficient use (Grafton et al., 2018). The goal of resilient agricultural systems should reflect a global need to reduce water scarcity (Rosa, 2017), with adaptations that are often context specific (Vanham et al., 2018). Successful management of water resources to protect against water scarcity requires consideration of the specific interactions of multiple actors (Keller and Keller, 1995).

One suite of solutions where water is scarce is to modernize irrigation technology to ensure that the greatest proportion of supplied water is used for beneficial crop growth (Gleick et al., 2011; Jägermeyr et al., 2015, 2016). Improving classical irrigation efficiency (CIE), the ratio of beneficial consumptive use to gross irrigation abstractions, is critical to meet agricultural production needs (Jägermeyr et al., 2016) and additionally has myriad co-benefits such as reduced energy use and improved water quality (Gleick et al., 2011; Vanham et al., 2018). However, more efficient irrigation systems tend to counterintuitively increase total water consumed, or at least do not decrease use to the degree expected. As efficiency increases, usually at a cost to the irrigator, the water available from reduced losses can be applied for higher value and more water intensive crops (rebound) or for expanding crop areas (slippage) (Contor and Taylor, 2013; Grafton et al., 2018; Pfeiffer and Lin, 2014; Tran et al., 2019), especially when users are encouraged to extract a full water allotment by legal doctrines such as the Prior Appropriations system used in the US West. Increasing CIE also tends to reduce non-consumptive losses that downstream users rely on (Foster and van Steenbergen, 2011; Frederiksen and Allen, 2011; Grafton et al., 2018; Grogan et al., 2017; Simons et al., 2015).

Non-consumed losses, the fraction of water applied by irrigation that flows back to the landscape, follow different pathways. The term *irrigation returns* can refer to flow or flow structures conveying non-consumed water off irrigated fields and back to a canal system (e.g. Lin Y. and Garcia L. A., 2012), percolation back to a source aquifer (e.g. Dewandel et al., 2008), or more generally to all water not consumed by irrigation water application or delivery (Grogan et al., 2017; Keller and Keller, 1995; Simons et al., 2015); we adopt the latter meaning when referring to *irrigation* or *incidental returns*.

Investigators of water resources argue that the reuse of incidental returns increases the basin or global efficiency of supplied water, making technological investments that increase CIE less effective when considered at basin-scales rather than at farm-scales (Keller et al., 1996). The effect has been observed empirically in well-studied basins (Simons et al., 2015). Increasing CIE is almost certainly a critical component to maintain the resiliency of agricultural systems when only surface or groundwater supplies irrigation and will necessarily reduce the incidental return to the system. In settings that conjunctively use both surface and groundwater resources, managed aquifer recharge (MAR) can increases the adaptability and resiliency of irrigated agriculture (Dillon et al., 2020). MAR adds water to aquifer storage when available, eliminates the need for infrastructure associated with surface reservoirs, minimizes surface evaporation, and can be less expensive than surface storage (Arshad et al., 2014; Dillon, 2005; Dillon et al., 2019; Maliva, 2014; Scanlon et al., 2016). MAR as part of a conjunctively managed water resource system has been demonstrated to maintain water supplies for irrigated agriculture during drought (Foster and van Steenbergen, 2011; Guyennon et al., 2017; Niswonger et al., 2017; Scanlon et al., 2016; Tran et al., 2020). However, water used for MAR tends to reduce flow leaving a catchment (Yaraghi et al., 2019), which may have important downstream consequences. In other

cases, MAR may affect annual flows slightly (e.g. Niswonger et al., 2017), but can shift timing of baseflow entering rivers from the aquifer to summer months, providing important temperature refugia for aquatic species (Van Kirk et al., 2020).

Despite the potential benefits from coupling MAR with conjunctively managed water sources, there remain challenges in uptake of the practice (Dillon et al., 2020) to address globally declining aquifer storage (Bierkens and Wada, 2019). Outside specific regulatory intervention, the practice of MAR can marginally reduce the cost to pump groundwater such that MAR would be expected to result in rebound and slippage effects (Tran et al., 2019) where more land is planted, or more water intensive crops are grown to utilize the available water. Benefits from the conjunctive management of water resources and MAR are projected to be greater in arid environments (Scanlon et al., 2016).

The two interventions presented above, increasing the efficiency of irrigation through technological modernization and MAR appear to synergistically alleviate the drawbacks of each practice. In the absence of slippage, increasing CIE can reduce incidental recharge (Simons et al., 2015), but retains greater flow within the river, whereas MAR increases recharge but reduces annual river flow (Yaraghi et al., 2019). Balancing the two interventions could potentially achieve greater resiliency of irrigated agriculture than either alone. To date there have been limited analyses to include both strategies in the same framework. Tran et al. (2019) account for specific efficient irrigation practices within the context of multiple potential drivers in an hydro-economic analysis. Other examples of mechanistic models applied to the problem of MAR and conjunctive resource management have assumed static efficiencies of irrigation technology and crops (Niswonger et al., 2017; Scherberg et al., 2014). Here we consider the coupled influences of irrigation technology modernization and MAR on water resources to assess the limits to which

either intervention could achieve aquifer stabilization, while maintaining downstream flows above critical thresholds.

We quantify the impact of changing irrigation efficiency on basin water stocks, aquifer recharge, downstream discharge, and within-basin water reuse using the test case of the Upper Snake River Basin (USRB) of Idaho, USA, an intensive agricultural setting in the semi-arid American west (Figure 11) that relies on both surface and groundwater. Arid and semi-arid agriculture can be very important economically; 31-36% of the nation's net farm income is produced in arid or semi-arid regions (Trabucco and Zomer, 2019; USDA NASS, 2014). Historic flood irrigation with river water elevated aquifer head above pre-irrigation levels, and in the latter half of 20th century aquifer storage has declined (Kjelstrom, 1995) due to increasing groundwater pumping and decreasing recharge of surface water as irrigation technology has modernized. Therefore, aquifer stabilization is critical for establishing resilience of the agricultural system (IWRB, 2016). Implementation of recharge (as MAR) and other management actions to ensure a resilient agricultural system in the USRB may provide important insights relevant throughout arid and semi-arid regions where surface evaporative fluxes are similarly high (Carr et al., 2010; Ghassemi et al., 1995; Tal, 2016).

Careful application of improved irrigation efficiency and MAR has been part of on-going management strategies in the USRB. Water is governed in Idaho under the Doctrine of Prior Appropriation, which allocates water to users according to the date when they first put water to continuous beneficial use. As aquifer drawdown has continued, aquifer recharge, water transfers, and other water conservation efforts have been classified as beneficial uses (Fereday et al., 2018). Water users have self-organized in an effort to stabilize the aquifer and optimize use of the water within the USRB, and have moved to a more conjunctive management of surface

and groundwater resources (Gilmore, 2019). A moratorium on ground water permits (Higginson, 1992) and conservation efforts have resulted in a reduction in consumptive ground water use, and motivated the adoption of targets for 0.3 km^3 y^{-1} of MAR as an intervention (IWRB, 2016); however, aquifer storage may still be declining. Simultaneously, maintaining sufficient downstream flow from the basin is strictly required for senior water rights holders and hydropower generation (IWRB, 1985). The USRB is an ideal setting to assess the trade-offs between within-basin aquifer storage and downstream supply through conjunctive management.

In this study, we frame a series of model parameterizations together to test hypotheses guided by the key constraints of water resource management in the USRB. We utilized a distributed model of hydrologic function and human water use to estimate the recharge required to a) stabilize the aquifer under present-day irrigation efficiencies, and b) offset reduced irrigation returns from continued modernization of irrigation technology. We performed simulations introducing progressively more efficient irrigation technology to a baseline representation of the USRB, which required reduced withdrawals from the Snake River, but which hastened aquifer drawdown by decreasing recharge of incidental returns. These simulations were paired with counterparts introducing enough managed aquifer recharge to ensure negligible change in aquifer storage (stabilization) over the same period. We hypothesized that only a fraction of the reduced incidental returns from modernizing technology would be needed to maintain aquifer volume if introduced as MAR. An alternative hypothesis is that asynchronicity in recharge water availability and irrigation demand, coupled with fast flow through the aquifer system, would require greater recharge rates than if water was introduced as inefficient irrigation and reused contemporaneously. For each simulation we calculated the total amount of previous incidental returns reused as gross irrigation, using the model's core capability of tracking water sources

through all pools of the hydrologic cycle. We hypothesized that simulations with additional MAR would exhibit lower reuse than simulations without MAR because a greater proportion of recent snowmelt would recharge the regional aquifer. Alternatively, additional MAR reduces surface water availability and may promote groundwater abstractions that would favour greater reuse as most irrigation returns percolate to recharge the aquifer.

Methods and Data

The following sections describe the setting, describe the formulation of hydrologic fractions used here, the experiments conducted, the Water Balance Model (WBM) – a distributed hydrologic model representing anthropogenic water uses, model input data, and validation criteria.

Upper Snake River Basin

The Upper Snake River Basin (USRB) is a semi-arid steppe ecosystem with a snowmelt dominated Mediterranean climate in western Wyoming, and southern Idaho, USA (Figure 11). The 92,700 km^2 basin is bounded to the east by the Teton Mountains, and to the north by the Sawtooth and Bitterroot Mountain ranges. Precipitation over the Snake River Plain is generally less than 250 mm/year but averages about 400 mm/year (or 46.3 km^3/year) over the whole basin with most water entering the river network as montane snowmelt. The basin is underlain by Quaternary basalts of the Snake River Group (Whitehead, 1992), which form the highly transmissive Eastern Snake Plain Aquifer (ESPA). Irrigation in the USRB began in the late 1800s and gravity drained flood irrigation was the primary mode of irrigation until the mid-1900s (Wulfhorst and Glenn, 2002; Lovin, 2002). Incidental recharge from the non-consumptive losses of irrigation water increased storage in the ESPA, and increased discharge from a dense collection of springs in the Snake River canyon between Milner and King Hill, Idaho (Kjelstrom, 1995). Through the latter half of the 20th century, aquifer head declined due to

increasing reliance on groundwater for irrigation and reduced incidental recharge (Moreland, 1976) as flood-irrigated land transitioned to sprinklers. Aquifer stabilization at today's head is a primary concern in the basin, even though head is above pre-irrigation levels. State agencies are practicing managed aquifer recharge (MAR), the deliberate infiltration of seasonally available water for use throughout the year, as one technique in the conjunctive management of water resources (IWRB, 2009, 2016).

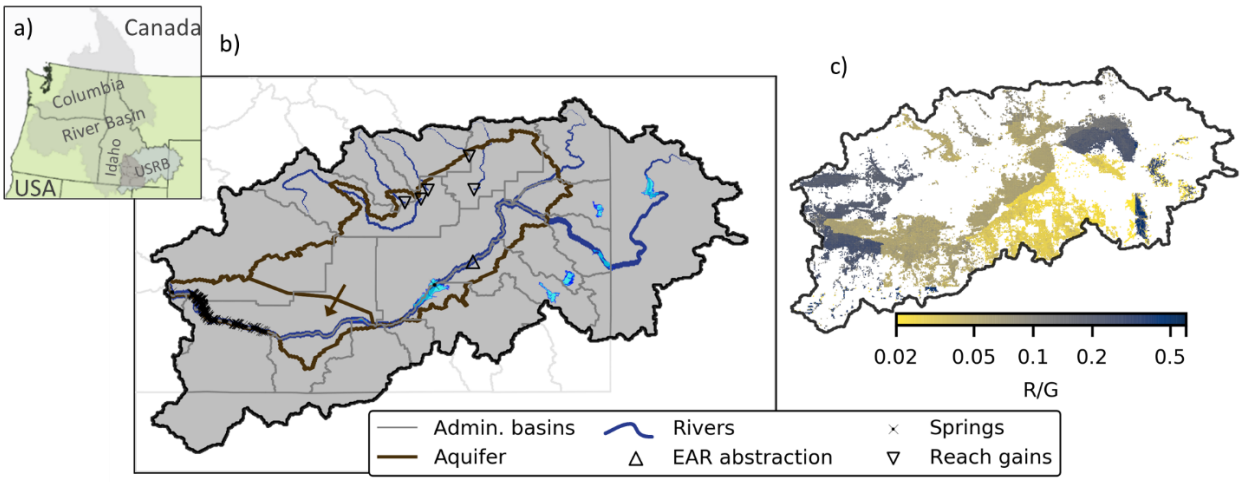


Figure 11: Location of Upper Snake River Basin (USRB) a headwater of the Columbia River in the US States of Idaho, and Wyoming (a). Configuration of major hydrologic features of USRB, including the two compartments conceptualized for the Eastern Snake Plain Aquifer (ESPA), locations of reach gains where losing rivers drain to the ESPA, the location of simulated abstractions for enhanced aquifer recharge (EAR), springs where flow from the ESPA drains back to the Upper Snake River, the river network scaled by mean annual flow, and extents of administrative basins (Administrative Basins, 2018) indicating areas using common surface water sources for irrigation (b). Average fraction of gross irrigation comprised of incidental returns (irrigation reuse - *R*) (c).

Groundwater age dating and geochemical analysis established that the downgradient portions of the aquifer consist of between 60 and 80% of water used for irrigation and derived from the Snake River (Lindholm, 1996; Plummer et al., 2000). A highly managed network of reservoirs

and canals convey about 12 km^3 y^{-1} of water to croplands (Maupin et al., 2014), equivalent to about 25% of annual precipitation to the basin. At least 5.5 km^3 of water is stored in the three largest reservoirs alone. An additional 2.5 km^3 y^{-1} is abstracted from the ESPA by irrigators (Dieter et al., 2018; Maupin et al., 2014), and approximately 5 km^3 y^{-1} of water returns from the ESPA to the Snake River through a series of springs (Covington and Weaver, 1991; Kjelstrom, 1995). Inflows to the ESPA include several losing rivers at the northern extent of the USRB and the Snake River, which loses water directly to the ESPA near American Falls reservoir (Lindholm, 1996; McVay, 2015). Spring flows out of the ESPA are critical for maintaining an aquaculture industry along the Snake River canyon and constitute a majority of Snake River discharge out of the USRB supporting critical aquatic habitats, hydroelectric generation potential, and irrigation of downstream agriculture. Water available from the Upper Snake River and the ESPA irrigate numerous agricultural products with dairy forage, beet sugar, and potato being the most economically-important (USDA NASS, 2014).

Hydrologic fractions and irrigation resource use

Defining efficiency of agricultural water use is complicated because water lost non-productively by one water user may be used productively elsewhere downstream in the basin, making terms describing efficiency or resource sufficiency specific to the spatial scale considered. We describe irrigation efficiency using hydrologic fractions that describe the fate of water abstracted from either surface or groundwater sources for the purpose of irrigation (Frederiksen and Allen, 2011; Haie and Keller, 2008; Lankford, 2012; Perry, 2011). Water abstracted as gross irrigation (G) can have three fates when added to irrigated pixels at the plot-scale: i) beneficial use (B) is the irrigation water used for beneficial crop growth; ii) non-beneficial consumption (N) is water evaporated non-beneficially from soil or canals; or iii) non-consumptive loss (L, herein

incidental returns or *incidental recharge*) is runoff or percolation as a liquid that remains in the basin (Figure 12). Of these plot-scale fates of gross irrigation water, B and N are both assumed terminal because liquid water leaves the domain as vapor or in crops. Incidental returns (L) on the other hand remain in the system and fates at the basin scale include export (X) via streamflow at the basin outlet, evaporation (E) from the surface water network, human use (U), reuse as gross irrigation (R), and net storage (S) primarily in the aquifer; however net storage in surface reservoirs, and soil is also calculated.

WBM tracks key component volumes, including incidental returns (L), to all terrestrial compartments of the hydrologic system permitting direct computation of gross irrigation water reuse (R). In our analysis we assume that all incidental returns are recoverable, and therefore do not make a distinction between recoverable and non-recoverable returns (as in Lankford, 2012), and directly assess the volumes of water *recovered* in gross abstractions. Gross irrigation reuse (R) is the weighted sum of abstractions consisting of incidental return in each source and is calculated daily Equation 2.

$$R_{i,j} = I_{Aqf} \cdot f^{irr}_{Aqf_{i,j}} + I_{Rsvr} \cdot f^{irr}_{Rsvr_{k,l}} \quad \text{Eq. 2}$$

where i and j are row and column indices for the point of irrigation water application; I_{Aqf} and I_{Rsvr} are the abstracted irrigation water from aquifer, and surface reservoirs, respectively; f^{irr}_{Aqf} and f^{irr}_{Rsvr} are the fraction of irrigation return flow water in aquifer and reservoir water, respectively; and k and l are row and column indices for the pixel of the surface supply reservoir. The metric was summed spatially and temporally and compared to total gross irrigation (G) to calculate a ratio of irrigation water reused within the USRB. Irrigation efficiency is calculated as classical irrigation efficiency (CIE) given by Equation 3, and effective

irrigation efficiency (EIE) following the quantity (Type N) model of Haie and Keller (2008) given by Equation 4.

$$CIE = \frac{B}{G} \quad \text{Eq. 3}$$

$$EIE = \frac{B}{G-R} \quad \text{Eq. 4}$$

Note, R does not quantify how many times a given parcel, or on average all irrigation water, is reused, as in the distinct definition of $\boldsymbol{R}$ as the index of unsustainable groundwater reuse in Grogan *et al.* (2017). Rather, it identifies what portion of total irrigation water has been through cycles of use (Figure 12).

Experiment structure

There is strong connection between the Upper Snake River and ESPA in the USRB, both through reach gains and sinks from the Snake River to the ESPA and from springs back to the Snake River. These connections are not unlike alluvial aquifers where conjunctive management of water resources is most common (Foster and van Steenbergen, 2011). We therefore use predictive inference (Ferraro et al., 2019) to assess the potential for trade-offs between downstream flow and aquifer drawdown as irrigation efficiency and MAR change independently. We should note that the experiments we perform potentially violate water law and precedent in the basin (Gilmore, 2019), so natural experiments (Penny et al., 2020) to interrogate similar processes are impractical. To test our hypothesis that only a fraction of reduced incidental recharge is needed as managed aquifer recharge (MAR) to increase water availability basin-wide, we simulate a suite of alternative model parameterizations to capture increasing irrigation efficiency (as CIE) paired with and without MAR. In WBM, we introduce a fraction of daily flow from the Snake River immediately above the American Falls Reservoir directly to the

ESPA to represent recharge as an intervention. Because our simulations also reflect changes in aquifer recharge related to changing flow in the river source, we use the term enhanced aquifer recharge (EAR) to refer to all induced changes in aquifer recharge in our model simulations. Specific changes to simulated irrigation technologies for each parameterization are described below. We then assess our hypothesis by calculating a management benefit (MB) metric that compares the change in net incidental recharge to the change in EAR required for aquifer stabilization by difference for each parameterization. The MB is the magnitude by which the increase in required EAR is less than the loss in net incidental recharge and calculated by Equation 5.

$$MB = (I_{rch}^{*} - I_{rch}) - (EAR - EAR^{*}) - \frac{dV_{ESPA}}{dt} \quad \text{Eq. 5}$$

where I_{rch} is net incidental recharge (incidental recharge minus groundwater abstraction), EAR is the enhanced aquifer recharge flux, and * represents the flux at the present-day baseline. For each parameterization, we compare the change in aquifer storage with the relative change in discharge from baseline to evaluate the combination of aquifer and streamflow capture needed to support irrigation abstraction at a given level of efficiency. In this manuscript, our definition of streamflow capture is general, any decreasing discharge out of the basin due to altered management practice, and does not specifically mean the change in streamflow and recharge resulting from increased groundwater pumping (Konikow and Leake, 2014).

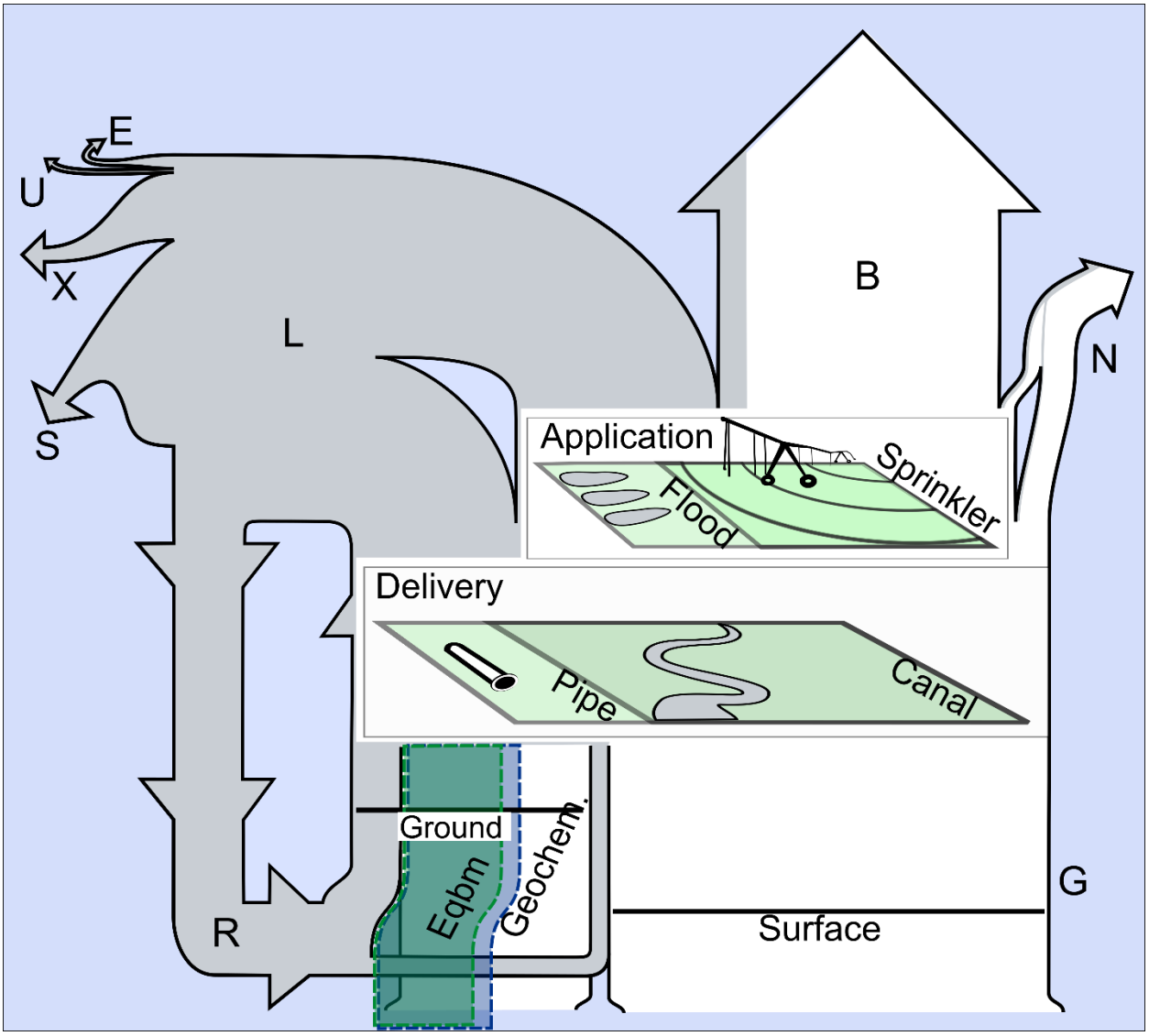


Figure 12: Diagram of fates of water abstracted for irrigation across the USRB. Flow-line widths are scaled proportional to fluxes across the simulation domain between 2008 and 2017 at the baseline parameterization. White depicts abstractions from pristine sources, whereas water lost non-consumptively from irrigation delivery or application during the model epoch is grey. Equilibrium (Eqbm) and geochemical (Geochem.) fractions of groundwater abstractions relax assumptions about aquifer water composition and are discussed below. Labels of irrigation fluxes are: G – gross irrigation abstractions, B – beneficial consumption by crops, N – non-beneficial consumption, L – non-consumptive losses or incidental returns, and the remaining fluxes refer to the fate of incidental returns: R – reuse in gross irrigation, E – evaporation, U – human use, X – export, and S – storage in aquifer, soils, and reservoirs.

For each simulated suite of irrigation technology parameters, we run paired simulations with and without EAR. For EAR simulations we target aquifer stabilization defined as a long-term (e.g. decadal) average change in groundwater storage of the entire ESPA close to zero during the contemporary time-period from 2008 through 2017 Equation 6.

$$\frac{dV_{ESPA}}{dt} \sim 0: \; -0.1 < \frac{dV_{ESPA}}{dt} < 0.1 \; [\text{km}^3 \; \text{y}^{-1}] \qquad \text{Eq. 6}$$

Once values for ESPA exchange were calculated for the baseline representation, simulations were conducted with these values for each of the nine more efficient irrigation technology parameterizations (Table 2). Then, additional EAR was estimated through manual calibration to achieve a stabilization of ESPA volume for the baseline and each of the efficiency parameterizations. For all model simulations, aquifer stabilization, basin discharge, and hydrologic fractions including reuse were calculated from hydrologic model output. In calculating MB, the change in aquifer volume is subtracted to account for small deviations from aquifer stability that remain after calibration.

Water Balance Model

We used the University of New Hampshire Water Balance Model (WBM) to characterize water balance and assess water resource fates (Vörösmarty et al., 1989; Wisser et al., 2010). WBM is a distributed hydrologic model utilizing conceptual soil, surface runoff, and shallow groundwater pools, a one-dimensional river network utilizing hydrologic routing schemes, and representations of human controls on the hydrologic cycle such as dams, impervious surfaces, irrigation, livestock, industrial, and domestic water use. WBM tracks specific components of water fluxes, notably irrigation returns, through each represented pool assuming each pool is well-mixed at each daily time-step (Grogan et al., 2017).

Several modifications were implemented in WBM for the present work; a more complete description of the fundamental WBM model structure is available elsewhere (Grogan, 2016; Grogan et al., 2017; Wisser et al., 2010). In previous applications of water tracking in WBM (Grogan et al., 2017), component stocks were adequately cycled as representative of the various components following model spin-up. To address concerns that water components retain a memory of assumptions at initialization owing to new groundwater representation described below, all stored water at model initialization was tracked as *relict* water, a measure of water remaining in the system prior to the dynamic model simulation epoch. We introduced an upper volumetric bound to the surface runoff pool to rectify a low bias in runoff during extreme precipitation and snowmelt events. The fraction of surplus soil water (soil water-content above field capacity) that flows to the shallow groundwater pool (γ,unitless), and it's complement (1-γ), which is directed to the surface runoff pool, are generally about 0.5 and robust in a range from 0.4 to 0.6 (Grogan et al., 2017; Samal et al., 2017; Stewart et al., 2013; Zuidema et al., 2018). Due to the highly permeable geology found along the Eastern Snake Plain, γ was increased to represent high initial infiltration rates common throughout the Eastern Snake Plain (IDWR, 2013). For our simulations, γ was spatially variable (ranging from 0.38 to 0.96, mean = 0.73) based on elevation as a proxy for the extents of the Eastern Snake Plain (Figure AI.1). Other parameters defining the major hydrologic controls were established by work across multiple scales (Grogan et al., 2017; Samal et al., 2017; Stewart et al., 2013; Wisser et al., 2010; Zuidema et al., 2018) and were not calibrated for this application in the USRB.

Several features were added to WBM to implement the experiment. To represent the intense management of USRB water resources, reservoir outflow from the three largest reservoirs were specified; therefore, WBM predicted reservoir volume as a consequence of managed release.

Irrigation technology was revised in WBM to a process-based representation that redistributes inefficient irrigation water via surface runoff flows, groundwater percolation, and evaporation during both delivery and application stages. The system explicitly represented non-beneficial consumption as evaporation of sprinkler mists and evaporation from canal and soil surfaces, using technology specific parameters reflecting county-wide averages from USGS water use statistics (Dieter et al., 2018; Maupin et al., 2014). A representative fraction of 4% of sprinkler applied water is evaporated as mists (Bavi et al., 2009; McLean et al., 2000; Uddin et al., 2010). Further, during the irrigation season, water is assumed to be evaporating at potential rates throughout the canal network. We assume crop ET is required (i.e. beneficial) for both transpiration and salt flushing, but water applied during an irrigation event in excess of daily crop demand wets soil above field capacity. Incidental losses during application followed Jägermeyr et al. (2015) and we used their estimates of the distribution uniformity parameter that prescribed excess water needed to satisfy net irrigation demand based on the type of technology, either drip, sprinkler, or flood. Excess water evaporates (non-beneficially) at the potential rate, and remaining liquid water is returned non-consumptively at the end of the timestep via either percolation, or runoff if vertical hydraulic conductivity is too low. The algorithm describing irrigation water fates is detailed in Appendix II.

We defined surface water sources for each administrative basin in Idaho (Administrative Basins, 2018) to come from one or more reservoirs based on the canal network's distribution (Figure AI.2). All daily surface abstractions for irrigation are made from the pool of source reservoirs providing water to each administrative basin proportional to their available storage. Groundwater abstractions for irrigation of croplands were calculated as the difference in demand not satisfied by surface water sources. A more detailed aquifer representation was needed here

than in previous WBM studies. We simulated the ESPA over the same extents as the ESPAM2.1 model (IDWR, 2013) using a lumped formulation that received distributed recharge from natural and incidental sources and reach gains from specific losing rivers (Figure 11), provided a pool of groundwater available for irrigation, and discharged to a series of 213 springs along the Snake River canyon (Covington and Weaver, 1991). Discharge from springs was head dependent and sub-daily head and outflow were calculated numerically using a third order scheme (Bogacki and Shampine, 1989). We represented the aquifer as upgradient (northeast) and downgradient (southwest) lumped compartments (Figure 11) to reflect two characteristic types of water identified by Plummer *et al.* (2000), old groundwater in the upgradient portion, and young water derived from incidental recharge of Snake River water in the downgradient or southwest portion. Storage parameters were estimated for each section: upgradient specific yield is 0.06 and thickness of the aquifer is 250 m; downgradient specific yield is 0.05 and thickness is 220 m (Garabedian, 1992; Whitehead, 1992; IDWR, 2013). We represented the hydraulic connection between the ESPA and the American Falls Reservoir (Garabedian, 1992; IDWR, 2013) as a drain/spring pair. Additional details of the implementation of the lumped aquifer solution are presented in Appendix II.

Input Data

We used a topological network of the Upper Snake River Basin (USRB) that covered an area of 92,900 km^2 at a spatial resolution of 30-arcseconds (approximately 780-m) based on HydroSHEDS (Lehner et al., 2008) but refined to better represent drainage as mapped by the US Geological Survey's National Hydrography Data (National Hydrography Dataset (NHD), 2019). Reservoir data was derived from the National Inventory of Dams (National inventory of dams, 2020) and updated manually to include additional dams, refine reservoir capacities, remove

secondary structures on reservoirs, and refine the locations and upstream drainage areas. Reservoir outflow came from observed flow data from USGS gaging stations located immediately downstream of three primary irrigation reservoirs: gage 13011000 in Moran, WY below Jackson Reservoir, gage 13032500 in Irwin, ID below Palisades Reservoir, and gage 13077000 in Neeley, ID, below American Falls Reservoir. No data regarding direct abstractions from reservoirs were available from these sources. Additionally, we increased the total capacities represented in WBM of these three reservoirs by 10% to approximate storage of their downstream canal systems. There were 128 dams and corresponding waterbodies in the USRB domain. WBM simulations used gridMET (Abatzoglou, 2013) for contemporary precipitation and temperature and MERRA2 for open water evaporation (Gelaro et al., 2017). We utilized a temperature based evaporation equation (Hamon, 1963) for calculating potential evapotranspiration (PET) and a temperature-index based snow accumulation and melt formulation (Willmott et al., 1985). Human population density, which controls both domestic and industrial water demand, came from SEDAC Gridded Population of the World (CIESIN et al., 2016). WBM simulations used Food and Agricultural Organization (FAO) estimates of livestock density for cattle (Steinfeld et al., 2006) at 5 minute resolution following Wisser et al. (2010). These data compared favorably with USDA National Agricultural Summary Statistics (NASS) for 2005 but exhibit more realistic spatial variability than county-level averages in NASS. Over the USRB domain, NASS livestock density is approximately 2 head/km^2 density representing a low bias of the FAO data of less than 1%. We utilized USDA Soil SURvey GeOgraphic (SSURGO) data to parameterize available water capacity for USRB soils. We specified a rate of 115 mm/day for percolation below land occupied by canals and irrigated lands exceeding saturation following findings from the Idaho Water Resources Board (2016).

WBM adapts FAO's methodology (Allen et al., 1998) to estimate crop water requirements based on reference ET, soil moisture, crop coefficient (k_c) and is detailed in previous work (Wisser et al., 2010; Grogan et al., 2017). Here, we utilized the US Department of Agriculture's Crop Data Layer (CDL) estimates of crop types and land cover at 30 m resolution (Han et al., 2012) after remapping crop groups (Table AII.1). The proportions of irrigation delivery technologies were spatially homogenous and reflected the average lengths of technologies in the USGS National Hydrography Dataset (nhd.usgs.gov). The relative proportions of application technology varied by county following USGS surveys (Maupin et al., 2014; Dieter et al., 2018). To address our first two hypotheses, parameterizations were defined that represent nine progressively more efficient suites of irrigation technology, identified here as parameterizations Eff.A through Eff.I. The nine parameterizations are controlled by the relative fraction of flood irrigation (with corresponding increases in sprinkler area), the relative fraction of drip irrigation (with corresponding decreases in flood and sprinkler area), the fraction of canals (with corresponding increases in pipes), and the percolation factor of canal bottoms (Table 2).

Model validation metrics

Model assessment used a composite objective function that described model-observation misfit across four primary metrics. We compared: 1) monthly flow from the springs draining the ESPA against total gains minus diversions between the Kimberly and King Hill, Idaho USGS gaging stations provided by the IDWR (Sukow, 2011, personal comm.); 2) annual gross and surface water abstractions for irrigation over the USRB aggregated by county for the years 2010 and 2015 and compared to USGS water use statistics (Dieter et al., 2018; Maupin et al., 2014); 3) seasonal river discharge at locations upstream of actively regulated reservoirs at USGS gages

13010065 (Flagg Ranch, Wyoming), 13137500 (Trail Creek, Ketchum, Idaho), and 13039500 (Henry's Fork, Lake, Idaho); and 4) seasonal storage within the actively regulated Snake River reservoirs against data from the US Bureau of Reclamation Hydromet database. A standard suite of statistics were used to assess each of these metrics, and we report percent bias and Nash-Sutcliffe efficiency (NSE) for the period between 2008 and 2017 (Table AII.2). Manual parameter calibration established reasonable estimates for the water exchange between the Snake River and ESPA near American Falls. Exchange between the Snake River and ESPA affects reservoir volume estimates, aquifer volume (and therefore spring flows), and can affect surface irrigation estimates as abstractions are necessarily curtailed if American Falls reservoir does not have sufficient storage to meet demand. Therefore, focusing on all four metrics to establish performance was necessary.

Table 2: Definition of efficiency parameterizations.

Parameterization	Fraction of baseline surface irrigation	Fraction of open-surface conveyances	Fraction of Unlined canals	Fraction of direct irrigation
Baseline	1.0	0.83	1.0	0.00
Eff.A	0.89	0.82	0.95	0.10
Eff.B	0.86	0.80	0.8	0.10
Eff.C	0.82	0.70	0.60	0.10
Eff.D	0.77	0.60	0.40	0.10
Eff.E	0.73	0.50	0.20	0.10
Eff.F	0.53	0.40	0.15	0.33
Eff.G	0.40	0.30	0.10	0.50
Eff.H	0.27	0.15	0.05	0.67
Eff.I	0.0	0.0	0.0	1.00

Results

Model Validation

Though most processes within the model were uncalibrated, WBM accurately represented observations of the key fluxes in the USRB that we tested. The spatial distribution of abstractions is accurate for both total and surface sources of irrigation water (Figure AI.3). Mean annual discharge from springs draining the ESPA is unbiased (Figure AI.4). Interannual variability in peak runoff is generally well captured (Figure AI.5a), and timing of peak runoff generation from snowmelt is accurate; however, the onset of snowmelt tends towards an early bias in most years. Though discharge from the headwaters of the Upper Snake River in the Teton Mountains was well characterized (e.g. monthly discharge NSE 0.72 and bias of 13% at Flagg Ranch, WY – Figure AI.5a), the intense management of reservoirs in the USRB results in cascading errors in simulating the timing, rates, and magnitudes of reservoir drawdown (Figure AI.5). Representing the hydrology of heavily managed basins, such as the USRB where most large reservoirs are managed as a single system (not just three reservoirs where we forced outflows to observations), with macro-scale models is challenging and development of robust representations of management of reservoir series are important directions of future research (Adam et al., 2007; Masaki et al., 2017).

To address our specific hypotheses, we compared a series of model parameterizations from a common baseline. Biases in model representation of the USRB from utilizing a minimally calibrated model are common between each hypothetical parameterization of changing irrigation efficiency. Therefore, the differences between model simulations are informative of the effect that interventions of irrigation technology have on semi-arid agricultural basins generally. Inferences specific to the USRB's response to similar management interventions are inevitable,

so it is worth considering how known model misfit could influence interpretations of the fate of incidental returns, irrigation reuse, and the effectiveness of coupling EAR with increased irrigation efficiency for the USRB specifically. We note several obvious biases between the model simulation at baseline and observations. First, WBM predicts the onset of snowmelt early in most years (Figure AI.5a), which leads to overfilling of the major reservoirs along the cascade of reservoirs through the Upper Snake River. Early season discharge leads to overfilling of reservoirs compared to observations, and then shunting of water downstream causing both a high bias in early season discharge at the basin outlet (Figure AI.5f), and less water in storage late in the season. Excess discharge at the outlet ranges between 0.65 to 8.75 $km^3\ y^{-1}$ with a median of 2.86 $km^3\ y^{-1}$. Furthermore, the early shift in snowmelt makes less water available in the reservoir cascade later in the year leading to overdraft of Palisades Reservoir late in the irrigation season in 2010, 2012, 2015, and 2016, and therefore less water available to American Falls reservoir in those years. Model simulations that more accurately captured the timing of snowmelt onset with the known reservoir management would retain more snowmelt in the reservoir cascade making more snowmelt available to maintain reservoir levels near observations, and for irrigation. Therefore, less groundwater would likely be used for irrigation resulting in less aquifer drawdown and a lower rate of gross irrigation reuse of incidental returns.

We also note that seasonal dynamics of the water table and therefore discharge through springs were highly damped relative to observations, which results from our lumped aquifer parameterization. Prior analysis shows that annual cycles in spring discharge results from fluxes that occur within 20 km of the springs (Boggs et al., 2010). Though mean spring discharge is unbiased, incidental recharge to the aquifer and pumping from the aquifer are spread over the two compartments of the aquifer and exceed the space scales that would create seasonal

dynamics. Suppressed seasonality of spring discharge could reduce seasonality of downstream flows; however, there are no major abstractions of surface water downstream of springs in our representation of the USRB. Moreover, seasonal head fluctuations could reduce pumping by either drying wells, or increasing pumping costs; however, these dynamics are unrepresented in the model, and have not been widely reported as affecting wells drawing from the ESPA. Therefore, we consider the results of our model would be unchanged if seasonal dynamics in aquifer head were more closely aligned with observations.

Simulated irrigation abstractions are generally low compared to USGS observations (Figure AI.3) and could be increased by either forcing lower efficiency of baseline irrigation practice or increasing evapotranspiration from crops. The efficiency of irrigation technologies is reasonably characterized empirically in the baseline parameterization; however, uncertainties with regard to specific technological parameterizations certainly exist. For instance, the distribution uniformity parameter that controls the amount of water applied to a field during an irrigation event can vary dramatically at field scales (Burt et al., 1997). Following the analysis of Jägermeyr et al. (2015), we use the parameters selected from their sensitivity analysis that optimized trade-offs between crop yield and water use for each technology type. The distribution uniformity, as well as parameters controlling percolation beneath canals and soil infiltration rates, could create less efficient irrigation technologies that would reduce the bias in irrigation water used; however, we avoided calibrating to avoid overfitting with respect to drivers of incidental returns and non-beneficial use. Though unbiased at global scales, the potential evapotranspiration calculation used here (Hamon, 1963) may underestimate the flux from the semi-arid environment of the USRB. Alternatives such as the Penman-Monteith (Monteith, 1965), resulted in poorer representation of the spatial variability in irrigation abstractions though the whole-basin total

abstractions were less biased. An increase in irrigation abstractions from higher potential evapotranspiration would increase the baseline CIE by increasing the beneficial consumption of crops, while increasing non-beneficial use and incidental returns only slightly. The excess volume of water lost via simulated discharge from early onset of snowmelt is less than the difference between WBM's and USGS's use estimates for gross irrigation water use in the USRB. Increased abstraction may reduce water available for EAR leading to greater trade-offs between changing streamflow capture and aquifer drawdown.

Comparison of baseline simulations with other studies

The fraction of incidental returns to the ESPA predicted by simulations is a critical factor for interpreting these results, and we compared simulations with both empirical estimates and previous modelling studies. The fraction of incidental returns in ESPA storage was lower than the fraction of incidental returns entering the aquifer as recharge because the aquifer equilibrates over longer timescales than the simulations were conducted. The composition of the aquifer was dominated by relict water because we identified all water in the system as relict at the end of spin-up in these simulations to permit tracking fate of all incidental returns. Following sufficient run-time, the model as parameterized at baseline should equilibrate to a composition of at least 60% irrigation returns (Table 3). In 1994 and 1995, isotopic and geochemical tracers showed that water in downgradient portions of the ESPA and in spring outflows consisted of approximately 75% incidental returns from the Snake River (Plummer et al., 2000). The fraction of incidental returns in ESPA recharge was lower in this analysis than estimated from tracers because a) our estimate represents a dilution of incidental returns over the entire ESPA, not local flow-paths sampled near the down-gradient portions of the ESPA where agriculture is concentrated, and b) CIE efficiency from changing irrigation technology decreased rates of

incidental recharge between 1994 and 2010 (Dieter et al., 2018; Maupin et al., 2014). These differing assumptions of the amount of irrigation return water in the ESPA is accounted for in our analysis.

The IDWR's ESPAM2.1 apparently predicted greater net recharge from irrigated agriculture to the aquifer; however, direct comparisons are complicated by differing simulation time periods and definitions of simulated fluxes (IDWR, 2013). The ESPAM2.1 estimates of net recharge accounted for all infiltration from irrigated croplands, whereas we report the infiltration explicitly from applied irrigation water. Net recharge predicted by ESPAM 2.1 was 3.4 $km^3\ y^{-1}$, greater than the 1.9 $km^3\ y^{-1}$ of net incidental recharge predicted by WBM at the baseline parameterization. The ESPAM2.1 simulation period was earlier than here (1980 through about 2008), but that model did not exhibit trends in net recharge that would make it consistent with WBM during the later simulation period used here. While the greater groundwater abstractions in the WBM baseline parameterization (2.7 $km^3\ y^{-1}$) compared to ESPAM2.1 (2.2 $km^3\ y^{-1}$) may partially explain the difference in net incidental recharge, groundwater abstractions were still lower than the 3.4 $km^3\ y^{-1}$ estimated by Frans et al. (2012). Both crop type data and meteorological data employed by ESPAM2.1 differ from the data used here. Wisser et al. (2008) found that combined influence of climate and crop landcover data resulted in uncertainty in crop irrigation demand of up to 50%, consistent with differences between WBM and ESPAM2.1. Considering the low bias in WBM's simulated gross irrigation compared to USGS water-use estimates (Figure AI.3), and lower rate of net recharge in WBM, we expect that our rates of incidental returns to the system and therefore irrigation reuse are likely underpredicted, at least with respect to the ESPAM2.1. Furthermore, the reduction in net recharge with modernization could be more significant than simulated here, making our estimates of the MB metric

Table 3: Comparison of irrigation related water fluxes in and out of the Eastern Snake Plain Aquifer and Upper Snake River Basin along a gradient of increasing efficiency of irrigation technology. Values without and with enhanced aquifer recharge depicted on left and right halves of each column, respectively.

	Baseline		**Eff. A**		**Eff.B**		**Eff.C**		**Eff.D**		**Eff.E**		**Eff.F**		**Eff.G**		**Eff.H**		**Eff.I**	
		EAR		*EAR*		*EAR*		*EAR*		*EAR*		*EAR*		*EAR*		*EAR*		*EAR*		*EAR*
Classical irrigation efficiency (%)	40.4	40.4	43.5	43.4	46.1	46.0	50.1	50.0	53.8	53.8	56.8	56.7	64.2	64.2	70.6	70.6	77.8	77.8	95.1	95.2
Incidental recharge (fraction)	60	53	56	48	54	45	49	40	44	36	42	31	35	25	29	20	23	15	5	3
Groundw. abstraction ($km^3\ y^{-1}$)	2.64	2.88	2.39	2.31	2.32	2.39	2.12	2.26	1.91	2.23	1.83	1.9	1.57	1.7	1.4	1.55	1.28	1.33	0.95	1.02
Net irrigated recharge (I_{rch} $km^3\ y^{-1}$)	1.83	1.67	1.52	1.55	1.25	1.12	0.86	0.72	0.65	0.42	0.42	0.34	0.12	-0.01	-0.12	-0.25	-0.36	-0.43	-0.77	-0.84
Irrigation reuse (R) ($km^3\ y^{-1}$)	0.86	1.27	0.7	1.12	0.59	1.01	0.47	0.86	0.38	0.87	0.32	0.83	0.2	0.61	0.14	0.45	0.09	0.3	0.01	0.05
Beneficial Reuse (BR/R) (%)	7.67	11.3	7.57	12.5	7.65	13.1	7.73	14.7	7.81	18.9	7.86	21.3	7.64	23.8	7.55	25.2	7.64	26.9	6.85	29.7

EAR – Enhanced aquifer recharge
USRB – Upper Snake River Basin

potentially low (i.e. conservative).

Baseline simulation water budget and fates

Major fluxes of irrigation abstractions are shown schematically in Figure 12. Beneficial consumption (B) was about 3.52 $km^3\ y^{-1}$, representing 40% of gross irrigation abstractions (G) at baseline irrigation. Nearly all incidental returns percolated due to the highly permeable geology underlying most of the USRB. About 10% of gross irrigation abstracted, or 0.86 $km^3\ y^{-1}$ of water (ranging from 0.43 to 1.11 $km^3\ y^{-1}$), was reused for irrigation each year under the baseline conditions (Figures 11 and 12). Figure 11 shows the spatial intensity of irrigation water reuse (R) for the baseline parameterization. Major controls on the spatial distribution of irrigation reuse included a) administrative basin extent and the balance of incidental returns in reservoirs acting as irrigation source, b) upstream catchment area, and c) presence of the ESPA. Reservoirs received incidental returns in runoff from upstream croplands. The fraction of incidental returns in surface irrigation for entire administrative basins reflect the fraction of irrigation returns stored within the collection of source reservoirs. Therefore, the reuse changed abruptly at administrative basin boundaries (Figure 11). Source reservoirs were not defined in Wyoming at the eastern margin of the model domain. Here, water was provisioned by locating the nearest downstream available water so R increased as incidental returns accumulated along downstream flowpaths. Irrigation reuse changed along the margins of the ESPA as the incidental recharge contributed by groundwater abstractions was characterized by the short-turnover shallow groundwater pool outside the ESPA region. Therefore, in the extreme west of the domain, shallow groundwater contained a high fraction of incidental recharge relative to the ESPA.

Beneficial (BR) and non-beneficial reuse (NR) is calculated explicitly in the model as the beneficial and non-beneficial fraction of gross irrigation reuse. The fraction of beneficial reuse

to gross irrigation reuse (BR/R) is roughly equal to basin-wide average classical efficiency (B/G), with slight spatial differences accounting for differing technologies in locations where reuse is more prevalent. Approximately 0.35 $km^3\ y^{-1}$ of beneficial irrigation consumption is derived from irrigation reuse under our baseline parameterization (Table 3) representing about 10% of total beneficial consumption (as BR/B), and 8% of total incidental returns (as BR/L).

Effects of irrigation modernization

Modernization of irrigation technology led to reduced aquifer storage and increased export of water from the basin. Specifically, we find that the modernization decreased plot-scale incidental returns from 4.6 to 0.2 $km^3\ y^{-1}$ (Figure 13a). As a result, the rate of loss from aquifer storage (drawdown) increases from about 0.7 $km^3\ y^{-1}$ to about 1.7 $km^3\ y^{-1}$ when simulated without EAR (Figure 13b), while average annual discharge leaving the basin increases from 10.8 $km^3\ y^{-1}$ to 12.2 $km^3\ y^{-1}$ (Figure 13d, Table 3). In these experiments, crop use is independent of irrigation process, so no changes in beneficial crop evapotranspiration are simulated (Figure 13a). Non-beneficial consumption decreased from 0.62 $km^3\ y^{-1}$ in the baseline to 0.01 $km^3\ y^{-1}$ for parameterization Eff.I. The high rates of percolation exceeded evaporative demand from bare soils so that incidental recharge was much greater than the non-beneficial consumption from bare soil evaporation.

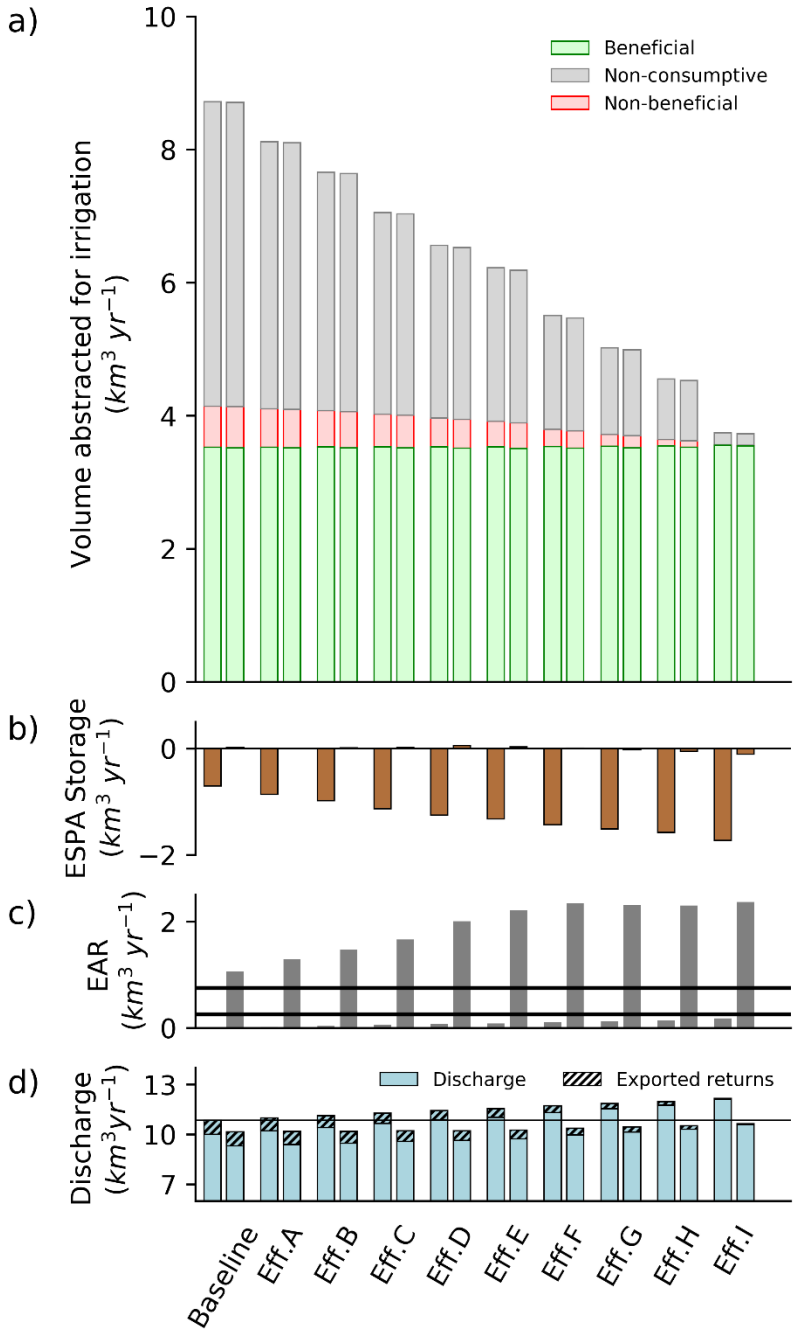


Figure 13: Critical water fluxes across efficiency scenarios paired by simulations without (left) and with (right) enhanced aquifer recharge (a). Component fractions of gross irrigation water for the USRB as 2008-2017 averages. Average change in volume of the ESPA (b). Enhanced aquifer recharge (recharge to the ESPA upstream of American Falls) required to stabilize the aquifer water balance (c). Horizontal lines represent target (0.26 $km^3\ y^{-1}$) and feasible (0.75 $km^3\ y^{-1}$) bounds on existing managed aquifer recharge practice and infrastructure (IWRB, 2016). Discharge and exported incidental returns at the watershed outlet at King Hill, Idaho. Horizontal line indicates average discharge at baseline (d).

Enhanced Aquifer Recharge

Simulated EAR ranged from 1.1 $km^3\ y^{-1}$ (baseline) to 2.4 $km^3\ y^{-1}$ (Eff.I) to maintain aquifer volume within 0.11 $km^3\ y^{-1}$ (Figure 13, Table 3). The 120% increase in EAR from baseline to the most efficient parameterization offset a loss of 4.3 $km^3\ y^{-1}$ from incidental recharge to the aquifer. Incidental recharge from irrigated crops was 4.5 $km^3\ y^{-1}$ at baseline, and net recharge from irrigated agriculture (incidental recharge minus groundwater abstractions) was positive at 1.8 $km^3\ y^{-1}$ at baseline with or without EAR. Incidental returns represented 60% of water entering the ESPA under baseline conditions (Table 3). As irrigation efficiency increased, incidental recharge to the aquifer decreased, declining to 0.18 $km^3\ y^{-1}$; incidental recharge flux did not depend on whether EAR was simulated or not. Groundwater abstractions also declined with increasing efficiency; however, for the most efficient parameterizations, abstractions exceeded incidental recharge and net recharge from irrigated crops (I_{rch}) became negative, declining to -0.8 $km^3\ y^{-1}$ (Table 3).

We hypothesized that the relative increase in EAR needed to stabilize the aquifer would be less than the loss of net irrigated recharge (I_{rch}) resulting from increasing irrigation efficiency. Simulated water balance supported the hypothesis (Figure 14a). For parameterizations more efficient than baseline, the increase in EAR for each parameterization was less than the loss of net irrigated recharge from baseline, and the relationship between the two metrics appeared to be non-linear (Figure 14a). Approximately 72% of the lost net irrigated recharge was required as EAR to stabilize the aquifer for parameterizations Eff.A through Eff.E, and then only approximately 17% of the lost net irrigated recharge was required as EAR for parameterizations Eff.F through Eff.I (Figure 14a). The abrupt change in the effectiveness of EAR to stabilize the aquifer corresponds with an increasing proportion of direct (drip) irrigation for Eff.F through

Eff.I (Table 2), reflecting the relatively larger reduction in the distribution uniformity parameter between sprinkler (0.55) and direct (0.05) than from surface (1.15) to sprinkler (0.55), causing rapidly decreased non-beneficial consumption. The magnitude by which the increase in required EAR is less than the relative loss in net irrigated recharge reflects the management benefit (MB) (Equation 5) that enhanced aquifer recharge, combined with efficiency, has on aquifer balance (Figure 14a). MB increased to a maximum of 1.3 $km^3\ y^{-1}$ for Eff.I, the only parameterization that exceeded the 1.06 $km^3\ y^{-1}$ EAR needed at baseline to stabilize the aquifer.

For all simulations conducted, the rate of aquifer drawdown (negative $\frac{dV_{ESPA}}{dt}$) was greater than the relative change in flow out of the basin from baseline (Figure 14b). Changing flow out of the basin represents a change in streamflow capture, or how use of water in the basin affects the flux leaving through the river. Simulations with EAR exhibited lower outlet discharge compared to baseline (greater streamflow capture or negative Q*-Q - Figure 14b) as a fraction of Snake River flow was diverted to aquifer replenishment. The rate of EAR controlled the rate of streamflow capture by explicitly adding water to the aquifer, and not through altering the head-dependent baseflow flux back to the river, since increasing EAR also increased baseflow. Note that we focus on changes in streamflow capture relative to baseline, and do not make inference to the absolute fluxes of streamflow capture associated with use of the ESPA. As classical irrigation efficiency increased both with and without EAR, the change in streamflow capture came closer to the rate of aquifer drawdown.

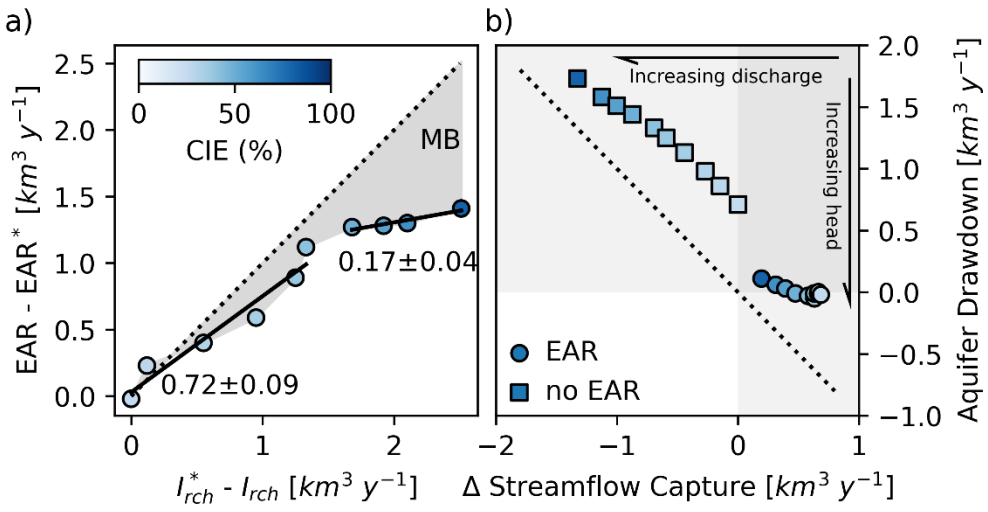


Figure 14: Enhanced aquifer recharge (EAR) above the EAR required at baseline (EAR*) to ensure aquifer stabilization plotted against the reduction in net recharge from irrigation ($\boldsymbol{I_{rch}}$) from baseline ($\boldsymbol{I^*_{rch}}$) as classical irrigation efficiency (CIE) increases (a). Dotted line represents equal increases in EAR and reductions in net recharge. All scenarios show that less additional EAR is required than is lost from net recharge as CIE increases, the magnitude is referred to as the management benefit (MB). Slopes of piecewise linear regressions (black lines) between two variables are shown with standard error of the estimate. Aquifer drawdown plotted against change in basin streamflow capture (Q*- Q) with dotted line representing equal changes to discharge from baseline and drawdown (b).

Discussion

Aquifer reliance on incidental irrigation for recharge

We found a non-linearity in the volumes of enhanced aquifer recharge (EAR) required to stabilize the aquifer as more efficient irrigation technologies were employed. That is, incrementally smaller volumes of enhanced aquifer recharge (EAR) were needed compared to the net irrigated recharge lost due to using more efficient technologies (Figure 14a). This applied only for incremental increases in EAR volumes above requirements for aquifer stability at the baseline parameterization. The volume represented by the efficiency of the combined system from pairing increasing CIE with EAR, the management benefit (MB - Figure 14a) demonstrates

that the rate of increasing EAR is less than the rate that net recharge declines. However, the volumetric benefit only exceeded the baseline requirement of EAR for the most efficient (Eff.I) case, and the benefit was not evident for parameterization Eff.A. The management benefit is predominately attributed to additional capture of Snake River discharge (Figures 13 and 14), and by way of increasing irrigation water reuse (Table 3) and decreasing incidental returns (Figure 13). This illustrates that in regions with conjunctively managed surface and groundwater sources like the USRB, increasing basin-wide water resource availability via combined implementation of managed aquifer recharge and changing irrigation efficiency can only be expected to capture more streamflow by transferring water to longer residence time compartments during seasons when water is more available.

Our simulations suggest several important implications of a conjunctive management strategy promoting aquifer recharge while increasing the efficiency of irrigation technology. The amount of EAR needed for a given technology parameterization always exceeded the corresponding rate of drawdown without EAR by 36 to 66%. The excess EAR was needed because the aquifer drainage continues between peak EAR, which follows peak river flow from March to May, and peak irrigation demand (July), and because diverting water away from supply reservoirs shifts reliance to groundwater, which in turn required additional EAR for stabilization. The shift to more groundwater utilization is the primary reason for greater irrigation water reuse for simulations with EAR compared to simulations without (Table 3). Furthermore, the rate of aquifer drawdown (up to 1.7 $km^3\ y^{-1}$ without EAR, and approximately 0 $km^3\ y^{-1}$ with EAR) more closely approximated changing streamflow capture as CIE increased, meaning that the system converted a greater proportion of the captured streamflow to aquifer storage bringing the change in fluxes into greater parity. Despite the increasing parity between the rate of drawdown and

capture with increasing CIE, drawdown always exceeded the magnitude of the change in streamflow capture (Figure 14b). This is an expected result because surface water is the dominant source for irrigation and the aquifer is naturally located upgradient of the basin's outlet; interventions that add water to the aquifer (decrease drawdown), will eventually lead to increased downstream discharge (decreased streamflow capture), but the converse is not generally true. Therefore, increasing CIE without EAR will act to deplete the resource relied on by groundwater irrigators more than the impact that EAR would have on downstream users, at least in terms of volumetric shortfalls.

The rate of change in aquifer storage $\left(\frac{dV_{ESPA}}{dt}\right)$ factors significantly in the preceding analysis, and over-estimation of the present-day rate of aquifer drawdown may shift values, but are unlikely to change the general conclusions. At baseline, we estimated $\frac{dV_{ESPA}}{dt}$ to be -0.71 $km^3\ y^{-1}$, which is a greater rate of drawdown than the -0.34 $km^3\ y^{-1}$ estimated by the ESPAM2.1 (IDWR, 2013), the latter being likely more accurate given our underestimates of percolation losses described above. The rates of EAR we identified to stabilize the lumped representation of the ESPA exceeded both targets and feasible limits of managed aquifer recharge (IWRB, 2016, Figure 14c). This potential limitation should be explored in future research focusing on evaluating specific management objectives.

Our simulations assumed a constant beneficial consumption, though use and efficiency are often positively correlated due to the economic incentives to use more water when it is made available locally through efficiency measures (Contor and Taylor, 2013; Grafton et al., 2018; Pfeiffer and Lin, 2014; Tran et al., 2019), and because prior appropriation doctrine requires that water rights holders use their full water right beneficially, essentially encouraging constant levels of water

withdrawal regardless of CIE. In the USRB, it is reasonable to assume negligible slippage and rebound effects. Frequent droughts, the collective action of irrigation districts, and legal agreements between water user organizations outside of the prior appropriation system, all work to incentivize reduced water withdrawals when possible (Gilmore, 2019). Moreover, a settlement between surface and groundwater irrigators (IDWR, 2015) details specific requirements for ensuring stable aquifer head for both irrigation and downgradient outflow from springs. To the extent that beneficial use could increase with higher CIE, greater EAR would be required to meet the mandate of aquifer stabilization, or aquifer drawdown would increase without EAR, for any given CIE relative to that simulated here.

The generalizable findings from these simulations have implications for similar semi-arid basins relying on a combination of groundwater and seasonably available surface water. Achieving aquifer stabilization and increasing downstream discharge from combining increased CIE with EAR as simulated here, would require significant investment in hydroinfrastructure of the basin. In some systems these investments may be a prerequisite for groundwater sufficiency (Scanlon et al., 2016). In these simulations, EAR was a prerequisite for aquifer stabilization because no tested CIE was able to create a stable aquifer with existing agricultural production and natural recharge alone. In the USRB, the current head level targeted for stabilization is greater than head existing in the basin prior to irrigation (Kjelstrom, 1995), and generates increased rates of baseflow from springs. This is a unique issue from many other semi-arid basins relying on groundwater for irrigation that are managed against aquifer depletion below pre-irrigation heads (Bierkens and Wada, 2019). The primary adverse externality of EAR, aside from technical considerations of feasibility, is decreasing watershed discharge on an annual basis, which would be undesirable for downstream users; however, increasing flow during the irrigation season can

be expected (Van Kirk et al., 2020). We found that decreased downstream flow simulated here with EAR, which at 10.2 to 10.6 $km^3\ y^{-1}$ still exceeds the observed record during the same period (7.35 $km^3\ y^{-1}$) and existing requirements for instream flow (e.g. 4.1 $km^3\ y^{-1}$) (IWRB, 1985) by greater than the existing model bias in outlet discharge.

Irrigation Reuse in the Upper Snake River Basin

Incidental returns from irrigation were a major component of the basin's water balance, and therefore are key to understanding basin-scale interpretations of system efficiency. Within the USRB, reuse of incidental returns generated during the model epoch currently makes up at least 9.9% of gross irrigation and would increase to 14.6% if EAR was used to stabilize the aquifer (Table 3, Figure 12). The baseline value of irrigation water reuse is likely underestimated due to the low bias in gross irrigation, lower rate of net agricultural recharge relative to ESPAM2.1, and a high fraction of relict water composing the ESPA water volume in our simulations. As irrigation efficiency increased and incidental returns decreased both with and without EAR, the total reuse of irrigation water declined (Table 3). However, the fraction of incidental returns that were ultimately used beneficially (beneficial reuse) exhibited very different behavior if EAR was simulated. With no EAR, beneficial reuse remained between 7 and 8% of total incidental returns for all efficiency parameterizations. With EAR, the beneficial reuse increased steadily with CIE to 30% of total incidental returns for parameterization Eff.I. As a result of the increasing beneficial reuse, basin-scale effective irrigation efficiency either increases faster (with EAR) or slower (without EAR) than classical irrigation efficiency (Figure 15).

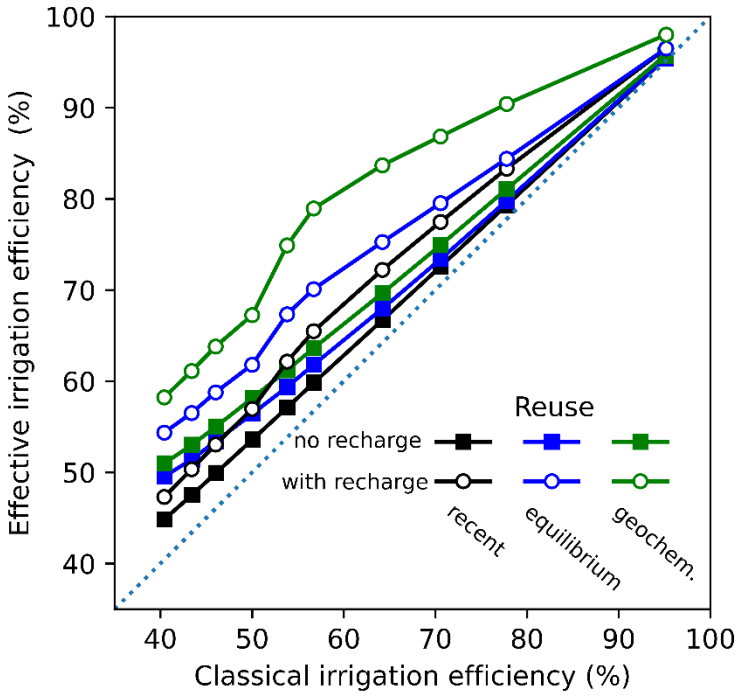


Figure 15: Effective irrigation efficiency plotted against the classical irrigation efficiency of each parameterization. Effective irrigation efficiency is calculated three ways based on the estimates of irrigation reuse: near-term – simulated reuse where incidental returns in aquifer abstractions is represented explicitly during the model epoch (3%), equilibrium – incidental returns in the aquifer abstractions are assumed to be at equilibrium in the aquifer at ratio of incidental recharge to total recharge (Table 3), and geochemical – incidental returns in aquifer abstractions are assumed to be represented by an average estimated geochemically (Plummer et al., 2000) (75%).

Metrics such as the effective irrigation efficiency (EIE) provide a unified metric of efficiency that captures the reusability of incidental returns at the watershed scale (Haie and Keller, 2008). Generally, EIE is calculated using assumptions of the recoverability of irrigation returns; however, we calculate basin-wide EIE using simulated recovered volumes and thereby incorporating explicit estimates of recovered returns. Water within the ESPA was primarily simulated as relict water; therefore, the simulations neglect a significant volume of irrigation returns stored within the aquifer from incidental recharge pre-dating the model epoch. Our estimates of irrigation reuse are therefore low and reflect only reuse of incidental return and

subsequent abstractions from the aquifer during the model epoch. We calculate three estimates of EIE, 1) near-term EIE: using the explicitly tracked incidental returns during the model epoch, 2) equilibrium EIE: assuming the equilibrium fraction of incidental returns in abstracted groundwater equals the ratio of incidental to natural recharge (Table 3), and 3) geochemical EIE: assuming that aquifer abstractions consist of a constant fraction of 75% incidental returns estimated geochemically by Plummer et al. (2000). Without enhanced recharge, the actual rate of recovery of incidental returns via irrigation was low, so effective irrigation efficiency is only slightly greater than classical irrigation efficiency at any parameterization (Figure 15) and reflects the large proportion of fresh snowmelt used to supply irrigation most years (Figure 13). With enhanced recharge, the added reuse increases EIE faster than CIE for parameterizations Eff.A through Eff.G. Assuming equilibrium or geochemical estimates of returns in aquifer water increases estimated EIE by 7 and 11% at baseline, respectively. Moreover, improving irrigation efficiency from baseline through parameterization Eff.E increases the rate that EIE improves. The parameterizations that correspond with an increasing rate of EIE improvements are the same parameterizations that show a smaller increase in the management benefit, e.g. a smaller amount of additional EAR compensating for the loss of net agricultural recharge (Figure 14a). Though the EIE captures a more complete picture of the effect of changing irrigation technology over the complete system, the high rate of increase in EIE for small changes in irrigation technology may overstate the benefits of intervention on the water balance of the entire basin as captured by the calculation of management benefit.

Incidental returns as a component of discharge at the basin outlet at King Hill, ID was 0.84 km^3 y^{-1} (approximately 8% of streamflow) at baseline and declined as CIE increased (Table 3). Therefore, the recoverable incidental returns can be used beyond the USRB. With EAR,

incidental returns in discharge were slightly higher than without for each technology parameterization and declined to 0.04 km^3 y^{-1} at Eff.I; thereby decreasing incidental returns as a fraction of flow to 0.3%. Modernization acted to increase the unabstracted fraction of discharge leaving the basin, therefore benefitting downstream users while increasing aquifer drawdown. The addition of EAR captured more unabstracted streamflow in the basin, while maintaining a similar flux of exported incidental returns.

Decreasing the fraction of incidental returns in river flow would be expected to improve water quality in the river. However, increasing irrigation reuse implies further recycling of agricultural runoff, which tends towards greater acute water quality threats such as salinization (Ghassemi et al., 1995; Qadir, 2016) and increasing nitrate concentration (Frans et al., 2012). Presently, neither soil salinization nor waterlogging are widespread in the USRB owing to existing conjunctive water abstractions and good drainage, but as irrigation technology modernizes in the USRB and excess irrigation water for flushing is reduced, isolated instances of salinization are becoming increasingly common (Ellsworth, 2004; Moore et al., 2011). While decreasing incidental recharge could exacerbate soil salinization if left unmanaged, irrigation reuse and incidental returns in USRB export both declined (Figure 14, Table 3), which could potentially improve water quality to the ESPA and downstream users. Our definition of incidental returns included canal seepage, a major source of recharge to the ESPA. Canal seepage only represents a source of contaminates if they receive poorly managed runoff, which is not evaluated here. Therefore, in the USRB incidental returns and reuse can only be loosely interpreted as an indicator of water quality, and fate and transport processes would be needed to assess the explicit fate of any agricultural contaminants. Considering the growing concerns of salinization associated with irrigated agriculture (Cañedo-Argüelles et al., 2013; Ghassemi et al., 1995),

especially in semi-arid and arid regions with increasing technological efficiencies (Banin and Fish, 1995; Carr et al., 2010; Tal, 2016), additional attention is needed to evaluate trade-offs of managing soil salinization and efficiency of irrigation technology.

Conclusions

Our simulations of the USRB characterize the limitations of relying exclusively on technological adaptation to address water shortfalls in semi-arid regions. Technological modernization does not by itself promote aquifer stabilization in some contexts. Modernization without managed aquifer recharge (MAR) resulted in a greater loss from aquifer storage and increased downstream flow, undermining the groundwater resource needed for agriculture resiliency in this semi-arid basin. Furthermore, we found that through combined application of MAR and increasingly efficient irrigation technology, the potential increase in downstream flow was always less than the increased drawdown in the aquifer, meaning that less streamflow capture than drawdown was needed for similar crop production in a conjunctively managed system. By increasing MAR to values likely difficult to achieve in practice (IWRB, 2016), the system utilizes only a portion of the net irrigated recharge lost by modernization to stabilize the aquifer. The simulations tested demonstrate the trade-offs inherent in reducing non-consumptive losses through modernization that have been explored in other regions with high gross irrigation reuse (Simons et al., 2015) and illustrate how modernization exports benefits to downstream users. The absence of clear evidence for significantly improved water availability with modernization that is predicted for global scales (e.g. Jägermeyr et al., 2015, 2016; Sauer et al., 2010) is because exported benefits (net increase in water availability) are absorbed by downstream users when analyzed at that scale (Grogan et al., 2017). However, in a single headwater semi-arid basin, there is a fundamental lack of parity between local groundwater users and downstream users; any intervention that

improves aquifer storage necessarily also benefits downstream users eventually, while the converse is not necessarily true. Also, potential policy and comprehensive water management initiatives which are likely to co-occur with modernization (Gleick et al., 2011) can provide additional benefits to basin water budgets not realized solely by modernizing irrigation technology (Jägermeyr et al., 2015). The ineffectiveness of technological modernization to stabilize the aquifer by itself may reflect the specific setting of the USRB that naturally favors non-consumptive loss to non-beneficial use via high percolation rates coupled with a straightforward avenue for local reuse via a productive aquifer and springs. Irrigation reuse declines as classical irrigation efficiency increases, but using MAR increased the reuse of incidental returns. Though we expected MAR to reduce reuse of incidental returns through the introduction of more pristine water to the aquifer, the larger effect of shifting irrigation reliance towards groundwater from surface water was observed, thereby increasing reuse at the basin scale. The added reuse from implementing MAR in our simulations lead to effective irrigation efficiency increasing faster than classical irrigation efficiency. We would expect the nature of gross irrigation reuse in the USRB to be neither an isolated instance, nor a general exemplar of water allocation issues, but it does provide an example of the complexity and lack of generalizability of specific interventions needed to achieve agricultural sustainability.

CHAPTER III

EXISTING WETLAND CONSERVATION PROGRAMS MISS NUTRIENT REDUCTION TARGETS

Increasing coastal hypoxia results from export of excess nutrients used for fertilization of row crops (Goolsby et al., 2000; Fennel and Testa, 2019; Goolsby et al., 2001; Tian et al., 2020), and has the potential to fundamentally change the character and habitability for marine species over the next millennium (Breitburg et al., 2018). Nutrient-rich runoff from agricultural lands in the Mississippi/Atchafalaya River Basin (MARB) needs to decline by about half to reverse the trend of expanding areas of bottom-level hypoxia in the Gulf of Mexico (GoM) (Dale, 2010; Goolsby et al., 2000; Marshall, 2018; Scavia et al., 2017). Nutrient mitigation measures should target hot-spots of excess fertilizer application or manure production (Roy et al., 2021). Subsurface drains (SSDs) are widely distributed throughout the mid-western United States (Sugg, 2007), and leachate from corn/soy rotations grown on SSD croplands represent the largest single source of excess nitrogen discharged to the Gulf of Mexico (Goolsby et al., 2001). Numerous studies (Marshall, 2018; Christianson et al., 2018; Liu et al., 2018; Ribaudo et al., 2001; Zimmerman et al., 2019; Santhi et al., 2014) identify common mitigation measures suggesting agreement about appropriate measures. Field margin interventions that intercept nitrogen-laden water between crop fields and streams, including in-situ bioreactors (Jaynes et al., 2008; Schipper et al., 2010), riparian infiltration (Jaynes and Isenhart, 2014), and constructed, treatment, or restored wetlands are often cited as effective nutrient reduction strategies (NRS) (Ribaudo et al., 2001; Christianson et al., 2013; Mitsch et al., 2001; Hansen et al., 2018; Iovanna et al., 2008; Crumpton et al., 2006; Cheng and Basu, 2017) as they exhibit high nutrient removal capacity, require low to moderate operational labor, and have numerous synergistic benefits such as flood regulation (Ameli and Creed, 2019) .

Despite the level of effort afforded to studying the denitrifying capacity of wetlands, recent prior studies at continental scales have used simplified scaling approaches to predict restoration effects

with divergent estimates of export reduction ranging from 1.6% (Marshall, 2018) to 54% (Cheng et al., 2020). There are numerous complexities in representing wetland dynamics such that differing assumption should be expected to generate a range of estimates of the potential conservation benefit. Denitrification (bacteria mediated reduction of nitrate to gaseous nitrogen and nitrous oxide) in wetlands is understood to be dynamic over time and space scales (Wollheim et al., 2014; Mulholland et al., 2008; Stewart et al., 2011), but is often assumed to be a temporally constant percent of inputs (denitrified fraction between 40%-50%) in macroscale studies that include wetland restoration (Marshall, 2018; Ribaudo et al., 2001; Christianson et al., 2013; Mitsch et al., 2001; Kovacic et al., 2000; Casey and Klaine, 2001). Hydrologic conditions interacting with wetland characteristics create complex distributions of transport time-scales affecting denitrification (Carleton and Montas, 2010; Werner and Kadlec, 2000; Lightbody et al., 2008). Such dynamics can create pulses of increased nitrate flux from wetlands to receiving waters especially during storms (Fisher and Acreman, 2004; Baker et al., 2018). Engineering constraints needed for system longevity further limit both the area and amount of runoff wetlands can process (Christianson et al., 2013; Tanner et al., 2010; Iovanna et al., 2008). In addition, the cumulative influence of wetlands affect both hydrologic (Rajib et al., 2020) and biologic (Johnston, 1991) response in streams, which may further increase denitrification in river networks (Evenson et al., 2021; Cheng et al., 2020). Considering the above complexities, it remains unclear whether a subset of assumptions describing these processes is responsible for the wide range of potential nitrate reduction predicted in prior studies, and better understanding which of these dominate whole system nitrate reduction potential can guide future research or conservation efforts.

Here we provide a whole-system assessment of field-margin wetland restoration using coupled process-based models to understand mechanisms that cause such a wide range in the potential impact of using wetlands to reduce nutrient export. Furthermore, the models provide a systems-scale analysis of nutrient reduction outcomes for two federal programs incentivizing the restoration of croplands to treatment wetlands in the United States. Our analysis, using coupled Earth system models, evaluates the efficacy of wetland restoration across gradients of adoption of two federal programs in the United States (the Farmable Wetlands Program and the Wetlands Reserve Program, FWP and WRP, respectively). Prior studies have not focused specifically on these restoration programs, nor provided as comprehensive of a representation of engineering constraints or limitations imposed by seasonal and storm-scale dynamics.

We find altogether unique results to prior studies when the agroecosystem is analyzed in an integrated framework that reflects the process of landscape restoration, while adhering to engineering constraints, and when the serial connections through the river network are considered. Furthermore, we account for daily time-scale dynamics of hydrologic flows and biologic uptake, factors rarely treated in prior analyses (Evenson et al., 2021). We ask whether restoring wetlands on restorable lands could reduce nitrate export to the Gulf of Mexico sufficiently to meet targets (Mississippi River/Gulf of Mexico Watershed Nutrient Task Force, 2008; Scavia et al., 2017) established to protect the marine ecosystem. We hypothesized that both seasonality and landscape design constraints not previously considered together at the whole basin-scale would be critically important when considering the effectiveness of wetland restoration to mitigate nitrate export.

Methods

Overview of models

We coupled macro-scale process models of agroecology (AgroIBIS - Kucharik, 2003) and hydro-biogeochemistry (WBM - Stewart et al., 2011; Wollheim et al., 2008a; Samal et al., 2017), and introduced new functionality to WBM that represented flow and nitrate transport from local croplands through field-margin treatment wetlands. The framework represents the equilibrium fluxes of nitrate through the MRB considering historic scenarios of wetland restoration under the two programs that focus on highly optimized treatment wetlands replacing subsurface-drained crops (Farmable Wetlands Program - Conservation Reserve Program, 2015), and opportunistic restoration of wetland systems where replacing croplands was ecologically feasible (Wetland Reserve Program -Natural Resources Conservation Service, 2021). Nitrate leachate from AgroIBIS was input to WBM following aggregation by crop area. Nitrate leachate flux at each pixel was an average of each crop type weighted by a representative area of each landcover.

Crop cover

Crop cover fractions of each 5-minute pixel were aggregated from Ramankutty et al. (2008) and split between irrigated and rainfed crops following Siebert and Döll (2010), both datasets representative of the year 2000. The total of irrigated and rainfed maize, soy, and wheat from these datasets represent 94% of all agricultural lands in the MRB. Therefore, we defined only a single “other crop” category and associated other irrigated crops using fluxes calculated for irrigated maize, and other rainfed crops using an average of values from rainfed maize and wheat.

To define the fraction of cropland undergoing annual maize/soy rotation, we utilized the USDA's Crop Data Layer (Han et al., 2012) to calculate the fraction of 30-meter pixels that transition from corn to soy or from soy to corn for each year of data from 2008 to 2019, as well as those pixels that remain maize or remain soy for those years. We then averaged those pixels up to our 5-minute resolution, to capture the relative fraction of continuously planted maize, continuously planted soy, and annual maize/soy rotation within each 5-minute pixel. We then assumed these averages are appropriate for the domain in the year 2000 and applied these fractions to the sum of maize and soy from Ramankutty et al. (2008).

Subsurface drains

Data on the spatial cover of subsurface drains (SSD) is available as county average fraction of cropland outfit with SSDs (f_{ssd}). Sugg (2007) calculated the area over which SSD is possible, then adjusted these values to match several different incomplete inventories to arrive at a reasonable approximation representative of the 1990s across the conterminous United States. This dataset was found to predict greater spatial coverage of SSD in northern Minnesota when compared against a high resolution dataset that combined remote sensing data with permits for construction (Cho et al., 2019). SSD area fraction (A_{SSD}) for each US county (U.S. Census Bureau, 2015) was rasterized to our model resolution, and at each pixel A_{SSD} was limited to the pixel intersection with total crop area (Ramankutty et al., 2008) to ensure consistency between our datasets. A complete description of SSD representation in WBM is provided in Appendix III, which includes a table describing each of the variables used here (Table AIII.1).

Field-margin wetlands model

Briefly, our model of wetland denitrification assumes a well-mixed system with denitrification occurring along the benthic surface parameterized as a temperature-dependent process, with

flow and nitrogen bypassing the wetland when water storage exceeded a specified maximum depth (Figure 16). The wetland pool was conceptualized as occupying the lowest areas of each pixel along the riparian margin nearest streams such that it can receive flow from surrounding uplands. Though the pool was calculated as a single volume, in practice restored wetlands would be distributed over multiple locations within a pixel. This assumption maximized the amount of crop runoff that can be treated, making our estimates of treatment wetland denitrification biased high, but is consistent with the identification of potentially restorable wetlands (Horvath et al., 2017). A complete description of the representation of upland wetlands is provided in Appendix III.

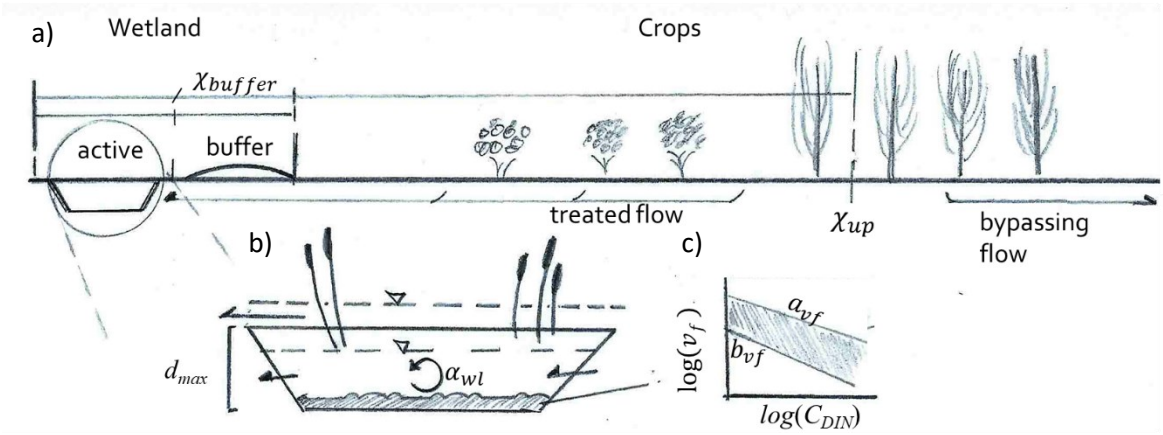


Figure 16: Field-margin wetlands and parameter definitions including catchment and buffer areas (a), maximum depth and wetland storage time-constant (b), and efficiency loss relationship between denitrification uptake velocity and concentration (c). Parameter definitions are presented as Table AIII.1.

Groundwater fluxes

A fraction of all leachates percolating to groundwater (χ_{lost}) is removed from the surface flow system. The term accounts for long-term net storage term in the vadose zone or deep

groundwater, or other forms of removal including nitrate assimilation, possibly along riparian fringes, and was introduced to reduce a high bias in watershed scale nitrate flux from the river system. The parameter was the only calibrated value in the study and was used to minimize the bias to observed riverine nitrate flux. We assume that this flux is independent of wetland processing, which is an extension of our assumption that wetlands are not interacting with deep groundwater flowpaths and as such does not interact with our counterfactual analyses of wetland restoration.

Temporal and spatial resolution

AgroIBIS was run at a 5-arcminute resolution across the conterminous United States. Spinup was performed from 1650 to 1947 to generate equilibrium soil biogeochemistry assuming pre-agricultural vegetation until 1850, then unfertilized wheat through 1947, while recycling climate input data available for 1948 to 2007. WBM was run at a 5-arcminute resolution over a geographic domain from 113°55' W (west) to 77°50'W (east), and 28°55'N (south) to 49°45'N (north). The evaluated domain covered the drainage basin defined by the MERIT 5-arcminute drainage network (Yamazaki et al., 2019). WBM simulations were performed at a daily time-step. To perform model spinup, WBM was forced with input data from 1996 through 1999 repeated five times, then run from 1992 through 1997; a total of 26 years of model spinup. We analyzed output data from 1998 through 2007.

AgroIBIS calculates water and nutrient balance in soils and calculates nitrate leaching from the root zone (Kucharik et al., 2000). Simulations used here were reported previously (Kucharik, 2003; Donner and Kucharik, 2008) but re-run with additional outputs needed for WBM at a daily time-step. AgroIBIS assumes all nitrate leaves the root zone via infiltration below the soil with no direct surface runoff. AgroIBIS output pixel specific fluxes for each of 8 land-cover classes

over agricultural areas of the MRB: 1) irrigated maize, 2) rainfed maize, 3) irrigated soy, 4) rainfed soy, 5) irrigated soy/maize rotation, 6) rainfed soy/maize rotation, 7) rainfed wheat, 8) natural vegetation.

Assessing model behavior

We compared monthly mean discharge and nitrate flux at USGS gaging stations throughout the MRB. USGS data were collected through the National Water Inventory System (U.S. Geological Survey, 2016), for all 58 gaging stations with greater than 200 nitrate samples collected since 1980 at co-located continuous discharge measurements. We reference the basin outlet at the USGS gaging station 07374000, which has continuous discharge data only since 2004, and backfill with chemistry data collected at St. Francisville, Louisiana (07373420) and daily discharge data collected by the US Army Corp of Engineers at Red River Landing (station 01100Q). Monthly flux data is calculated using LOADEST (Runkel et al., 2004) for each station using automated search for the best regression model and linear approximations for the standard errors. The selected regression models for most stations were most often quadratic functions of the natural log of discharge and time selected by the Akaike Information Criterion (Akaike, 1974). From the pool of 58 stations, LOADEST succeeded in determining a best regression model at 15 stations and were distributed throughout the entire MRB. We compared model performance to 12 of these stations, after removing station 06259000 below Boysen Reservoir on the Wind River because WBM is unable to recreate the active reservoir management, station 06752280 at Cache La Poudre, Tinmath, Colorado because of data errors, and station 06877600 on the Smoky Hill River, in Enterprise, Kansas where WBM predicts much higher discharge than observations likely owing to localized loss of water from stream channels to the aquifer (Ferrington JR., 1993), which violates WBM model assumptions. We compared the model with

daily discharge and nitrate flux at each station, as well as quarterly and long-term mean values across the entire pool of stations using percent bias (PBIAS), model percent error (MPE), Nash-Sutcliffe Efficiency (NSE - Nash and Sutcliffe, 1970), and Kling-Gupta Efficiency (KGE - Gupta et al., 2009). Manual calibration of χ_{lost} was used to minimize bias. However, the KGE was the metric maximized during calibration such that the baseline parameter set used in our analysis resulted in a slight negative bias in simulated monthly nitrate flux across the basin both across all tested stations (Figure 17) and at the basin outlet on the Mississippi River at Baton Rouge (Figure 18). The model captured seasonality of nitrate discharge when comparing observations to the baseline system configuration with 0 km^2 restored wetland area (Figure 18). Time-series of basin nitrate flux are presented for five levels of adoption of the Wetland Reserve Program (Figure 18) for reference to observed export.

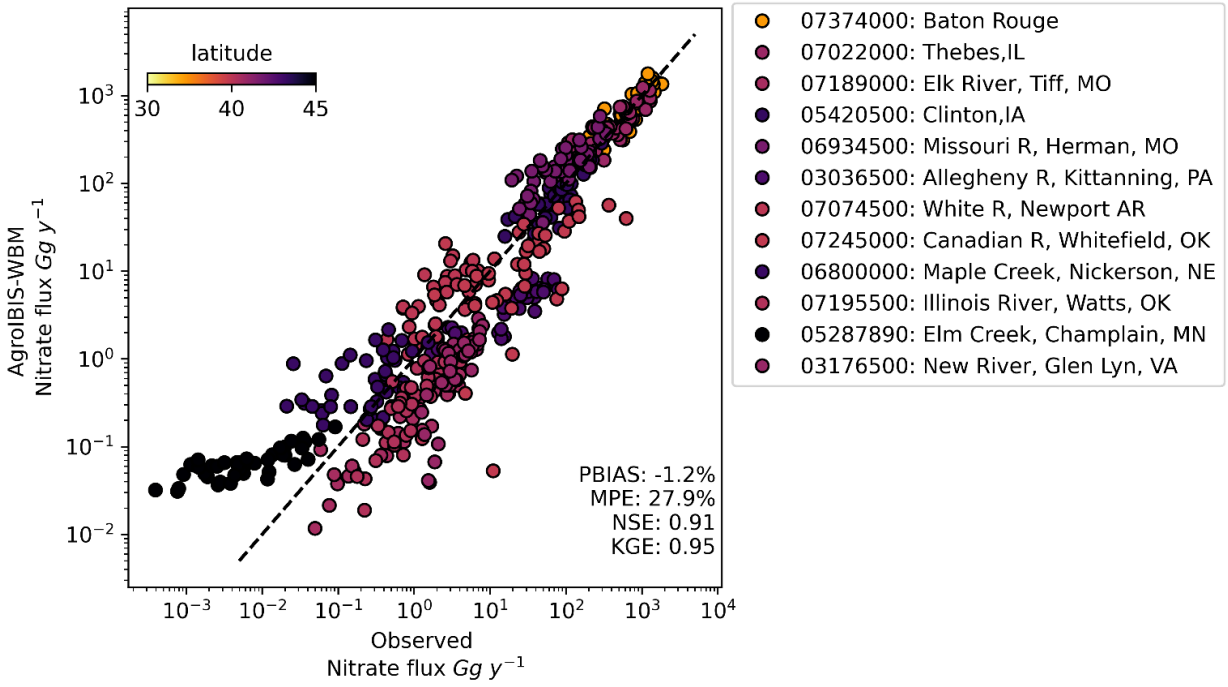


Figure 17: Comparison of monthly average AgroIBIS-WBM simulated nitrate flux and observations calculated using LOADEST from USGS concentration and flow observations at 12 stations. Dashed line depicts agreement between models and observations.

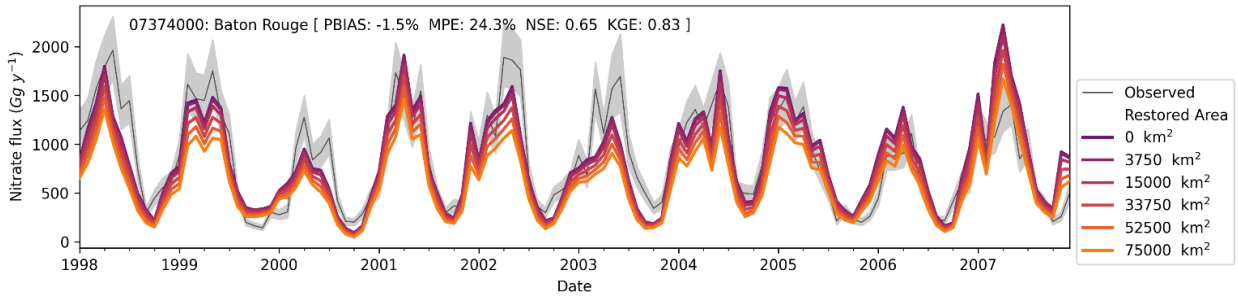


Figure 18: Monthly time-series of nitrate export from the Mississippi River at Baton Rouge for six scenarios of wetland restoration following the Wetland Reserve Program compared against observed nitrate flux with error bands in light gray estimated using LOADEST.

Wetland restoration

We selected areas for restoration from existing crops from the Potentially Restorable Wetlands on Agricultural lands (PRW-Ag) dataset (Horvath et al., 2017) selecting any areas identified as either moderate or high potential for restoration. In both restoration programs, the largest limiting factor in restored area was the amount of upstream catchment area remaining in agricultural production. We assumed that restored lands received drainage from only cropped areas; therefore, the maximum amount of crop runoff was treated for the area restored to wetlands. The maximum area for restoration via the Wetland Restoration Program (WRP) was limited to PRW-Ag areas with upstream rainfed crop area equal to 8.4 times the restored area; WRP wetlands (with buffers) were assumed to consist of 12% of their catchment area, consistent with mean wetland/catchment area ratios in the basin (Wu and Lane, 2017; Cheng et al., 2020). The maximum area for restoration via the Farmable Wetlands Program (FWP) was limited to tile-drained PRW-Ag areas with upstream tile-drained, rainfed crop area equal to 22 times the restored area; FWP wetlands (with buffers) were assumed to consist of 4.5% of their catchments (Iovanna et al., 2008). We assume that the limiting geographic areas (tile-drained lands, rainfed

crops) in our analysis are located entirely upstream of restorable wetlands within each pixel. We assume that both PRW-Ag restoration practice would focus on rainfed crops, as we assume that irrigated crop areas are likely too dry to be appropriate sites for maintaining levels of inundation needed for wetland restoration. Furthermore, because the PRW-Ag data correlates with tile-drained lands at county and coarser scales (Horvath et al., 2017), we assume the extrapolation of restorable wetlands to be located on tile-drained lands at sub-pixel scales is also appropriate. We created counterfactuals of adoption where each pixel was independently increased between the unrestored baseline land-cover and its maximum potential restored landcover at 0%, 5%, 20%, 45%, 70%, and 100% adoption of each program.

Affect of modeling assumptions

By assuming that potentially restorable wetlands were located at the topographically lowest portions of all pixels, and that all crop leachate could be routed to flow through these wetlands, the maximum amount of treatable mass was assumed available to any restored or natural wetlands at each pixel, thereby maximizing the amount of denitrification possible via wetland restoration. Assumptions underlying the upland wetland solution detailed in Appendix III do not uniformly assume maximum denitrification. For instance, our formulation of wetlands as a well-mixed system produces less denitrification than either tanks-in-series (Crumpton, 2001; Kadlec and Knight, 1996) or transient storage (Cheng and Basu, 2017; Wollheim et al., 2014) models applied to natural wetlands. As discussed below, the fraction of mass denitrified within wetlands (30 – 50%) estimated by the model compares favorably with these formulations and in-situ measurements. Other assumptions have potentially ambiguous effects on the total denitrification estimated by the model. By assuming water temperature from soil leachate and neglecting radiative heating or cooling, we likely over-predict denitrification in the early spring, and under-

predict denitrification in summer. We also assume that wetlands do not exchange with groundwater, which is a reasonable assumption for treatment wetlands restored through the Farmable Wetland Program but is questionable as a uniform assumption over the entire domain of the Wetland Reserve Program implementation. Characterizing connectivity of upland wetlands with groundwater will require further analysis. Baseflow entering wetlands would make more mass available for wetland denitrification while potentially lowering nitrate concentration (increasing denitrification efficiency) following removal in groundwater, and elevating spring temperatures within the wetland, each could act to increase the mass denitrified. However, steady flow from groundwater could also decrease water temperature during summer which would decrease the denitrification uptake rate coefficient. Groundwater input to wetlands could also increase depth of flow in the wetlands, which would decrease benthic exchange of dissolved nitrate. Moreover, a flowpath of discharge from wetlands to groundwater is not considered, which could result in higher denitrification rates due to efficiency gains at lower nitrate concentrations entering groundwater, although such processing would likely be negigible. These complex interactions with groundwater will vary by specific local context at resolutions difficult to represent at the resolution of our model, have countervailing effects, and should be a focus of further study.

Sensitivity analysis

We performed a sensitivity analysis to assess the influence of six critical parameters defining performance of the wetland system on total denitrification and resulting change in export of nitrate to the Gulf of Mexico. The Method of Morris (Campolongo et al., 2007; Morris, 1991) provides an estimate of global sensitivity of a computationally expensive model $f(\boldsymbol{X})$ to a suite of parameters or model inputs. For a vector $\boldsymbol{X}$ of model inputs, a suite of M trajectories through

parameter space each perturb a single parameter x_i by a value Δ_i. Each element along sample trajectory j varies each single input successively. The Elementary Effect (EE) of input *i* on a specific model output for trajectory j is calculated by calculating a sensitivity of the model from the trajectory's base condition to the perturbance according to Equation 7.

$$EE_i^j = \frac{f(x_1 \ldots, x_i + \Delta_i, \ldots x_n) - f(x_1 \ldots, x_i, \ldots x_n)}{\Delta_i} \quad \text{Eq. 7}$$

The mean (μ_i), standard deviation (σ_i), and mean of absolute values of EE_i^j ($\mu_i^\star$) Equations 8-10, respectively, provide important indices for estimating global sensitivities, and therefore the relative importance of individual inputs through the sampled space.

$$\mu_i = \frac{1}{M}\sum_{j=1}^{M} EE_i^j \quad \text{Eq. 8}$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{M}\left(EE_i^j - \mu_i\right)^2}{M}} \quad \text{Eq. 9}$$

$$u_i^\star = \frac{1}{M}\sum_{j=1}^{M}\left|EE_i^j\right| \quad \text{Eq. 10}$$

The efficiency in Method of Morris is derived from developing trajectories for input perturbation that sample the selected input space with the fewest number of model evaluations. We applied the method of Ruano et al. (2012) that selects a subset of trajectories optimized from a pool of many candidate trajectories. We selected $M = 40$ trajectories from 1000 candidates. The Method of Morris was applied here using the Sensitivity Analysis Library (SALib v1.3) for Python (Usher et al., 2020). Six parameters were varied over plausible ranges and applied uniformly across the basin (Table 4). We evaluated the three sensitivity indices μ, σ, and $\mu^\star$ to a single model output representing the decline in nitrate export to the GoM at 100% adoption of the WRP (Figure 19). On Figure 19, wedges describe whether a parameter had a relatively

independent control on the metric of interest (monotonic or almost monotonic), or if it behaved strongly non-linearly across parameter-space or interacts with other parameters (not monotonic). The constant describing uptake velocity at a nitrate concentration of 1 mg $N\text{-}NO_3$ L^{-1} was the most sensitive parameter of the six tested and justifies focusing on the univariate sensitivity of this parameter. The second most sensitive parameter is χ_{buffer} which creates a linear increase in the area available for wetland surface area for a given amount of restoration and the value of 0.777 selected is appropriate to protect wetland landscapes from soil infilling and storm damage (Christianson et al., 2013).

Table 4: Description of parameters varied in sensitivity analysis. Citations as follows: C2013 – Christianson et al. (2013), M2018 – Marshall et al. (2018), KK1995 – Kadlec and Knight (1995), I2008 – Iovanna et al. (2008), T2010 – Tanner et al. (2010), W2014 – Wollheim et al. (2014), M2009 – Mulholland et al. (2008), CB2017 – Cheng and Basu (2017), BK1993 – Beven and Kirkby 1993.

Symbol	Function	*A priori* value	Tested range	Source
χ_{buffer}	Fraction of restored area needed as buffer ($m^2\ m^{-2}$)	0.777	0.62,0.93	C2013,M2018
χ_{up}	Maximum multiplier of treatable upstream area ($m^2\ m^{-2}$)	22	5.5,88	I2008,C2013,M2018
d_{max}	Maximum depth of flow within restored wetland (m)	0.8	0.2,2.5	KK1995,T2010
α_{wl}	Hydrologic time constant within active wetland store (d^{-1})	0.06	0.024,0.08	CB2017
ϕ	Hydraulic geometry exponent (width to discharge) (-)	0.51	0.48,0.62	BK1993
b_{vf}	Uptake velocity ($m\ y^{-1}$) a DIN concentration of 1 mg L^{-1} (-)	18.	4.,56.	W2014 (M2009-W2014)

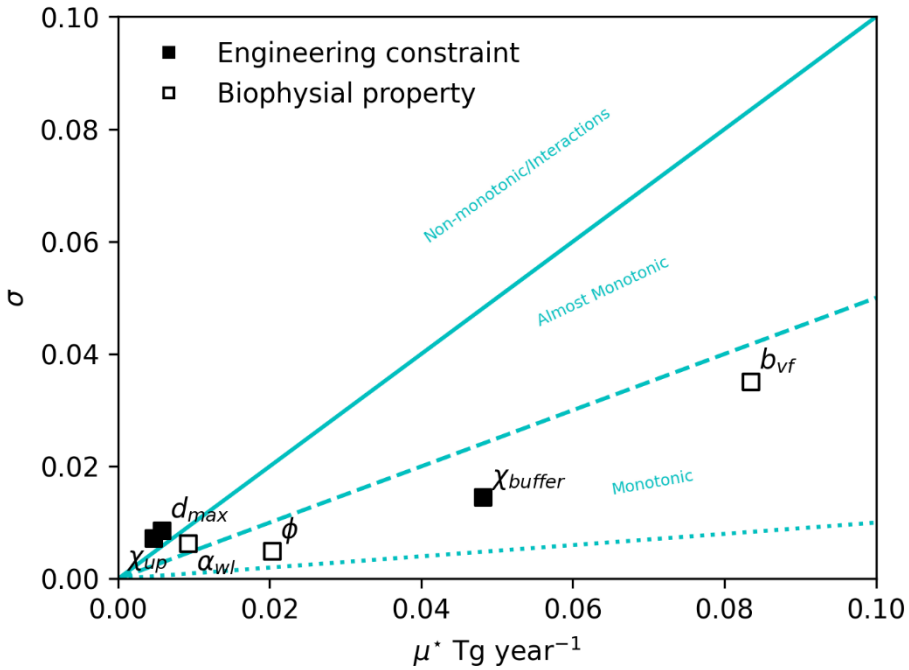


Figure 19: Morris Method screening results depicting standard deviation of elementary effects against mean of absolute elementary effects for each tested model parameter for engineering constraints and biophysical properties. Elementary effects are calculated as the reduction in DIN export to the GoM between the baseline and 26% restored wetland conditions. Parameters are described in Table 4.

Findings

We found that wetland restoration through existing federal programs that incentive the restoration or construction of field-margin wetlands could not reduce nitrate by the 45 to 60% needed to restore ecosystem health in the Gulf of Mexico even with assumptions that should maximize the treatment of nitrate runoff (Figure 20). At complete adoption, FWP and WRP wetlands could reduce nitrate export by 6.2% and 27%, respectively. These levels of restoration increased denitrification by 85% and 430% over removal by natural wetlands. Assuming nitrate leachate rates would decrease to rates of naturally vegetated landscapes such as grassland or forest if crops were fallowed, nitrate reduction from the two programs was 3.4 (FWP) and 3.0

(WRP) times greater than just retiring the same cropland (Figure 20). These high ratios of export reduction from wetlands relative to crop retirement illustrates why wetland restoration is an ideal intervention for mitigating nutrient pollution in highly agricultural basins. Simple extrapolation of wetland denitrification to retirement reduction ratio suggests that by reducing crop leachate 20% through restoration could meet target reduction of 60% and there seems to be sufficient restorable lands to meet this reduction (Horvath et al., 2017). Prior studies have suggested that wetland restoration alone could potentially reduce nitrate to targets (Mitsch et al., 2001; Cheng et al., 2020), but we find that restoration following the WRP and FWP cannot. The geospatial separation of restorable wetlands with croplands, reduced denitrification during storms, and leaching to deeper flow-paths through the surface that bypass wetlands all contributed to reducing the simulated nitrate removal below their potential values.

Potentially restorable wetlands exist throughout the MRB; however, there remained sufficient geospatial separation between restorable areas and existing crops to limit the amount of runoff intercepted and treated. Of the 0.79 Tg y^{-1} of nitrate runoff via surface flow-paths from agricultural lands, treatment wetlands restored by the FWP intercepted 0.19 Tg y^{-1}, 29% of intercepted nitrate was denitrified, and 57% of nitrate runoff bypassed both natural and restored wetlands (Figure 21). Bypassing flow in both programs was a consequence of insufficient restorable wetlands (Horvath et al., 2017) in pixels containing crops (Figure 21). The relatively modest 6,000 km^2 of wetlands restored through the FWP would intercept runoff from 121,000 km^2 of cropland. Nitrate runoff intercepted via this program averaged 21 kg $N\text{-}NO_3$ ha^{-1} y^{-1}, and therefore, denitrification added by this program exhibited areal denitrification rates (37 g $N\text{-}NO_3$ m^{-2} y^{-1}) which were consistent with high observed values for treatment wetlands (Mitsch and Day, 2006). Despite high areal rates of denitrification, much higher than the average

denitrification rate from treatment wetlands restored by the WRP (13 g N-NO_3 m^{-2} y^{-1}), the mass fraction of nitrate denitrified by FWP wetlands was lower than typical assumptions in macroscale analyses. The relative effectiveness of this program reflected the concentrated effort to intercept and treat subsurface drainage effluent, previously identified as being the largest single source of nitrate to the coast (Goolsby et al., 2001).

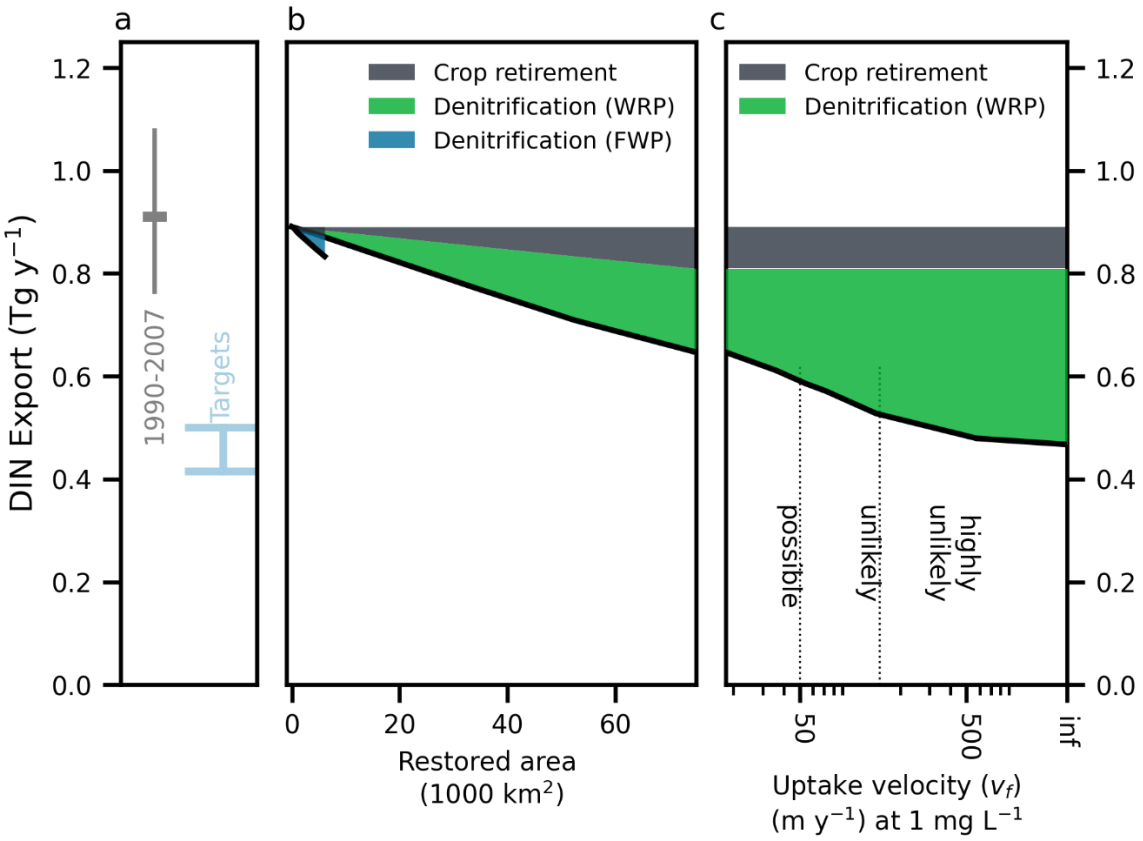


Figure 20: Export of nitrate from the Mississippi River Basin. Observed and target (Scavia et al., 2017; Mississippi River/Gulf of Mexico Watershed Nutrient Task Force, 2008) nitrate export (a). Nitrate export following adoption of two wetland restoration programs (b). Color patches depict the reduction in export attributable to reduced leaching as croplands are retired, and additional denitrification within restored wetlands. Nitrate export following complete adoption of the Wetland Reserve Program across a range of wetland denitrification rates (c).

Restoration through the WRP provided about half of the target nitrate reduction to the Gulf of Mexico. Following complete adoption of the WRP, 0.47 Tg y^{-1} nitrate was intercepted, 48% of that nitrate was denitrified, and 20% of nitrate in surface runoff bypassed wetlands (Figure 20). The WRP removed only 4 times more nitrate than the FWP despite the WRP restoring more than 12 times the cropland and removing about 12% more of nitrate intercepted (Figure 21). WRP wetlands removed a greater fraction of intercepted nitrate because catchments were smaller making flows lower and exposure to denitrifying benthic surfaces greater. Furthermore, intercepted leachate was derived from a greater proportion of croplands of lower intensity, which decreased average intercepted nitrate concentration. Because wetlands followed an efficiency loss relationship with nitrate concentration, a decreasing denitrification rate with increasing nitrate concentration (Wollheim et al., 2014), they exhibited higher proportional denitrification at low leaching intensity. In contrast, the FWP targeted restoration at intensively leaching crops (Figure 21b) and targets treatment of more crop area per restored area. Given a choice in restoring wetlands such that a greater area could be treated, or a smaller area treated with a higher fraction of inputs removed, we show greater denitrification will always occur where more mass was treated with a decreasing fraction of removal (Appendix III).

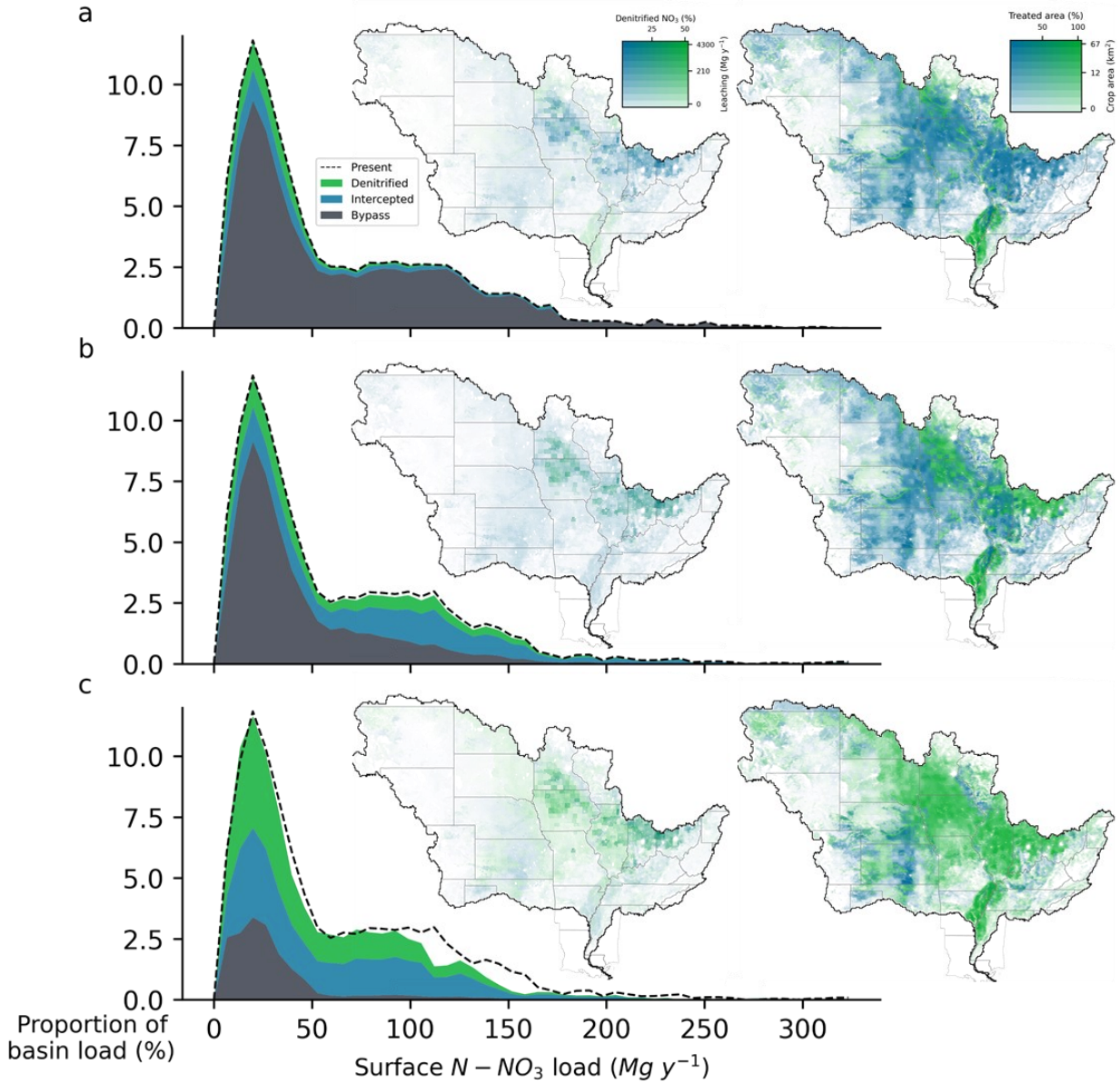


Figure 21: The geospatial distribution of restorable wetlands cannot intercept all nitrate runoff. Distribution of nitrate yield and fate of nitrate prior to entering the river is presented as density plots (left column) and maps (center) for present day existing wetlands (a), treatment wetlands restored through the Farmable Wetlands (b), and Wetland Reserve Programs (c). The distribution of treated and untreated croplands for three scenarios (right).

Our analysis suggests a more modest expectation of the role of wetland restoration on reducing hypoxia than other analyses. Previous scaling studies (Cheng et al., 2020; Cheng and Basu,

2017) have found substantially greater removal than estimated here when higher fractions of leachate (e.g. 30-50%) (Cheng et al., 2020) are assumed to be intercepted by restored wetlands. AgroIBIS-WBM was unable to achieve satisfactory correspondence with observed nitrate flux in rivers with parameterizations that created such high levels of connectivity between cropland leachate and river flows. A large fraction of nitrate percolating towards subsurface removal or long-term storage was needed for consistency with observed fluxes (Figures 17 and 18), which precluded treatment of this mass at field-margins. Assumptions from the FWP and WRP reflected in this analysis, as well as in integrated assessments (Marshall, 2018; Hansen et al., 2021) find even more modest opportunity for field-margin wetland denitrification to mitigate GoM export than we report here. Studies that assume wetland restoration cannot intercept subsurface-drainage intercepted only an additional 4% of catchment area to wetlands in the Upper Mississippi River Basin (Evenson et al., 2021). By neglecting leachate through defined areas of subsurface-drainage (Sugg, 2007), 40% of MRB leachate, predominately in the corn-belt of the midwestern US, is left untreated thereby eliminating treatment of the most concentrated source of nitrate in the basin (Goolsby et al., 2001).

Seasonality and storm-scale dynamics pose an additional constraint on whole-watershed nitrate removal. The seasonal peak in the capacity of wetlands to remove nitrogen did not coincide the highest rates of nitrate flux into wetlands (Figure 22). The peak of the basin-average flux to restored wetlands occurred in May, when denitrification removed only 26% (FWP) and 39% (WRP) of intercepted nitrate, and only 10% (FWP) and 31% (WRP) of nitrate in surface runoff during the month. Nitrate mass bypassing wetland treatment when water depth exceeded typical or design depths (Tanner et al., 2010) totaled 6.9% (FWP) and 3.2% (WRP) of total annual runoff, and the majority of this flux bypassed wetlands during the late winter and spring (Figure

22). Seasonally lower denitrification due to low water temperatures during the winter and spring accounted for 17% of the higher seasonal nitrate runoff. During periods of constant high runoff in spring, wetland depth was consistently higher within wetlands, which increased the hydraulic load in the wetland and reduced contact between nitrate dissolved within the water column with the benthic surface thereby reducing denitrification (Figure 23).

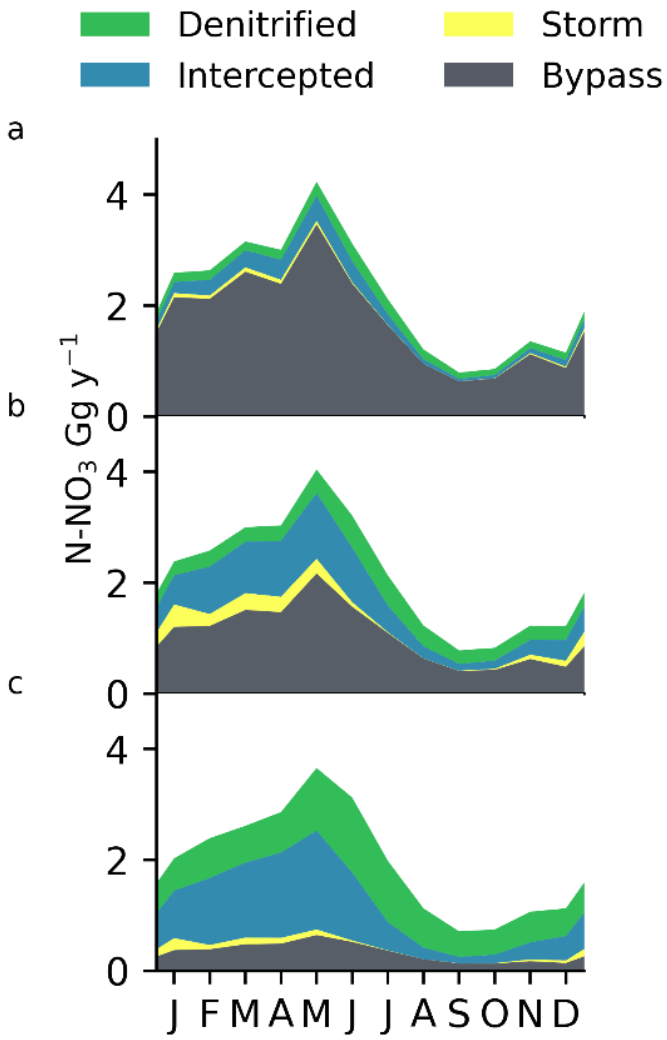


Figure 22: Annual peak leachate flows precede seasonal wetland denitrification maximums. Monthly means of total basin fluxes for present-day unrestored wetlands (a), complete adoption of the FWP (b), and complete adoption of the WRP (c). Sum of four fluxes represent total surface nitrate runoff from croplands. Storm bypass (yellow) represents mass entering wetlands when water depth exceeded the specified maximum depth.

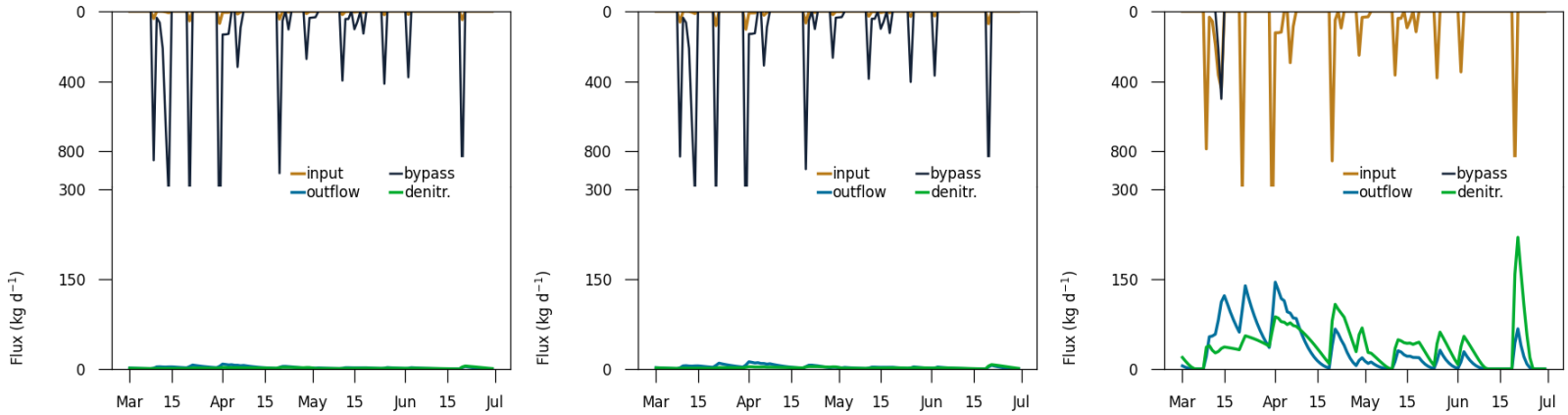


Figure 23: Example time-series of wetland states and fluxes within a single pixel near West Lafayette, Indiana with no restored wetlands (a), complete adoption of the FWP (b), and complete adoption of the WRP (c

The importance of representing temporal dynamics in scaling estimates of treatment wetland effectiveness was demonstrated for croplands without tile-drains (Evenson et al., 2021). The significantly greater amount of mass bypassing wetlands and lower denitrified fraction in the FWP program than in the WRP program reflected the greater sensitivity of denitrification of tile-drain runoff to storm-scale dynamics, because of fast runoff from tile-drained lands. When considering such storm-scale dynamics, we find that total annual wetland denitrification for the FWP is marginally lower than typical estimates of 40-50% denitrification of intercepted input mass; however, mean annual denitrified fraction of intercepted mass for WRP is very consistent with typical estimates.

When considering the fraction of flow-paths intercepted by field-margin treatment wetlands, coastal hypoxia targets could be met by wetland restoration alone only if denitrified fraction approached 100% (Figure 20). Intrinsic denitrification rates are input into WBM as estimates consistent with rivers (Mulholland et al., 2008) and typical natural wetlands (Racchetti et al., 2010; Cheng and Basu, 2017); however, both natural and treatment wetlands have exhibited much higher rates in prior field studies (Wollheim et al., 2014; Bachand and Horne, 1999). Sensitivity of the model through the range of physical plausibility was tested (Figure 20c), following verification that the parameter was the most sensitive to export reduction (Figure 19). Even at values exceeding physically realistic limits of about 400 m yr^{-1}, complete adoption of WRP was able to only reduce nitrate export by 48%. A substantial fraction of nitrate mass is directed towards the subsurface in our conceptualization, less so in areas with subsurface-drainage (Figure 24). Nitrate entering the subsurface is assumed to enter a shallow groundwater pool experiencing denitrification at fixed rates (Green et al., 2008) or return to streams (bypassing wetlands), or is removed from the surface water network. Between nitrate bypassing

restored wetlands and loading from baseflow, as well as inputs from natural vegetation and other anthropogenic sources (including domestic waste, but excluding livestock waste, which was not included in our analysis), export remained substantial at about 0.5 Tg y^{-1}. The fraction of total leachate removed from further surface water interactions is the dominant flux in the simulations equaling between 2.0 and 2.1 Tg y^{-1} or about 58% of all leachate (Figure 25). The flux was defined by a single parameter and was the most sensitive parameter in controlling model biases with respect to observed riverine fluxes (Figure 17). Potential reservoirs for storage of this mass include unsaturated soils below the rooting zone, potentially responsible for storing up to 1.8 Tg y^{-1} over the conterminous US (Ascott et al., 2017), or increases in groundwater storage associated with increasing groundwater concentrations (Puckett et al., 2011), potentially responsible for storing up to 2.4 Tg y^{-1}. Importantly, a substantial fraction of such legacy storage in groundwater would be expected to discharge to rivers over long timescales (Dupas et al., 2020). Storage of 3.5 Tg y^{-1} of nitrogen within the rooting zone of the MRB (Van Meter et al., 2018; Meter et al., 2016) is another potential storage mechanism; however, a consistent rate of storage in the rooting zone is captured by AgroIBIS. Characterizing the role of long-term storage of subsurface nitrogen, any transformation reactions occurring in the subsurface, and the return of this contaminant to surface water at long time-scales, may be the single most important consideration defining the success of current nutrient reduction strategies (Van Meter et al., 2018; Basu et al., 2022).

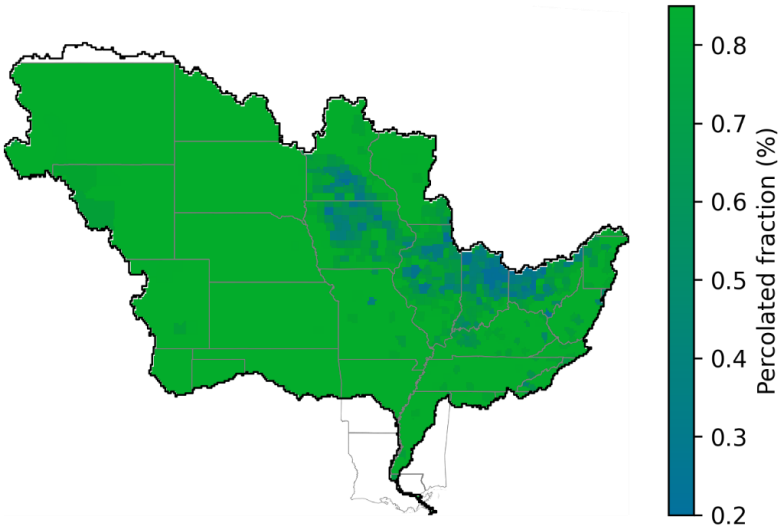


Figure 24: Infiltrated fraction throughout the domain varies based on presence of subsurface (tile) drainage.

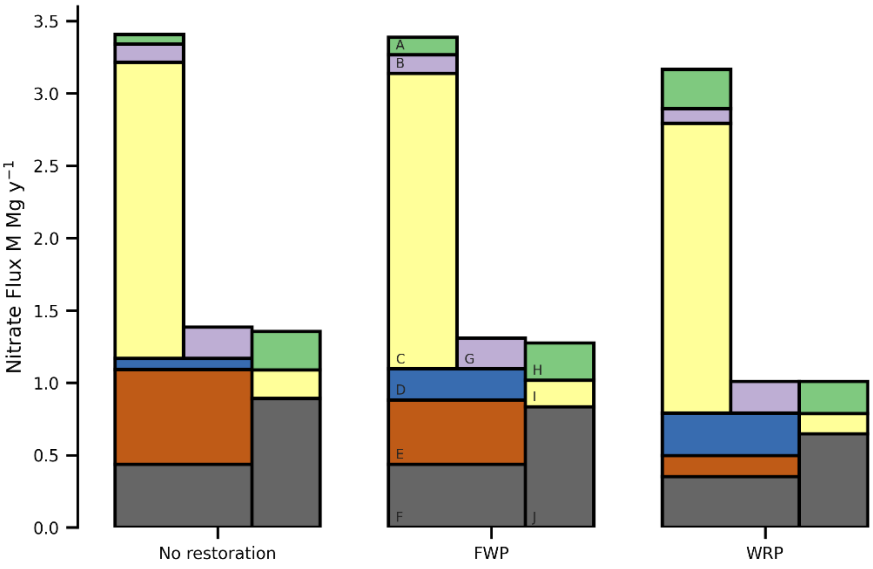


Figure 25: Distribution of nitrate fluxes throughout simulations for the conditions of no restoration, and complete adoption of the Farmable Wetlands Program (FWP), and Wetland Reserve Program (WRP). Fluxes labeled in the central image correspond to: A – denitrification within all wetlands, B – denitrification within shallow groundwater, C – storage or removal in other subsurface pools, D – flow out of wetlands to streams, E – flow bypassing wetlands to streams, F – baseflow nitrate from shallow groundwater to streams, G – other domestic inputs to streams from non-point source suburban development and waste water treatment plants, H – denitrification within streams, I – nitrate extraction from streams via water abstractions, and J – nitrate export to the Gulf of Mexico.

Conclusion

Restored wetlands in our analysis are limited to crop margins, meaning that treated mass is generated proximal to any restored wetlands. Complementary findings show that restored fluvial wetlands that treat leachate from all upstream croplands are substantially more effective than crop-margin wetlands (Hansen et al., 2021). Studies that make different assumptions about buffer areas, position on the landscape, or focus on wetlands receiving inflows as part of downstream riverine transport (Hansen et al., 2018; Czuba et al., 2018; Hansen et al., 2021) differ in target restoration practice from the programs analyzed here, and find treatment wetlands far more effective when treating larger contributing areas. The entire spectrum of connectivity of wetlands to upstream river flow, to subsurface-drainage effluent, and to undrained cropland runoff exists within the MRB, and representation of such heterogeneous connectivity in scaling studies of nutrient reduction demands novel approaches beyond what has been possible to date. However, existing policy programs investigated here assume direct connectivity to cropland runoff by field-margin intervention and are limited in their effect to about half of the required nutrient reduction to protect the GoM. We should prioritize policies that facilitate interventions that make use of larger contributing areas such as fluvial wetlands (Hansen et al., 2021) or floodplain restoration (Kroes et al., 2015), especially because such downstream interventions have greater capacity to intercept and treat legacy nitrate that returns to the river network via groundwater discharge far from croplands where it originated (Basu et al., 2022), while maintaining dialogue with stakeholders to address how conservation measures benefit and impact local communities (Gourevitch et al., 2020). Field-margin treatment wetlands will remain an important approach towards mitigating Gulf of Mexico hypoxia, because they have the

capacity to reduce a meaningful fraction of nitrate and have existing conservation mechanisms in place to facilitate adoption.

CHAPTER IV

CONCLUSION

The three studies presented analyze a cross-section of policy interventions, and the potential efficacy of each reflects the constraints imposed not only by the degree or magnitude that the intervention can be applied, but also by the geographic distribution over which they can be applied. It is not common for literature on watershed management to acknowledge such geographic constraints, and distributed modeling of watershed management as shown here can represent and explore such constraints directly. In this conclusion, the nascent scientific discipline of sociohydrology, which treats human decision-making as a fully integrated part of the hydrologic cycle, is argued to be an appropriate framework for research characterizing the efficacy of watershed management and restoration. A background on the philosophy of experimental design used here that applied models in sociohydrologic studies is drawn from futures science and the study of social ecological systems. Finally, a framework is described that provides a simple lens on how interventions in watershed management complement one another that can be used to evaluate needs in management practice.

Much of the most critical research foci in the geosciences occur at the intersection of Earth system processes and human society (Blöschl et al., 2019). Humankind can affect but not altogether change natural processes, so our ability to affect sustainable positive change on the Earth system is largely confined to our decisions about watershed management. Watershed management entails the information gathering, goal-setting, and policy-making that we use to facilitate the sustainable use and conservation of the ecosystems that provide the services we need to thrive (Gregersen et al., 2007). Representation of water management activities through hydrologic modeling is both a critical aspect for understanding hydrology (Nazemi and Wheater, 2015), as well as providing sound understanding of the hydrologic phenomena to facilitate sound management (Gregersen et al., 2007).

The discipline of sociohydrology (Sivapalan et al., 2012) acknowledges the endogenous role of humans in the hydrological cycle . The discipline captures the ways in which people respond to hydrologic changes by altering watershed management strategies (Di Baldassarre et al., 2015, 2013; Chang et al., 2014; van Emmerik et al., 2014). Sociohydrological studies aim to include human decision-making in hydrologic functioning. Models of human responses to date generally lump geographic variability in both watershed characteristics, human values, and importantly, management response (Di Baldassarre et al., 2013; van Emmerik et al., 2014). Other studies have sought to identify the geographic distribution of human responses to hydrologic alteration from coarse watershed-scale interventions. Economically appropriate changes in crop selection were a modeled response to reservoir management policies (Giuliani et al., 2016). People's decisions on home values appear to be partly informed by local policies and outcomes of stream riparian area protection (Chang et al., 2014). Optimal selection of managed aquifer recharge and surface water irrigation practices reflected both economic and geologic controls in the lower Mississippi River basin (Tran et al., 2019, 2020). Although the ultimate goal of sociohydrology is to incorporate the endogeneity of humans in the hydrologic cycle, studies focusing on one-way causalities such as human effects on watersheds, or human response to watershed change, are valid lines of inquiry and necessary to progress the discipline (Troy et al., 2015).

Empirical or observational studies that examine hazards such as pollutants or ecosystem changes tell us plainly why we must change our interactions with the Earth system. For example, excess nutrients applied throughout watersheds are mechanistically linked to observed changes in marine ecosystem structure (Breitburg et al., 2018), just as elevated chloride salts are linked to changing freshwater aquatic ecosystem structure (Cañedo-Argüelles et al., 2016, 2013; Hébert et al., 2022). Intensification of droughts in the US southwest (Williams et al., 2022) combine with

overdraft of water (Tidwell et al., 2014) to present a potentially hazardous strain on watersheds' potential for provisioning water. These studies provide examples of primarily empirical research that identifies problems and mechanistic interpretations of their effects.

Though empirical studies that establish how societal changes reflect watershed management are necessary (Sivapalan et al., 2012), so are experiments using environmental models that facilitate studies not possible through direct observational and empirical methods. The use of simulation models throughout the natural sciences incorporates aspects of both theoretical and experimental scientific methods (Dowling, 1999); model structures are formulated *in silico* according to a hypothesized or theoretical concept of processes, and then applied as instruments of falsification to perform experiments. The inference gained from such experiments is then extended to the material world (Beven, 2002).

The management practices investigated in each of the three studies comprising this Dissertation each are intended to provide a net benefit to the health and happiness of near future societies, but tradeoffs exist and have the potential to materially harm humans. Direct experimentation of these practices come with ethical concerns and need debate and dialogue with affected and involved communities. Experimentation *in silico* comes with less direct impact on communities (Peck, 2004). Both empirical and modeling-based studies are both subject to ethical concerns when the understanding gained informs decision-making that may ultimately harm people or an ecosystem. Direct empirical evidence of intervention efficacy can suffer from ethical concerns but can also lack direct causal inference. In contrast, experimentation with models has obvious limitations as to how well we can map reality with model representations (Beven, 2002); however, causality is addressed explicitly in modeling analyses, albeit hypothetically. Decisions that result from hypotheses and assumptions in the model structure therefore reflect any potential

biases of the model creator, and therefore themselves carry important ethical considerations (Almada and Attux, 2018).

Modal narratives as experiments

Beyond ethical concerns for utilizing models for sociohydrologic investigation, empirical approaches are only available for current or recent phenomena. Experiments using simulation models are virtually required when considering alternatives at future times (De Jouvenel, 2000) and deep [pre-]historic time (Perry et al., 2016). To frame the experimentation within each study presented here, each hypothesis can be viewed as a *modal narrative*, or a description of a possible alternative condition targeted to evaluate effects of some management decision, or tradeoff between competing decisions (Booth et al., 2009). Following the distinction of Booth et al. (2009), model experiments set within a historic context are described as *counterfactuals* and those set within a future time-frame as *scenarios*.

Each of the presented studies developed counterfactuals or scenarios to understand consequences of human management of watersheds, or human management of processes with direct consequences on watersheds (Table 5). The associated *modal narratives* span several classifications in the typology of Börjeson et al. (2006) which are defined by whether factors defining the narrative are controlled within the system or without, and whether they answer questions about probability, possibility, or preference (Table 6). These studies focus on a limited range of potential motivating forces that exemplify specific decisions about management rather than the completely realized and comprehensive scenarios that were part of the larger research programs the presented studies were associated with (e.g. Borsuk et al., 2019; Cronan et al., 2021; Mavrommati et al., 2017; Samal et al., 2017). In the larger research program, transdisciplinary groups of scholars and stakeholders collectively built scenarios considering

unique sets of variables, target objectives, and using application specific methodology, while drawing on the work presented here.

Table 5: Modal narratives developed in each study presented as part of this dissertation. Clauses inside brackets and separated by pipes [|] denote alternative practices under examination.

Chapter	Description
1	Future managers [reduce | do not reduce] road salt application to recommended values. Future zoning requirements [force infilling development | favor large building lot sizes | remain constant]. Climate follows from [moderate and declining | constant high] emissions pathways.
2	Irrigation used [existing | increasingly efficient] technology. Enhanced aquifer recharge [was never instituted | was sufficient to stabilize aquifer head].
3	Wetlands [were | were not] restored through increasing degrees of adoption of the [Farmable Wetlands Program | Wetland Reserve Program]

Table 6: Typologies of modal narratives as classified by Börjeson et al. (2006) and examples of scenario building used in Dissertation research and supporting research programs.

Category	Types	Descriptive	Used in my work
Predictive (probable)	Forecasts	Expected result pre-conditioned on the most likely future state	
	What-if	Expected result preconditioned as bifurcating along a specified dimension	Chapter 3
Explorative (possible)	External	Possible result influenced by changing external factors	Samal et al. 2017, Cronan et al. 2021
	Strategic	Possible result preconditioned on a specific action under the influence of changing factors	Chapters 1, 2
Normative (preferrable)	Preserving	Target result achieved by manipulating an existing system	
	Transforming	Target result achieved by fundamentally altering the system to facilitate goal meeting	Chapter 2

Summary of the Model Experiments

The three studies presented here vary considerably in direct topic of discourse, which provides a cross-section of problems in watershed management. Each study included in my research program uses a version of the University of New Hampshire Water Balance Model (WBM) or

the related Framework for Aquatic Modeling of the Earth System (FrAMES). The core functionality of the model is documented elsewhere (Grogan et al., 2022). Each chapter describes how the model developed for the specific topic relates to and departs from this core functionality.

In Chapter 1, a simulation model of chloride within watersheds investigated the potential for road salt application to affect aquatic biota (Zuidema et al., 2018). Comparison of model estimates and observed measures of chloride in stream water were used to solve an inverse problem to determine an expected rate of road salt application. Exploratory scenarios illustrated the expected influence on chloride impairment of streams from changing road salt application rates and buildout patterns using projected urban intensities from a detailed land-cover change and scenario analysis (Thorn et al., 2017). Each of these scenarios represented a narrative where build-out follows one of three possible paradigms, road salt application rates remain constant or are reduced, and climate warming follows from one of two emissions trajectories (Table 5). The scenarios were used to investigate the relationship between managing build-out strategy and road salt application rates in mitigating harm to aquatic species.

In Chapter 2, a model of agricultural water reuse in the Upper Snake River Basin (USRB) in southern Idaho investigated the interplay of changing irrigation technologies and enhanced aquifer recharge on surface and groundwater availability for agriculture (Zuidema et al., 2020). In the manuscript, strategic exploratory counterfactuals assessed changes in riverine flow and aquifer head as irrigation technology was progressively modernized. Each of these individual counterfactuals was paired with a transforming normative counterfactual that required that aquifer head remain unchanged according to policy decree (IDWR, 2015). The combined evaluation of the two forces dictating future water availability in the USRB, modernizing

irrigation technology to reduce abstraction from the river and enhanced aquifer recharge to minimize drawdown of the aquifer, illustrated how competing policies dictating instream flow (IWRB, 1985) and aquifer drawdown (IDWR, 2015) are compatible at rates of enhanced aquifer recharge considered feasible (IWRB, 2016).

In Chapter 3, a model of treatment wetland processing of field-margin crop leachate throughout the Mississippi River Basin explored the efficacy of two federal programs (Conservation Reserve Program, 2015; USDA, 2021) for restoration of wetlands to reduce nitrate in riverine export to the Gulf of Mexico. A suite of predictive what-if counterfactuals for each program captured expected system response to progressively greater levels of adoption for each program. Geographic and engineering constraints on wetland design limited the amount of nitrate reduction that was possible. A comprehensive accounting of the entire suite of flow-paths through the watershed characterizes a lower estimate of the benefits of this intervention compared to other scaling studies (Mitsch et al., 2001; Cheng et al., 2020).

Epistemology of the evidence presented

The degree to which inference from simulation experiments qualifies as scientific observation is necessarily entangled in value judgements about how 'good' the model was that generated the results. The problem of model *validation* goes beyond the semantics with which we describe the correspondence of our models with observational data to the epistemological underpinnings of what we can learn from them (Barlas and Carpenter, 1990; Beven, 2002; Kleindorfer et al., 1998; Konikow and Bredehoeft, 1992; Oreskes et al., 1994). Ultimately simulation models are best considered a heuristic tool that facilitates experimentation where observational data are unavailable as is the case when developing modal narratives to assess potential effects of policy decisions and degrees of adoption of a specific management practice. Regarding the semantics,

the term *validation* is widely used in the scientific literature in describing the correspondence between a model and comparable observations (Eker, 2018) and the studies presented in Chapters 1 through 3 follow this convention. It is generally understood that terms like *validation* or *verification* are inconsistent with their epistemological meaning and should instead be read as *corroboration* (Oreskes et al., 1994).

Since the middle of the 20th century, two prevailing doctrines in the philosophy of science define an axis of views towards the use of models for creating knowledge (Kleindorfer et al., 1998; Barlas and Carpenter, 1990; Reed et al., 2022). From a foundationalist (or objectivist) perspective, models are heuristic tools for constructing theories that challenge the limits of empirical data despite absolute falsification in open environmental systems being impossible (Oreskes et al., 1994; Beven, 2002, 2006; Konikow and Bredehoeft, 1992). A relativist perspective considers the conversational aspect of integrating theoretical components within models to provide meaningful insight. Modern scientists engaged in modeling value both perspectives (Eker, 2018) and the strategy for model development used here reflects this common approach. A variety of model calculated metrics were compared to relevant observations throughout the development process to corroborate the structure and behavior of the model; but the model was then applied as a heuristic to explore hypothesized conditions over which no possible corroborating data could be generated.

The corroboration between model and observation leads directly to questions of uncertainty. The tapestry of uncertainties in representing open environmental systems described by Vrugt and Sadegh (2013) include uncertainties on parameters, input data, initial states, model structure, outputs, and updated states. The Approximate Bayesian Computation framework (Sadegh and Vrugt, 2013; Vrugt and Sadegh, 2013) and the Generalized Likelihood Uncertainty Estimation

(Beven and Binley, 1992; Stedinger et al., 2008) each attempt to define and account, respectively, for the suite of uncertainties in hydrologic modeling. Successful applications are generally limited to simple model structures in conceptual rainfall-runoff models over nearly pristine catchments. The Framework for Understanding Structural Error (FUSE - Clark et al., 2008) has similar limitations and goals; however, is more focused on comparing a structured formalization of possible conceptual hydrologic models. The approach taken by these practitioners strives for high explanatory power with fairly limited depth to explain a diversity of processes or conditions (Beven, 2002).

The studies presented here target a much deeper set of processes than have been treated historically with full uncertainty characterization. A valid criticism of the research presented is the absence of formal and complete characterization of uncertainties on the core predictions of the experiments. Arguments against formal uncertainty characterization focus on limited computational resources; an argument that has been losing validity with contemporary computing resources. Still computation would be problematic. Typical simulation times across my studies ranged from 6 hours to 7 days for a single evaluation. With spatially varying parameters and multiple input data sets, millions of model evaluations would be needed for proper characterization in each study (Reed et al., 2022) equating to decades of computation time (and tons of greenhouse gas emissions). The exploratory value of these studies was instead supported by deductive analysis of each models' limitations. Full uncertainty characterization and quantification would be recommended for models used to answer direct questions relevant to decisions on any topic considered here.

Each study developed substantial new simulation capacity targeting unique aspects of human interaction with the Earth system. History-matching (Konikow and Bredehoeft, 1992) to

corroborate model characterization with observational data was an integral part of the developmental process of each study. The level of sophistication of history-matching reflected the needs of each individual study, from simple univariate sensitivity analyses (Chapter 3), to Approximate Bayesian Computation (Vrugt and Sadegh, 2013) to estimate a specific parameter value from joint probabilities of relevant parameters (Chapter 1). Following manual calibration and sensitivity analyses of hypothetically important parameters, global optimization algorithms (Wales and Doye, 1997) were used parameterize the model, or to determine whether the selection of parameters could produce a calibrated model (Chapter 1 and 2). More recently (Chapter 3), evaluation of whether specific parameters provided sufficient control of the model system was performed through formal sensitivity analyses (Morris, 1991; Ruano et al., 2012).

Discussion and Conclusions

The results from the presented modeling experiments each suggest difficulties in realizing success in watershed management when accounting for complex geographic tradeoffs. Such difficulties are common within the practice of watershed management (Gregersen et al., 2007). Examining the role of tradeoffs between competing (Chapter II) or complementary (Chapter I) interventions is an important role for modeling. Such what-if analysis is useful for identifying likely system constraints and uncertainties. The experiments contextualized the effect of decisions and suggest the levels of investment that must be made to approach target outcomes. In each case, meeting the motivating targets (reducing chloride impairment, maintaining stable aquifer levels, or reducing nitrate in riverine export) was achieved at the extreme adoption of each tested interventions when applied according to the constraints and assumptions imposed by the model. Each study identified key uncertainties that will need consideration for successful watershed management.

The ability for each intervention to meet its target reflected how well the tested intervention or suite of interventions addressed the *intensity* and the *extensity* of the target concerns. The *intensity* of an intervention refers to the magnitude that the problem in watershed management is addressed at a point in space, such as the reduction in deicer application rate, or the amount of increased classical irrigation efficiency associated with a new suite of irrigation technology. The *extensity* of an intervention refers to the portion of the generative source of the problem it can be applied, such as limiting the construction of new impervious surfaces, or the geographic coverage of potentially restorable wetlands. A graphical representation of how the two concepts relate to one another is presented as Figure 26. The aspects of each presented study that dictate the intensity and extensity of tested interventions are summarized by Table 7. The intensity and extensity of interventions tested in Chapters 1 and 2 were sufficient to meet target conditions, whereas they were too constrained in Chapter 3 for completely meeting target reductions in nutrient export.

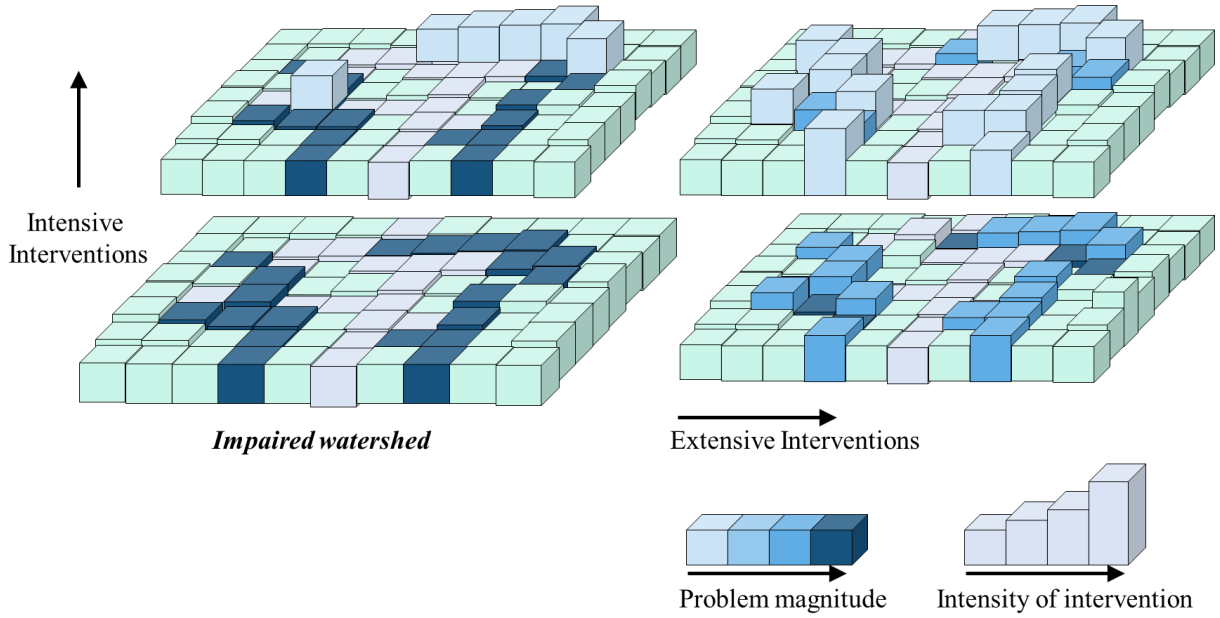


Figure 26: Relationship between the intensity of an intervention and the extensivity of the adoption of an intervention along hypothetical landscapes. Each of the four panels illustrates a portion of the landscape with some arbitrary problem, the impact of which is indicated by the deepness of the blue color. The intensity of an intervention that reduces the impact of the problem is depicted as the relative height of boxes making up the landscape. The geographic extent or coverage of the intervention is depicted along the horizontal axis. Interventions that are adopted intensively and over an extensive area (top right) virtually eliminate the problem from the watershed.

Table 7: Interventions tested and factors or assumptions that controlled the intensity and extensity of mitigation effects. Descriptors of intensity and extensity scale include *weak* (less than 20% change in problem attribute, either contaminant or inefficient water loss, or extent of problem coverage or flux), *moderate* (20% to 60% change in attribute or extent), and *strong* (greater than 60% change in attribute or extent).

Chapter	Goal	Intervention	Intensity control	Extensity control
1	Successfully reduced chloride impaired river length from present-day	Deicer application rate	Strong intensity control. Direct and proportional reduction in contaminant.	Applied uniformly, all impervious surfaces are affected consistently with intensity control.
		Zoning, new development minimizes or adds impervious area (Thorn et al., 2017)	None	Strong extensity control. Infilling limits new impervious area. Lot size minimums adds impervious area.
2	Successfully stabilized aquifer level while maintaining river flow above regulatory limits	Irrigation technology modernization	Strong intensity control by reducing irrigation water demand, but negative feedback by limiting incidental recharge	Applied uniformly, all irrigated croplands are affected consistently with intensity control
		Enhanced aquifer recharge	Strong intensity control by directing increasing aquifer level, but negative feedback by reducing river flow	Weak extensivity control, limited to extent of groundwater irrigators, negative feedback to downstream users
3	Unsuccessfully reduced nitrate export from Mississippi River by 45 to 60%	Wetland restoration following the Farmable Wetlands Program	Weak intensity control by partially reducing nitrate runoff from only highest leaching croplands	Weak extensivity control due to low geographic coherence between nitrate runoff and restorable wetlands. Did not intercept all flow-paths
		Wetland restoration following the Wetland Reserve Program	Weak to moderate intensity control by partially reducing nitrate runoff from croplands across whole distribution of leaching rates	Moderate extensivity control due to wide geographic coherence between nitrate runoff and restorable wetlands. Did not intercept all flow-paths

In Chapter 1, mitigation was achieved by pairing strong intensity and extensity controls through reduced road salt loading and zoning restrictions (Figure 27). For some combinations of tested scenarios (e.g., under a high A1FI emissions pathway), just the strong intensity control road salt reduction was sufficient to meet impairment reduction goals regardless of the extensity control of buildout paradigm. Reducing the input of a simple contaminant reduced its adverse impact on aquatic habitat, leading towards a simple evaluation. The analysis does not provide embedded tradeoffs of such an intervention. The potential for reduced travel associated with decreasing deicer application could manifest as a greater rate or severity of vehicular crashes (Fu and Usman, 2014). Furthermore, the assumption that deicer application rates are uniform is not supported even by recommendations (Environment Canada, 2004; Salt Institute, 2007), which could decrease the uniform extensity of the intervention over different roadway and surface types.

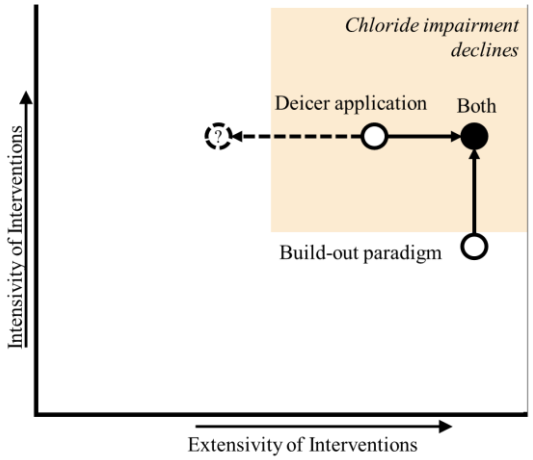


Figure 27: Intensity and extensivity of road salt and build-out strategy to maintain or decrease chloride impairment in future climate and land-cover change contexts. The yellow square in the top right corner reflects a region where the interventions adopted could be successful. Deicer application was assumed to be applied uniformly across all impervious surfaces in the basin in Chapter I, therefore, the reduction of deicer exhibits as much extensivity as it does intensity and plots along a line of equality between the two axes. Because reducing road salt application rates may not be equally applicable over all surfaces, the extensivity of this intervention may be lower than represented in the analysis (indicated by dashed line and oval) and may require additional interventions such as build-out controls to achieve watershed management goals.

The nearly successful implementation of conjunctive management in the Upper Snake River Basin (Chapter II) paired enhanced aquifer recharge with an extreme adoption of modern irrigation technology. Distinguishing the factors that define these interventions as intensive or extensive is less clear; managed aquifer recharge is limited geographically, but technological modernization occurred across all croplands. Practical considerations would; however, affect the intensity and extensity of interventions beyond what was characterized in the study (Figure 28). For instance, managed aquifer recharge is considered for discrete recharge basins at specific locations throughout the watershed, which would increase groundwater availability locally (IWRB, 2009) over smaller areas than represented by the lumped aquifer parameterization of Chapter 2. In addition, the intensity of enhanced aquifer recharge as an intervention is limited by

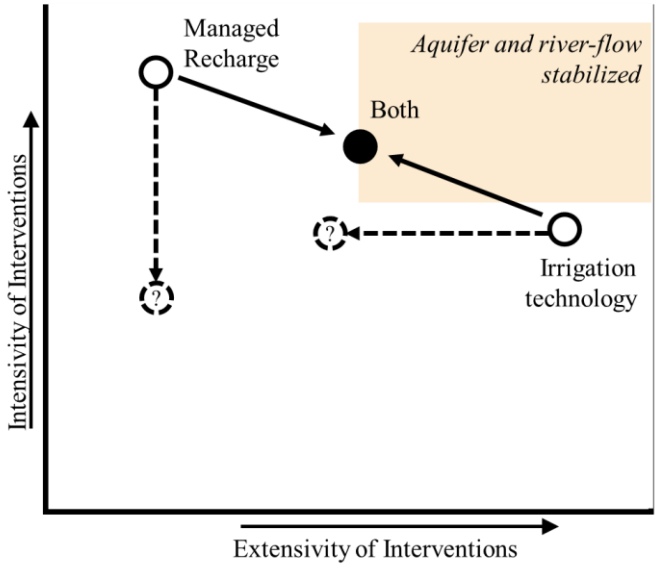


Figure 28: Intensity and extensivity of the irrigation technology modernization and managed aquifer recharge interventions to achieve stabilization of aquifer and streamflow. Modernization of irrigation technology is applied extensively and uniformly across all crop areas; however, as indicated by the dashed line and oval, is likely only applicable to a subset of existing crops. Therefore, the intervention would likely be applied much less extensively than assumed in the study. Similarly, managed aquifer recharge was intensively applied over a geographically important area; however, is unlikely to be achievable at rates that appear necessary for achieving aquifer stability as indicated by the dashed line and oval.

infiltration capacity of the recharge basins (Figure 13), which would preclude enhanced aquifer recharge to levels below the rates required at any tested counterfactual of irrigation technology. Finally, the extensity of technology modernization would be strongly impacted by the crop choice within the basin; dominant crops grown in the region such as maize, alfalfa, potato, beets (Leytem et al., 2021) are poorly suited to drip irrigation (Jägermeyr et al., 2015). Therefore, crop selection would need to change to accommodate more modern forms of irrigation, or the extensivity of this intervention would be diminished.

Export of nitrate to the Gulf of Mexico did not meet target reductions through either program of wetland restoration (Chapter 3) because intensity and extensity were not complementarily moderate or strong (Figure 29). Weak extensity of the FWP, and weak to moderate intensity of the WRP resulted in interventions meeting only 14 or 60% of the goal in reduction, respectively. In Chapter 3, the core constraints on implementation through the lens of intensity and extensity were implemented. The main finding from the study was that interventions that increase the extensity of intervention, likely through processing of nitrate already in the river network, will be needed to meet the targets for nitrate export (Mississippi River/Gulf of Mexico Watershed Nutrient Task Force, 2008; Scavia et al., 2017). Denitrification in wetlands, although only a moderately intensive intervention (30 to 50% nitrate is typically removed), remains one of the most intensive interventions available for nitrate reduction (Hansen et al., 2021). Even fallowing of croplands entirely would be less effective as legacy nitrogen stored in soil leaches to streams (Van Meter et al., 2018). A suite of interventions will be necessary to reduce nitrate export to targets (Marshall, 2018), each intervention spanning unique combinations of intensive and extensive control.

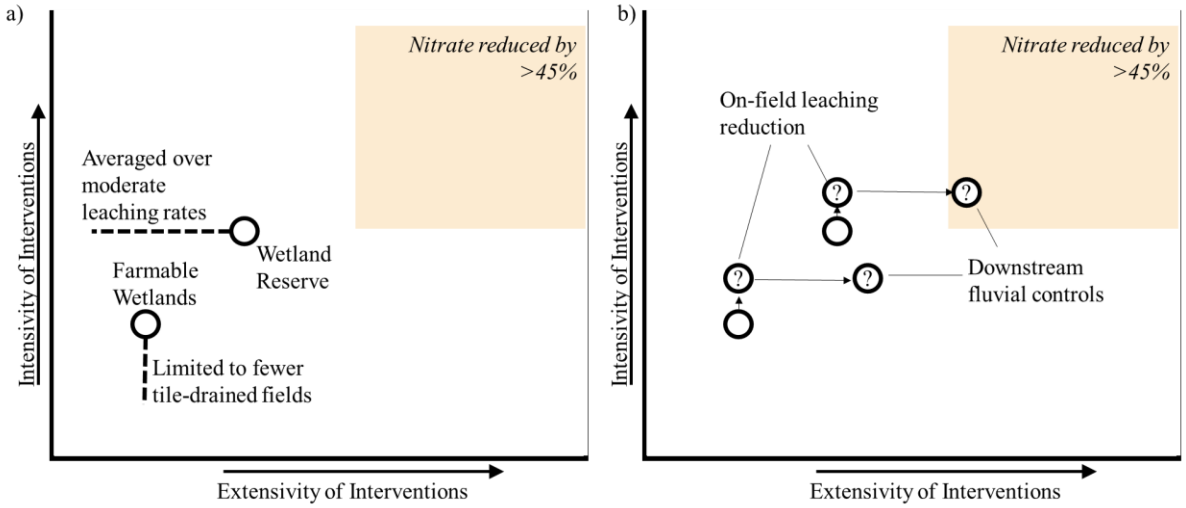


Figure 29: Intensity and extensivity of two programs of wetland restoration to reduce nitrate export to the marine ecosystem. Neither program provided intensive enough nitrate removal over a sufficient area to achieve the watershed management goal (a). Neither program intercepted sufficient nitrate to achieve the goals, meaning the intervention was not applied extensively enough, and only the Wetland Reserve Program exhibited sufficiently high denitrification rates to achieve the target reduction, albeit locally. To achieve watershed management goals (b), pairing the two programs with on-field interventions that reduce runoff, and downstream interventions that treat more runoff may be effective due to the additive efficiencies in denitrification biogeochemistry (Christianson et al., 2018).

The concepts of *intensive* and *extensive* controls on interventions in watershed management seem obvious and even self-evident; however, there is little explicit discussion of these controls in watershed management literature. An important factor in watershed management is that hydrologic systems provide a natural way by which the extensivity of an intervention can be increased. Fluvial or downstream interventions provide a natural mechanism to intervene against inputs or contaminants discharge far upstream. In the context of nutrient reduction, the greater efficacy of fluvial wetland restoration than field-margin wetland restoration is a prime example (Hansen et al., 2021; Czuba et al., 2018). In such a context, there are conceptual similarities with the iSIDES model for understanding the impact of an environmentally degrading practice that generates some pollutant (Hale et al., 2014). The framework identifies similar ideas of pollution intensity (in contrast to the intervention intensity discussed here), and ameliorative effects of

dilution and ecosystem services that accumulate a greater influence at scale. The concept of extensivity presented here acknowledges the spatial heterogeneity of actions beyond those provided by restorative ecosystem services, and utilizes the scale accessed by an intervention on the river network to increase the extensivity of a management practice.

Related ideas of intensity and extensivity are discussed in environmental economics, particularly agricultural economics. For instance, Soule et al. (2001) review agricultural economics literature to assess whether government sponsored crop insurance has reduced the input of agricultural contaminants such as fertilizers and pesticides. Their comprehensive literature review suggests that fertilization rates likely declined on aggregate when farmers enrolled in insurance programs. However, farmers enrolled in the crop insurance programs were also more likely to bring marginal lands into production. In this case, it appears that crop insurance may have increased the intensity control while simultaneously causing a negative feedback on any potential extensity control, leading to an inconclusive assessment of the potential environmental impact of crop insurance programs (Soule et al., 2001).

Evaluating contemporary work in socio-hydrology through the lens of intensity and extensity controls on management practices illustrates the potential value in formalizing the concept. Sociohydrology aims to integrate human decision-making as an integral part of hydrologic analysis (Sivapalan et al., 2012). Socio-hydrologic modeling studies share common themes where policy and practice evaluations are assessed to inform resilient watershed management (Di Baldassarre et al., 2015; Giuliani et al., 2016; van Emmerik et al., 2014). As the field of socio-hydrology matures and analyses move beyond simple and coarse-scale analyses, the concept of intensity and extensity of interventions will become invaluable assets in evaluating the efficacy of proposed practices.

The year 2022 marks the end of hydrologic research decade called Panta Rhei (*everything flows*) by the International Association of Hydrological Sciences, which focused attention on the co-evolution of humans and watersheds (Montanari et al., 2013). New research directions will incorporate the explicit representation of humans on the landscape as agents of hydrologic processes, and distributed modeling of sociohydrological systems will be a critical method in developing new insight (Troy et al., 2015). We have seen the beginning of such spatially explicit representations of sociohydrologic modeling with studies that apply hydro-economic models, and natural experiments that represent economic and policy constraints on human-decision making within the context of water availability and water resource management (Tran et al., 2019, 2020; Penny et al., 2020; Hansen et al., 2021).

APPENDIX I

SUPPLEMENTAL MATERIAL FOR CONTROLS OF CHLORIDE LOADING AND IMPAIRMENT AT THE RIVER NETWORK SCALE IN NEW ENGLAND

Supporting information includes additional background on the observation data used, on the FrAMES-NACL model, presents methodology of the Markov-Chain Monte Carlo experiment, provides details of the calculation of the informal likelihood function, and presents additional validation data.

Sensor data used in study

Raw time-series data and more information about the Lotic Volunteers Temperature, Electrical Conductance, and Stage (LoVoTECS) data used in the analysis are available at lovotecs.sr.unh.edu. Chloride concentration data for synoptic sampling events in 2013 analyzed by the University of New Hampshire Water Quality Analysis Laboratory according to USEPA method 300.1 are related to hourly average station data (Figure AI.1). Ordinary least square regression of the specific conductance to measured chloride concentration yields Equation 1 of the Chapter I and is shown on Figure AI.1. We calculate similar relationships as shown on Figure AI.1 if stations were binned based on river order or impervious cover (with river order 5 and impervious cover of 25% demarcating bins) or if all specific conductance and chloride data pairs were averaged by station.

Mass balance calculations in FrAMES-NACL

Within FrAMES, the Water Balance Model (WBM) controls vertical water transfer and terrestrial runoff generation (Figure AI.2a) which is routed through a 1-D simulated topological network (STN) river system using the Water Transport Model (WTM) (Vörösmarty et al., 1998; Wisser et al., 2010). Precipitation and snowmelt are transferred either to the root zone or directly to the stream network from hydrologically connected impervious (hci) areas. Snowpack accumulation and melting follow a simple temperature index method (Willmott et al., 1985). Evapotranspiration is drawn from the root-zone (Hamon, 1963). Saturation excess, generated

when rooting zone soil water exceeds field capacity, is directed to two exponential runoff generating pools representing surface (quickflow) and shallow groundwater (baseflow) flow-paths.

Two modifications were made to FrAMES model structure to accommodate NACL functionality. First, the hci parameter, which controls impervious connectivity to stream networks, was decreased during winter precipitation compared to liquid rain events to account for greater routing of road salt to the subsurface during winter storms (e.g. snow plowed from roads is piled up on sides which infiltrate soils). Second, we added a groundwater storage pool to represent long-term storage in the subsurface. Previous versions of FrAMES conceptualized a shallow groundwater compartment with an exponential transit time distribution (TTD) to generate the hydrodynamic response of catchments. We represent longer time-scales of transport through the subsurface as mass exchange between the shallow groundwater compartment and an immobile zone (discussed below).

Chloride concentrations are calculated within each compartment of NACL and based on incoming water and their associated concentrations. Fluxes between compartments are defined by flow between compartments assuming instantaneous mixing throughout the compartment:

$$\frac{dM_\chi}{dt} = \sum_{j=0}^{m} \dot{m}_\chi^j - \sum_{k}^{n} Q_\chi^k \, C_\chi \qquad \text{Eq. AI.1}$$

where M_χ [kg] is chloride mass within compartment χ (e.g. snowpack, rooting zone, surface runoff pool, or shallow groundwater (baseflow) runoff pool), $\dot{m}_\chi^j$ [kg Cl d^{-1}] is the j^{th} mass inflow to compartment χ, Q_χ^k [$\mathrm{m}^3\mathrm{d}^{-1}$] is the k^{th} water outflow from compartment χ, and C_χ [kg m^{-3}] is the chloride concentration in compartment χ during the timestep given by:

$$C_\chi = \frac{M_\chi + \sum_{j=0}^{m} \dot{m}_\chi^j}{V_\chi + \sum_{k=0}^{n} Q_\chi^k} \quad \text{Eq. AI.2}$$

where V_χ [m^3] is the water volume in compartment χ. Water and chloride compartments and transfers are illustrated graphically in Figure AI.2.

FrAMES-NACL calculates a chloride mass exchange flux between the mobile (shallow groundwater) and immobile zone ($\dot{m}_{imm}$ [kg Cl d^{-1}]) following the time-step linearization technique of Silva et al. (2009). Exchange is defined by concentrations within the mobile (C_m) and immobile zones (C_{imm}), exchange coefficient (α_{imm} [d^{-1}]), and immobile zone volume (β_{imm} [L^3]), using concentrations in each zone from the previous timestep ($t-1$):

$$\dot{m}_{imm} = \beta_{imm}\alpha_{imm}C_{imm}^{t-1}\, exp(-\alpha_{imm}dt) - \beta_{imm}\alpha_{imm}C_m^{t-1}\, exp(-\alpha_{imm}dt) \quad \text{Eq. AI.3}$$

The concentration within the mobile zone (C_m) is updated at time t using a modified mobile zone storage term (β^*) and outflow from the mobile zone through the baseflow flux (Q_{BF}).

$$C_m^t = \frac{C_m^{t-1} + \frac{\dot{m}_{imm} + \sum_{j=1}^{m} \dot{m}_j}{\beta^*}}{1 + \frac{Q_{BF}}{\beta^*}} \quad \text{Eq. AI.4}$$

where $\dot{m}_j$ represents the m individual chloride fluxes to the mobile groundwater zone. The modified mobile zone storage term (β^*) is the sum of the mobile zone volume (β_m) and immobile storage accessible over the duration of the time-step:

$$\beta^* = \beta_m + \beta_{imm}[1 - exp(-\alpha_{imm}dt)] \quad \text{Eq. AI.5}$$

Finally, NACL updates the immobile zone concentration at time t (C_{imm}):

$$C_{imm}^{t} = C_{imm}^{t-1}\, exp(-\alpha_{imm} dt) - C_{m}^{t-1}[exp(-\alpha_{imm} dt)] + \frac{C_{m}^{t} - C_{m}^{t-1}}{dt}\left[1 - \frac{1}{\alpha_{imm}}(1 - exp\,(-\alpha_{imm} dt))\right] \quad \text{Eq. AI.6}$$

We use a single immobile zone of fixed size (β_{imm}) with a single mass exchange coefficient (α_{imm}) at each gridcell; however, the method could be extended to multiple immobile zones to generate longer-tail transit time distributions. The immobile domain is fixed in size, which assumes deeper groundwater volumes remain constant. As the groundwater pool drains during summer low-flows, mass exchange from the immobile zone dominates salt concentration in groundwater resulting in maintenance of high seasonal salt concentrations, even in response to short term storm events.

Parameterizing FrAMES-NACL

The simulated topological network (STN) representing the river network of the study domain used the HydroSHEDS digital elevation model (Lehner et al., 2008) aggregated to a geographic resolution of 45 arc seconds (approximately 1.5 km resolution) through network scaling (Fekete et al., 2001). We used daily gridded air temperature from the Modern-Era Retrospective Analysis for Research and Applications (Rienecker et al., 2011) and downscaled temperature to match STN resolution, using an elevation lapse of -6.4°C/km (NOAA et al., 1976). For precipitation forcing data, we re-interpolated (geographic inverse distance weighting with $p = 2$) MERRA surface precipitation and NOAA Global Historical Climate Network ground station data with equal weighting. We used land-cover data derived from NOAA's Coastal Change Analysis Program (Vogelmann et al., 1998), imperviousness from national land cover data (Xian et al., 2011), and population density data derived from U.S. census data (Thorn et al., 2017)

NACL represents four sources of chloride (Figure AI.2b) with specific parameters defining loading provided on Table AI.1. Chloride input via precipitation ($\dot{m}_{ppt}$ [kg Cl d^{-1}]) is calculated from daily precipitation [mm d^{-1}] and weekly chloride concentration in wet deposition (C_{ATM} [kg Cl m^{-3}]). We interpolated (geographic inverse distance weighting with $p = 2$) weekly mean concentrations of chloride in precipitation (C_{ATM} [kg Cl m^{-3}]) measured through the National Atmospheric Deposition National Trends Network (http://nadp.sws.uiuc.edu/NTN/) at stations within (NH02 at Hubbard Brook Experimental Forest), and surrounding our study domain (MA08, MA13, ME02, ME08, ME96, and VT01). We associated all wet precipitation with C_{ATM} throughout the study domain and assume no loading from dry deposition.

Chloride input via road salt application is calculated from per area application rate per unit snowfall and the fraction of impervious surfaces that are treated with deicer (f_{DEI}). Snowfall rate in snow water equivalents (P_W [mm d^{-1}]), defined by the precipitation rate on days with daily mean air temperature less than $-0.29°C$ (Stewart et al., 2013) controls road salt applied. Treated impervious areas receive chloride at a rate specified by the deicer loading parameter (C_{DEI} [kg Cl $mm^{-1}m^{-2}$]), which assumes chloride salt deicer application rate is constant per depth of frozen precipitation. The mass flux of deicer ($\dot{m}_{dei}$) applied to a grid-cell is then given by

$$\dot{m}_{DEI} = f_{DEI} \times A_I \times C_{DEI} \times P_w \qquad \text{Eq. AI.7}$$

where A_I [m^2] is the total impervious area of the grid cell.

The mass flux from domestic salt loading ($\dot{m}_{POP}$ [kg Cl d^{-1}]), including septic and sewer waste and water softeners leachate, is calculated from a per capita loading rate (L_{POP} [kg Cl $P^{-1}d^{-1}$]) and population density (ρ_{POP} [P km^{-2}]):

$$\dot{m}_{POP} = L_{POP} \times \rho_{POP} \times A \qquad \text{Eq. AI.8}$$

where A [km^2] is the grid-cell area. We use estimates of domestic chloride loading (L_{POP}) derived from monitored sewage effluent (Struzeski, 1971; Novotny et al., 2009), but account for chloride introduced by water softeners (Godwin et al., 2003; Kelly et al., 2008; Trowbridge et al., 2010). We do not distinguish direct transport of chloride to river systems through sewered infrastructure, and we consider chloride added during waste treatment to be negligible (Kelly et al., 2008). By neglecting sewer flow, domestic chloride travels longer through municipal centers than is realistic; however, because domestic loading is considered constant, overestimating transport time doesn't affect our simulations.

Agricultural salt loading is represented as a constant areal loading rate (L_{AG} [kg Cl $m^{-2}d^{-1}$]) for fractions of grid cells with agricultural land cover (f_{AG}):

$$\dot{m}_{AG} = L_{AG} \times f_{AG} \times A \qquad \text{Eq. AI.9}$$

Agricultural loading (L_{AG}) estimates were taken from two Midwestern US studies and likely overestimate loading from the pastures that constitute most croplands in the study region (Novotny et al., 2009; Stites and Kraft, 2001).

Markov-chain Monte Carlo calibration and experiment

To address our first hypothesis, we use FrAMES-NACL to calibrate deicer application rate (C_{DEI}), and compare to empirical estimates of deicer application rates obtained from three studies (Godwin et al., 2003; Sander et al., 2007; Trowbridge et al., 2010) that inventoried road salt

usage. We estimated deicer application rates [$g\ \mathrm{Cl}\ mm^{-1}m^{-2}$] from inventories by dividing total salt usage reported by the inventory studies by total received snow (typically annual),scaled to an estimate of treated impervious area. If total frozen precipitation was not reported in a study, we used data from nearby NOAA Global Historic Climate Network stations. Where application rates were reported in lane length, we calculated lane area assuming a lane width of 3.66m. In parameterizing road salt application in FrAMES-NACL, we assume 60% of impervious areas are used for pedestrian and vehicle travel (Cappiella and Brown, 2001; Southworth, 2003; Shuster et al., 2005) and receive deicer application (e.g. $f_{DEI} = 0.6$) (Figure AI.3), and that the remainder of impervious area (e.g. roofs) receives no deicer.

We calibrated road salt chloride loading rates to be consistent with the distributions of chloride observed in streams using FrAMES-NACL. A Markov-Chain Monte Carlo (MCMC) analysis [Goodman and Weare 2010, Foreman-Mackey et al. 2013] informed C_{DEI} estimates in the context of uncertain hydrologic storage and flow. The MCMC analysis characterized posterior parameter distributions for 6 free hydrology and chloride-related parameters summarized on Table AI.1. We used the acceptance ratio (R_A) (ρ in Sadegh and Vrugt (2013 eq. 10)) on probabilities of non-exceedance (PONE) of target variables (R_A) as an informal Bayesian likelihood to test model behavior. Observations were defined as PONE from daily data for both discharge (Q) and specific conductance (k_0) at 9 discrete probabilities (0.1, 0.2, … 0.9) similar to methods used by Yu and Yang (2000), Westerberg et al. (2011, 2014), and Vrugt and Sadegh (2013).

The MCMC was started with normally distributed ensembles about each parameter's initial estimate assuming most parameters initial variability was 10% of the estimate. For road salt loading, the initial value and variability of C_{DEI} came from the average of the inventory

estimates. The prior distribution for C_{DEI} was log-uniform on bounds that encompassed 1.5 times the range in inventory estimates.

The affine-invariant sampler (Goodman and Weare, 2010; Foreman-Mackey et al., 2013) explored the posterior distribution for a total of 111 iterations equating to 13440 model runs for the MCMC. The integrated autocorrelation time (τ_E) of ensemble means (Goodman and Weare, 2010) for various metrics of the convergence process including parameter values and all fit summary statistics, including R_A, suggest that initial parameter estimates and variance were close to the posterior stationary estimates. 34 iterations (4080 model evaluations) were discarded as burn-in following the suggestion of Sokal (1997, p.8) to discard $20\tau_{\text{exp}}$ iterations, and assuming that the ensemble mean autocorrelation time τ_E of Goodman and Weare (2010) is a reasonable approximation of the exponential autocorrelation time τ_{exp} from Sokal (1997). The final 77 iterations (9240 model evaluations) sampled the stationary posterior distribution for the MCMC experiment.

Calculation of informal likelihood

We use the acceptance ratio (R_A), the fraction of model predictions that fall within a band of acceptability defined by observation data (Q or k_0) (Vrugt and Sadegh, 2013). The acceptance ratio $[0,1]$ represents the proportion of simulated values within error bounds (based on observational and estimated model structural error) of observational data. We selected stations with at least 170 days of data throughout the entire year from catchments of order 5 or less, retaining higher order stations for validation (k_0:n=5, Q:n=9). The larger pool of stations measuring specific conductivity weights k_0 almost twice as heavily as discharge. *A priori* heteroscedastic estimates of observational error (σ_Q or σ_{k_0}) were 10% of observation value for

discharge (Q) and 6% of observation value for specific conductance (k_0), the latter based on manufacturer reported instrument error of instantaneous measurements at 95% confidence:

$$\sigma_Q = 0.1Q \quad \text{Eq. AI.10}$$

$$\sigma_{k_0} = 0.06k_0 \quad \text{Eq. AI.11}$$

A multiplicative factor κ equal to 4 accounted for additional model structural error. We double the value of κ from analysis of Vrugt and Sadegh (2013) because we are incorporating information from numerous subcatchments of the Merrimack and Piscataqua Rivers; however, parameterizations are not spatially explicit and apply to both watersheds as spatially homogenous expected values.

Station data from lower (1-5) orders defined model performance whereas we reserved higher order station data for validation. For each j station, time-series of the lower and upper error bounds are:

$${}^{-}Q_j = Q_j - \kappa\sigma_{Q,j} \text{ and } {}^{-}k_{0_j} = k_{0_j} - \kappa\sigma_{k_0,j} \quad \text{Eq. AI.12}$$

$${}^{+}Q_j = Q_j + \kappa\sigma_{Q,j} \text{ and } {}^{+}k_{0_j} = k_{0_j} + \alpha\sigma_{k_0,j} \quad \text{Eq. AI.13}$$

The model's ability to recreate non-exceedance probability distributions of the two variables of interest (discharge and specific conductance) is the primary metric for quantifying performance. Here, for each j station with a minimum of 200 daily observations of a test variable, we rank lower (e.g. ${}^{-}Q_j$) and upper (${}^{+}Q_j$) bounds on observations using Cunane plotting positions (Stedinger et al., 1992). Runoff and specific conductance are interpolated to ($M = 9$) specific non-exceedance probabilities (p) ($p = [0.1,0.2, \dots ,0.9]$) and for each station define ${}^{-}Q_j^p$, ${}^{+}Q_j^p$, ${}^{-}k_{0_j}^p$, and ${}^{+}k_{0_j}^p$ used for comparison with model generated equivalents ($\tilde{Q}_j^p$ and $\tilde{k}_{0_j}^p$). $\tilde{Q}_j^p$ and

$\tilde{k}_{0_j}^{\ p}$ represent model generated equivalents of observed variables y_j and k_{0_j} ranked and interpolated to p non-exceedance probabilities for the subset of days with observational data for each station j. An acceptance ratio R_A, which is a system level model assessement equivalent to the distance measure used in Sadegh and Vrugt (2013), summarizes model behavior for a particular scenario:

$$R_A = \frac{1}{(N_Q + N_{k_0})M}\left(\sum_{j=1}^{N_Q}\sum_{p=1}^{M} A\left(\tilde{Q}_j^p,\ {}^{-}Q_j^p,\ {}^{+}Q_j^p\right) + \sum_{j=1}^{N_{\tilde{k}_0}}\sum_{p=1}^{M} A\left(\tilde{k}_{0_j}^{\ p},\ {}^{-}k_{0_j}^{\ p},\ {}^{+}k_{0_j}^{\ p}\right)\right) \quad \text{Eq. AI.14}$$

where N_Q and N_{k_0} are the number of sites with sufficient data to test runoff and specific conductance, respectively, and the acceptance function A counts the number of instances where a model generated variable at a given non-exceedance probability is bounded by the estimates of that value from observations:

$$A = \begin{cases} {}^{-}\theta_j^p \le \tilde{\theta}_j^p \le {}^{+}\theta_j^p, & 1 \\ \tilde{\theta}_j^p < {}^{-}\theta_j^p, & 0 \\ \tilde{\theta}_j^p > {}^{+}\theta_j^p, & 0 \end{cases} \quad \text{Eq. AI.15}$$

where θ represents either Q or k_0. Stated simply, R_A is the number of sites and non-exceedance probabilities where model generated predictions for two variables are consistent with observations, expressed as a proportion total sites, probability bins and variables. R_A is defined as a range between 0 and 1, with 1 defining a perfect simulation.

We calculate the informal likelihood using the PONE of target variables in favor of time-series because our metric of chloride impairment is closely related to the PONE of specific

conductivity and the input data has irreducible error that contaminates parameter identification using simulated time-series. We compare concentrations to a fixed threshold (§3.3) throughout the productive season, and not at points in any species' life cycle, obfuscating a requirement for predictions at specific times. Error in individual storm magnitudes represented by input data contributes significantly to total model error in hydrologic and transport simulations (Kavetski et al., 2006a, b; Vrugt et al., 2008; Renard et al., 2010, 2011; McMillan et al., 2012; Andino et al., 2016). MERRA precipitation intensities exhibit considerable error at shorter than monthly frequencies (Rienecker et al., 2011), and poor resolution of storms in small observational catchments follows from MERRA's coarse (0.5°x0.67°) resolution. Utilizing exceedance probabilities for our objectives eliminates the latent parameters typically employed when representing input data errors explicitly (Kavetski et al., 2006a, b; Vrugt et al., 2008; Renard et al., 2010).

Definition of Impairment Threshold

Prolonged exposure of freshwater with salinities or chloride concentrations less than 400 [$\mu S\ cm^{-1}$] reduces richness of community structure in fish (Morgan et al., 2012), fen gramminoids (Richburg et al., 2001), and stream macroinvertebrates (Blasius and Merritt, 2002). At prolonged exposure at higher salinities (e.g. > 800 [$\mu S\ cm^{-1}$]), zooplankton richness and reproduction is reduced, and invertebrate mortality increases (Findlay and Kelly, 2011b; Cañedo-Argüelles et al., 2013). We selected a threshold that bisects these two levels 600 [$\mu S\ cm^{-1}$].

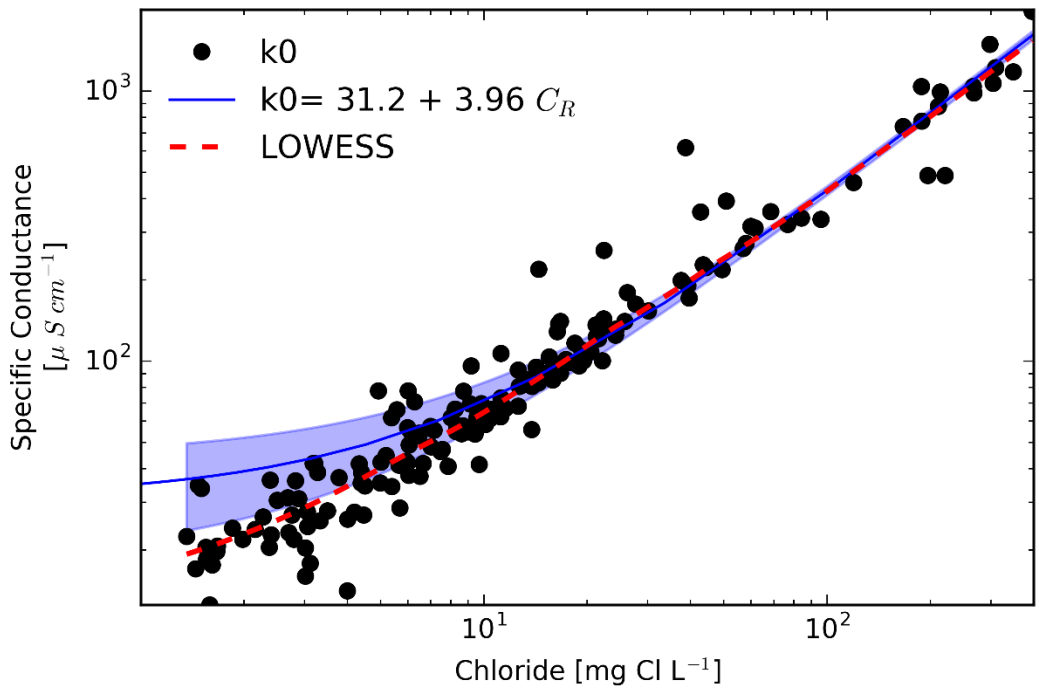


Figure AI.1: Scatter-plot of hourly station specific conductance versus chloride concentration sampled across the LoVoTECS network on three dates in 2013. Line of best fit and 95% confidence interval about the mean prediction are shown in blue. Locally weighted scatter-plot smoothing (LOWESS) shows OLS exceeds specific conductance estimates from chloride at the lowest concentrations of chloride.

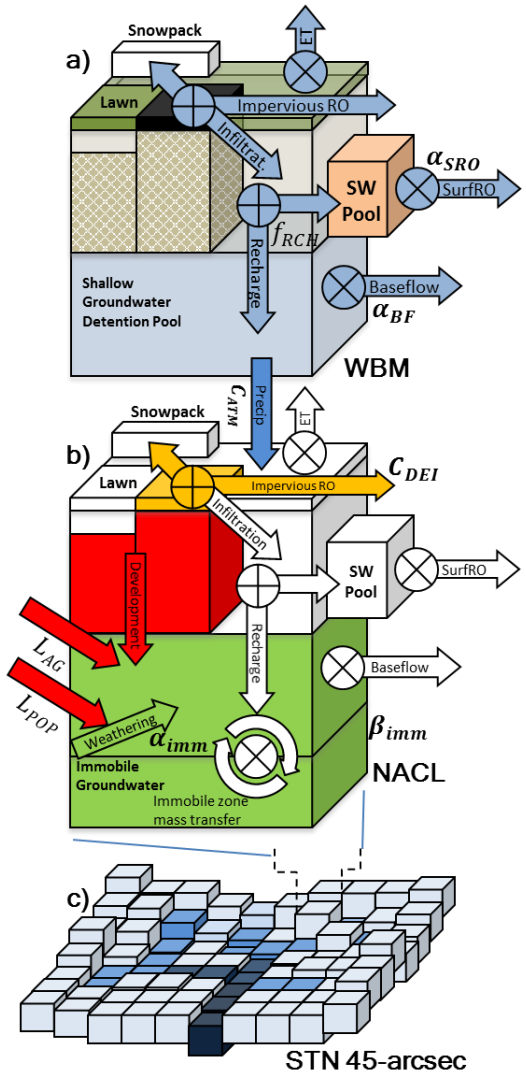
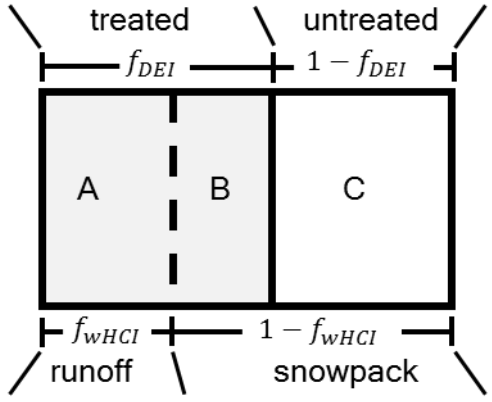


Figure AI.2: Conceptual diagrams of a) Water Balance Model (WBM), b) the Non-point Anthropogenic Chloride Loading (NACL) model, and c) simulated topological network (STN) defining 1-D drainage network. Major hydrologic fluxes and stocks are labeled in a; RO – runoff. The 6 parameters controlling runoff generation and chloride explored via MCMC are in bold italics. Parameters with diagonal crosses are time-scale parameters with units including t^(-1), and parameters with vertical/horizontal crosses represent divergence parameters separating flow paths.

Table AI.1: Initial (x_0) and posterior (x_n) credible ranges (CR) for FrAMES-NACL parameter values. Parameters with no reported posterior credible ranges were not included in the Monte Carlo analysis.

Variable	Definition	x_0 (base)	Units	Source	x_n (95% CR)
C_{DEI}	Chloride application from deicer per depth of frozen precipitation per area	8.0 (0.2,33)	g Cl $mm^{-1}m^{-2}$	Opt	6.5 (5.3,11.)
α_{SRO}	Release rate from surface pool (quickflow)	0.25 (0.0003,0.5)	d^{-1}	Pre[4]	0.31 (0.17,0.46)
α_{BF}	Release rate from groundwater pool (baseflow)	2.5 (0.03,14)	$\times 10^{-2} d^{-1}$	Pre[4]	2.4 (1.7,3.1)
α_{imm}	Exchange rate with deep groundwater pool	2.5 (0.0008,18)	$\times 10^{-3}$ d^{-1}	Opt	2.4 (0.5,11)
β_{imm}	Size of deep groundwater pool in equivalent water stored	2.4 (0.05,7.4)	$\times 10^{3}$ mm	Opt	2.3 (1.0,5.2)
a_w	Fraction of hydrologically connected imperviousness during winter	0.15 (0.01,0.99)	–	Opt	0.13 (0.03,0.25)
L_{POP}	Domestic per capita chloride loading	0.05 (–)	kg Cl $P^{-1}d^{-1}$	Emp[5,7,8]	–
L_{AG}	Per area agricultural chloride loading	1.8 (–)	$\times 10^{-5}$ g Cl m^{-2}	Emp[5,6]	–
f_{RCH}	Fraction of soil saturation entering baseflow pool (compliment to quickflow)	0.5 (–)	–	Pre[4]	–
f_{DEI}	Fraction of impervious areas treated with chloride deicer	0.6 (–)	–	Emp[9]	–

Empirical (Emp), Previous (Pre), or trial Optimized (Opt) indicate source of parameter values at x_0. Empirical and Previous sources include: 1 - (Godwin et al., 2003); 2 - (Sander et al., 2007); 3 - (Trowbridge et al., 2010); 4 - [*Stewart et al.*, 2011]; 5 - (Struzeski, 1971); 6 - (Stites & Kraft, 2001); 7 - (Novotny et al., 2009); 8 - (Kelly et al., 2008); 9 - (Shuster et al., 2005).



A. $C_{DEI} \times f_{wHCI} \times A_I \times P_w$
B. $C_{DEI} \times \max(f_{DEI} - f_{wHCI}, 0) \times A_I \times P_w$
C. $C_{ATM} \times \min(1 - f_{DEI}, 1 - f_{wHCI}) \times A_I \times P_w$

Figure AI.3. Distribution of chloride flux loading and fates from impervious areas ($\boldsymbol{A_I}$). Frozen precipitation ($\boldsymbol{P_w}$) falling on road salt treated impervious areas defined by $\boldsymbol{f_{DEI}}$ (shaded) either melt directly to stream runoff carrying roadsalt at concentration of $\boldsymbol{C_{DEI}}$ (A), or accumulate in snowpack with chloride concentration of $\boldsymbol{C_{DEI}}$ (B). Frozen precipitation falling on untreated impervious surface accumulates in snowpack with chloride concentration $\boldsymbol{C_{ATM}}$ (C). The condition $\boldsymbol{f_{wHCI} > f_{DEI}}$ is not encountered in this study.

Table AI.2. Summary of fit metrics for headwater (river orders 1-5) and mainstem (river orders 6-8) stations for non-exceedance probabilities of specific conductance ($\boldsymbol{K_0^p}$) and discharge ($\boldsymbol{Q^p}$). Reported fit metrics include the number of observations (number of stations times 9 probabilities tested), acceptance ratio ($\boldsymbol{R_A}$ $[-]$) used in the MCMC, root mean square error ($\boldsymbol{RMSE}$ $[\boldsymbol{m^3 s^{-1}, \mu S\ cm^{-1}}]$), and the median residual ($\boldsymbol{\hat{r}^{0.5}}$ $[\boldsymbol{m^3 s^{-1}, \mu S\ cm^{-1}}]$). RMSE of model simulations are provided for windows of 1, 5, and 10 days.

			Probability of Non-Exceedance				RMSE averaging		
	Network	n	R_A	NSE	$\hat{r}^{0.5}$	$RMSE$	D	5D	10D
K_0^P	Headwater (calibration)	306	0.53	0.75	9.6	180	320	230	205
	Mainstem (validation)	45	0.38	0.07	17.4	32	74	58	54
Q^P	Headwater (calibration)	171	0.88	0.93	-0.04	0.50	3.3	2.7	2.4
	Mainstem (validation)	81	0.85	0.99	-0.15	7.8	72	54	46

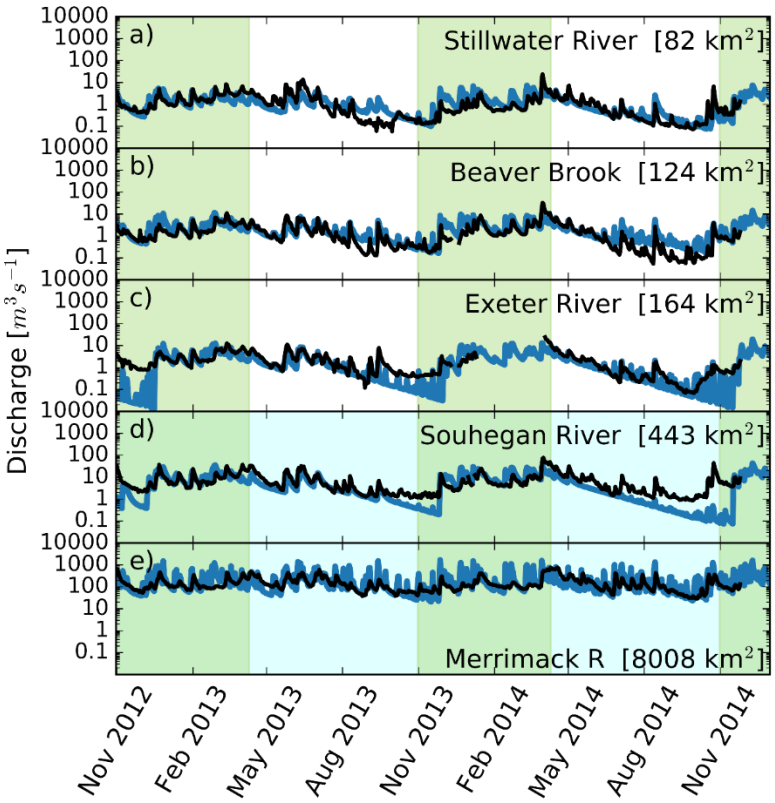


Figure AI.4: Time-series of discharge at five stations comparing FrAMES-NACL (blue) with observational data (black). Values in brackets denote catchment area. Light bands indicate the productive season (mid-April through October).

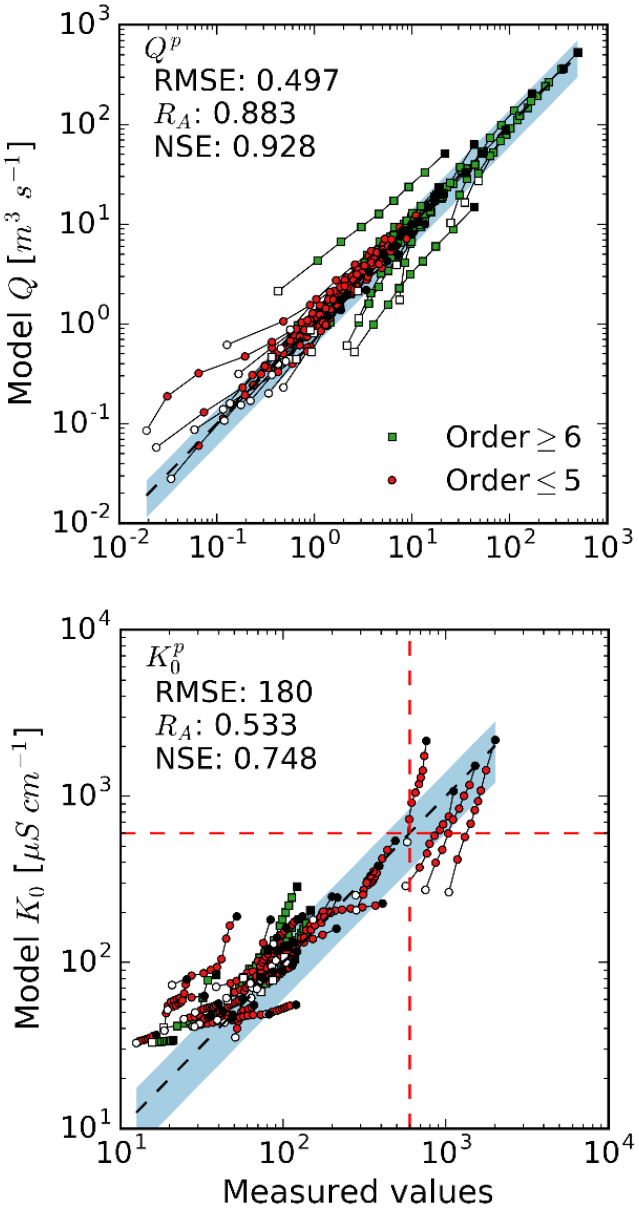


Figure AI.5: Model predicted versus observed discharge ($\boldsymbol{Q}$) and specific conductance ($\boldsymbol{k_0}$) at nine non-exceedance intervals $\boldsymbol{p} = [\mathbf{0.1}, \mathbf{0.2}, \ldots, \mathbf{0.9}]$ for stations with a minimum of 200 days of data. Red circles and green squares represent headwater (order 1-5) and mainstem (order 6-7) stations, respectively. Open (white) symbols indicate the 0.1 exceedance probability value for a station, the closed (black) symbols indicate the 0.9 exceedance probability value. Lines connect symbols for each individual station. Blue bands surrounding dashed 1:1 line represent the limits of acceptability bands. Fit metrics including root mean square error, acceptance ratio, and Nash-Sutcliffe efficiency are calculated using headwater stations.

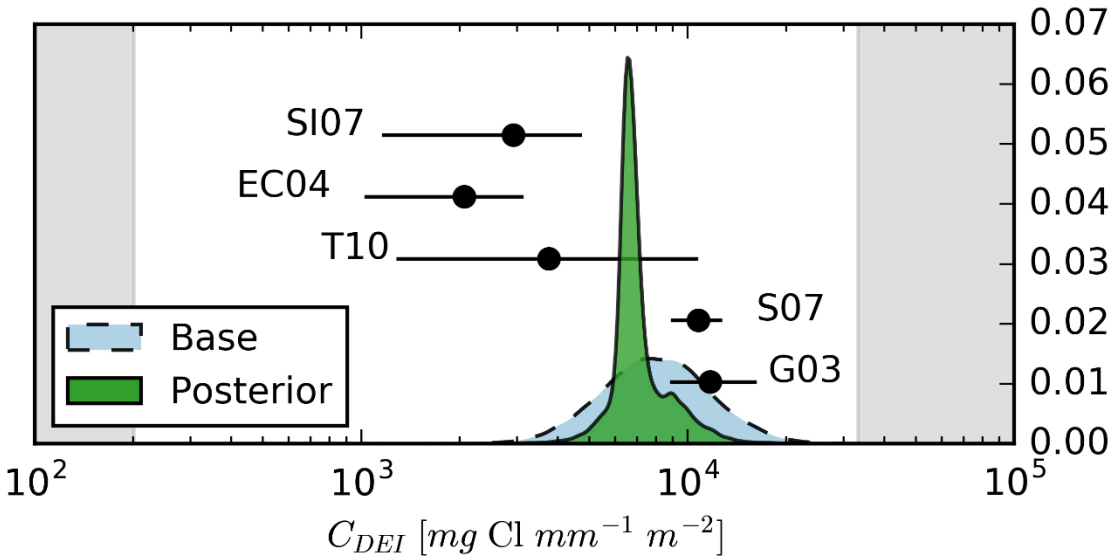


Figure AI.6: Initial parameter value (dashed blue) and posterior (solid filled) kernel densities for $\boldsymbol{C_{DEI}}$ $[\boldsymbol{mg\ \mathrm{Cl}\ L^{-1}}]$. Literature loading best estimates (●) and ranges (—) are overlaid on the kernel densities. Horizontal axes represent the parameter units. Vertical axes represent the relative kernel density of parameter values (integrals = 1) and do not apply to empirical estimates or ranges. Gray regions show areas outside the prior distribution. The sources for empirical loading estimates are: G03 (Godwin et al., 2003) (Mohawk River Valley, NY), EC04 (Environment Canada, 2004), S07 (Sander et al., 2007) (TCMA, MN), SI07 (Salt Institute, 2007), T10 (Trowbridge et al., 2010) (Southeastern NH).

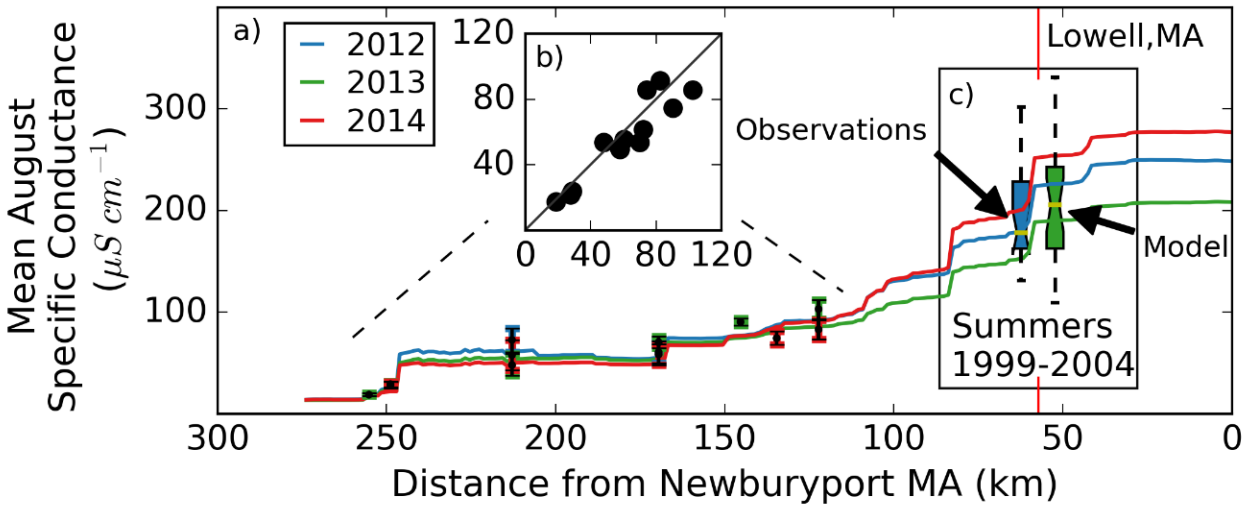


Figure AI.7: Profiles of mean August specific conductance along Pemigewasset-Merrimack River profile compared to observations in the LoVoTECS network. The red vertical line indicates the location of the USGS station at Lowell. Panel b) shows the model versus observed mean August specific conductance from LoVoTECS data along the Pemigewasset-Merrimack profile (shown in Panel a). Panel c) Box-and-whisker plots of discrete grab samples of summer specific conductance (USGS), and modeled specific conductance between 1999 and 2004 at Lowell, MA.

APPENDIX II

SUPPLEMENTAL MATERIAL FOR INTERPLAY OF CHANGING IRRIGATION TECHNOLOGIES AND WATER REUSE: EXAMPLE FROM THE UPPER SNAKE RIVER BASIN, IDAHO, USA

Additional information and data included in this Appendix supplements Chapter II. The spatial distribution of infiltration fraction and dams provisioning water to each region of the Upper Snake River Basin are depicted in Figures AII.1 and AII.2, respectively. Complete descriptions of the representation of irrigation technology and lumped aquifers is provided. Crop input data and methods for calculating spring discharge are detailed. Graphics presenting model validation and behavior are presented, and finally the mass balance of all irrigation returns are illustrated.

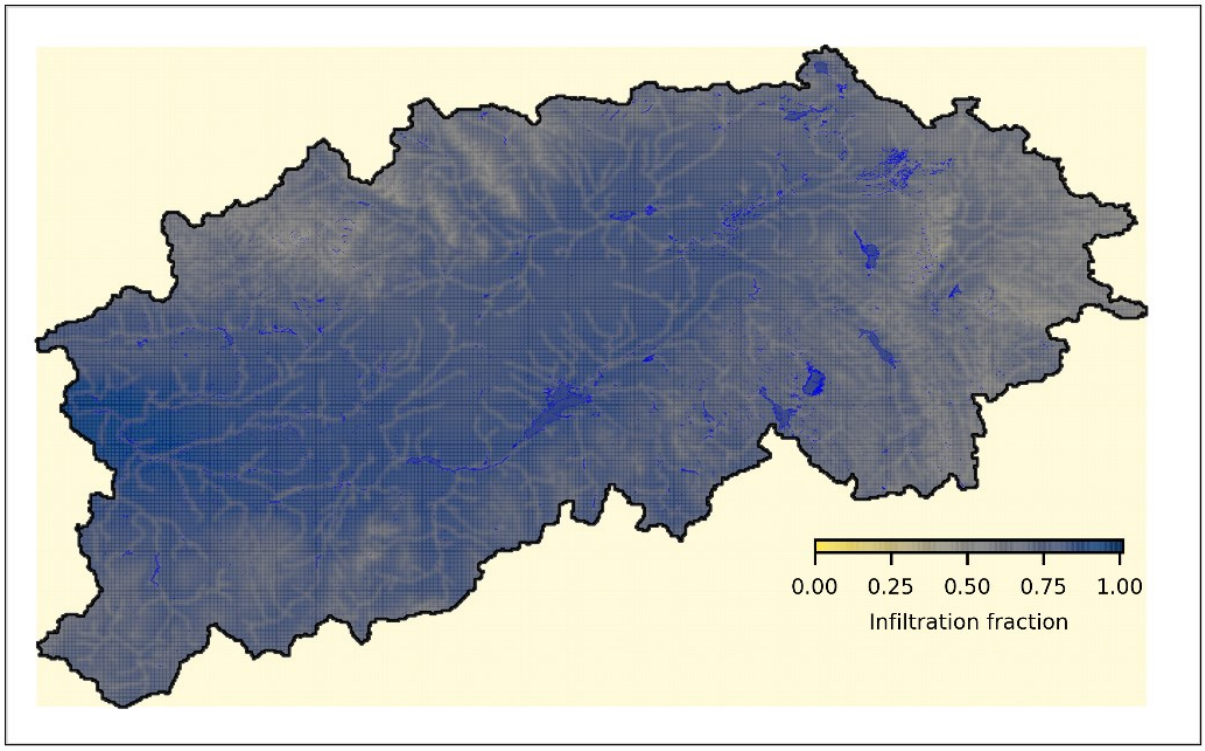


Figure AII.1: Infiltration fraction defining the proportion of saturation excess that infiltrates to shallow and deep groundwater reservoirs; the complement runs off the soil surface. The distribution of infiltration fraction qualitatively exhibits observations that runoff occurs proximal to moderate size rivers (greater than order 4), and tends to decrease with elevation.

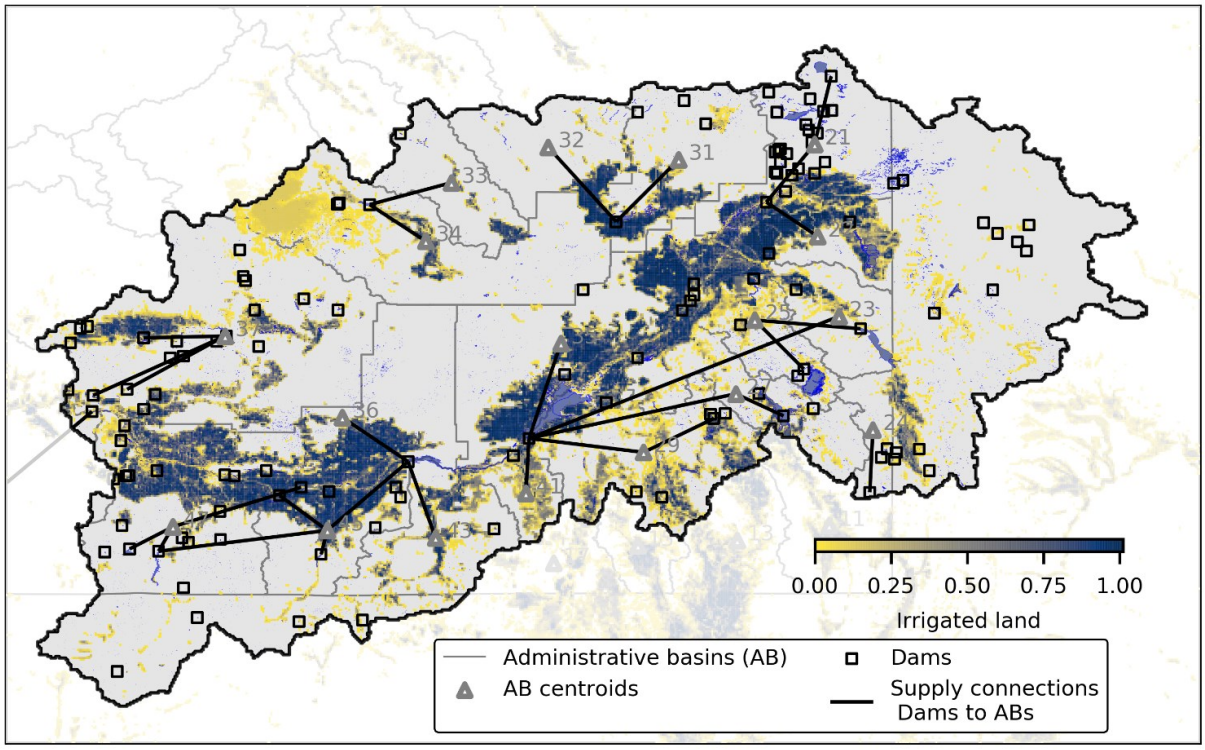


Figure AII.2: Map showing Idaho administrative basins, dam locations, connections between specific dams and administrative basins used for provisioning surface water, and total irrigated lands as a fraction of pixel area. Mapping specific reservoirs to each AB was done manually by tracing canal courses from National Hydrography Data (nhd.usgs.gov) from reservoirs through irrigated areas. In addition, we derived irrigation water for AB 23 (eastern border of Idaho) from American Falls Reservoir because abstractions from this area of the state are taken below Palisades Reservoir, the next downstream reservoir is American Falls.

Irrigation technology and modernization

Irrigation technology was revised in the UNH Water Balance Model (WBM) to a process-based representation as an alternative to the prior conceptual formulation where non-beneficial fates were specified as a fraction of gross irrigation (Grogan et al., 2017; Wisser et al., 2008, 2010). The process-based formulation redistributes inefficient irrigation water via surface runoff flows, groundwater percolation, and evaporation during both delivery and application stages. The system explicitly represented non-consumptive losses using technology specific parameters

applied to proportions of irrigated croplands operating each technology. Losses during delivery were calculated from conveyance surface area (as a fraction of irrigated cropland), daily open water evaporation, and percolation. Conveyance methods included pipes with no evaporation or percolation, and open conveyances such as canals and ditches that percolate at a fraction of local infiltration rates and evaporate from their surfaces. Incidental losses during application follow Jägermeyr et al. (2015) and use the distribution uniformity parameter that described excess water needed to satisfy net irrigation demand based on the type of technology, either drip, sprinkler, or flood. The distribution uniformity parameter was maintained at the values originally estimated for surface/flood, sprinkler, and direct/drip agriculture (Jägermeyr et al., 2015).

The process of calculating non-beneficial use (N) and non-consumptive returns (L) via application of irrigation water is performed throughout the WBM time-step cycle. Following calculation of net crop water demand (I_{net}), additional delivery and application requirements are calculated accounting for technology specific inefficiencies. Then, an initial estimate of delivered water is based on estimated water availability and if available water is determined to be insufficient to meet demand (plus inefficient use and loss), all associated irrigation fluxes are scaled downward linearly by the provisional irrigation supply factor (X_{irr}). At this stage, WBM performs the river routing calculation, and estimates of provided water are updated according to actual water availability. Finally, excess water introduced to irrigated crop fields is partitioned between non-beneficial evaporation, non-consumptive runoff, and non-consumptive percolation. What follows is a more detailed description of each of these steps. Unless specified otherwise, all calculations described in this section are distributed spatially across irrigated crop areas as grid operations.

WBM can run any number of individual technologies simultaneously using data of irrigated land fraction for which each of the technologies is used

$$\begin{cases} \sum_i f_i^{d,irr} = 1 \\ \sum_i f_i^{a,irr} = 1 \end{cases} \quad \text{Eq. AII.1}$$

where $f_i^{d,irr}$ and $f_i^{a,irr}$ are fraction of land served by technology i within irrigated land, and superscripts d and a denotes delivery and application technology group, respectively.

Irrigation Delivery

Inefficient fluxes from conveyances rely on calculated daily open water evaporation rates (function of air temperature, humidity, and wind speed), and percolation rates of saturated soil. These rates are spatially and temporally distributed to the fraction of surface area of the irrigation delivery system ($f_i^{d,A}$) relative to the irrigated area (A^{irr}, m^2) for each i delivery technology. These non-beneficial fluxes are calculated at each pixel on each day crops demand irrigation water. Crop water demand functionality of WBM is described by Grogan *et al.* (2017). We assume that there is no surface runoff from any irrigation water delivery technology.

Evaporation of delivery water (N_{evap}^d) is calculated for days when irrigation demand is required as

$$N_{evap}^d = A_{fw} E_{fw} \quad \text{Eq. AII.2}$$

where E_{fw} is evaporation rate from free water surface (m/d), and A_{fw} is a weighted calculation of the pixel area undergoing free water evaporation through irrigation delivery systems:

$$A_{fw} = A^{irr} \sum_i^n f_i^{d,irr} f_i^{d,A} \varepsilon_i^{evap} \quad \text{Eq. AII.3}$$

where $f_i^{d,A}$ (-) is the fraction of area relative to irrigated area that irrigation delivery systems occupy on the ground, and ε_i (-) is a parameter that describes the fraction of an irrigation delivery technology that experiences free-surface evaporation. For the ε_i^{evap} parameter we suggest using values approaching 1.0 for ditch and canals (because both have water surface exposed for evaporation), and approaching 0.0 for pipe delivery technology as the only water exposed to air for evaporation in pipes consists of pipe leakage. All parameters can be spatially explicit; however, in our representation of the USRB only A^{irr} is spatially explicit; total irrigated areas and the fraction of delivery technologies are described in the main text. The fraction of canal areas are modified by the technology parameterizations (main text). The fraction of area coverage in the presence of a specific delivery type ($f_i^{d,A}$) and is assumed to be 1.2% for canals, which equates to an average 8 m wide canal traversing pixels that are completely irrigated. Defining spatially explicit estimates of canal coverage was beyond the scope of this study; but may be an important consideration for refinement of the baseline model.

Percolation is calculated from unlined irrigation conveyance (canal or ditch) benthic surface in a method similar to the calculation for evaporation.

$$L_{perc}^{d} = A_{perc} P_{perc} \qquad \text{Eq. AII.4}$$

where P_{perc} is percolation rate from the base of an irrigation delivery system to saturated soil, and A_{perc} is a weighted calculation of the pixel area undergoing saturated canal percolation under irrigation delivery systems:

$$A_{perc} = A^{irr} \sum_i^n f_i^{d,irr} f_i^{d,A} \varepsilon_i^{perc} \qquad \text{Eq. AII.5}$$

where ε_i^{perc} fraction of canal area to which percolation is applied by technology i. For the ε_i^{perc} parameter we suggest using 1.0 for ditch (no lining at the bottom of the ditch), a value representing the fraction of canal bottom areas in the domain that are un-lined (e.g. ~ 1 for canals assuming 100 % of bottom area are exposed to percolation in the e.g. USRB), and zero for pipe delivery technology as its water is isolation from percolation in pipes. The percolation factor for canals is adjusted in our technology parameterizations (main text). The P_{perc} rate is a specified parameter described in the main text.

Both N_{evap}^{d} and L_{perc}^{d} are scaled by the provisional supply factor (X_{irr}). It should be noted that L_{perc}^{a} is introduced to the model at the location of the irrigated fields and not explicitly at the locations of canals. Furthermore, water that percolates beneath canals is considered a non-consumptive loss associated with irrigated agriculture, and is therefore a component of irrigation reuse (R) described in the main text.

Irrigation Application

Process-based modelling of irrigation water losses by application technology is implemented following an approach similar to Jägermeyr *et al.* (2015). Differences between the two approaches reflect additional processes introduced here, as well as accommodating unique structures of the two hydrologic models.

The first stage of estimating inefficient fluxes during application of irrigation water is to estimate inefficient runoff from excess application, which follows calculation of crop irrigation requirement, and concurrent with estimation of inefficient delivery fluxes N_{evap}^{d} and L_{perc}^{d}. Excess irrigation supply (I^{a}), analogous to the *Application Requirements (AR)* parameter of

Jägermeyr *et al.* (2015), is calculated for each crop group (k, which can be either specific crop functional groups or pre-processed average land-cover groups described below):

$$I^a = \sum_i^n \sum_k^m \begin{cases} \max\left(0.5 S_{AWC}^k \overline{DU}_i - W_{irr} - L_{perc}^{rice}, 0.0\right) & where\ I^{demand,k} > 0 \\ 0 & where\ I^{demand,k} = 0 \end{cases} \quad \text{Eq. AII.6}$$

where S_{AWC}^k is a grid of crop (k) specific available water capacity (mm) that accounts for soil properties, $\overline{DU}_i$ is the application technology specific distribution uniformity coefficient (Jägermeyr et al., 2015), W_{irr} is the storage in the irrigation runoff retention pool (whose balance is calculated like the surface retention surface runoff pool of WBM, but applies only to the irrigated pixel fraction), and L_{perc}^{rice} is percolation associated with rice paddies, which is calculated separately (Grogan et al., 2017) and only applies over pixels with identified rice paddy, and $I^{demand,k}$ is the crop group specific irrigation demand. Existing storage in the irrigation runoff retention is subtracted assuming that irrigation requirements are reduced by whatever volume exists in pixels above field capacity assuming that existing excess volume in the irrigation retention pool is shared by all crops at a given pixel. Soil porosity defining soil saturation above field capacity is not presently a parameter input to WBM; therefore, we estimate the volume of additional water above field capacity that saturates soil as $0.5 S_{AWC}^k$. The distribution uniformity parameter ($\overline{DU}$) is a fraction of the crop field to which this soil saturation applies. $\overline{DU}$ for flood irrigation is close to 1 (all the soil in a crop area gets saturated) while for sprinkler irrigation about half of the possible saturation volume is actually applied. In the case of drip irrigation, a very small amount of water goes above W_{cap} and so $\overline{DU}$ is very low.

A fraction (ε_{mist}) of water delivered to irrigated crop fields can be lost non-beneficially above crop canopy from enhanced evaporation of, for instance, sprinkler mists. The flux of mist enhanced evaporation (N^a_{mist}) is calculated for each technology (i):

$$N^a_{mist} = (I^a + I^{demand,k})\varepsilon_{mist} \quad \text{Eq. AII.7}$$

Parameterization of ε_{mist} depends on local climate and specifics of sprinkler technology such that they can vary widely from 0 to 40%, with most analyses estimating losses to be less than about 5% (Bavi et al., 2009; McLean et al., 2000; Uddin et al., 2010). For the present study, we kept ε_{mist} at a constant value of 4% considered reasonable for the semi-arid region of the USRB, but reflects an important area to consider for either refining baseline representation, or improving overall water resource utilization (which we did not consider in this analysis).

Application and delivery inefficiencies are summed to net irrigation demanded by crops to estimate an initial gross irrigation flux (G^*):

$$G^* = I^{demand} + I^a + N^a_{mist} + N^d_{evap} + L^d_{perc} \quad \text{Eq. AII.8}$$

A variety of functions are associated with sourcing available irrigation water in WBM, which yield a fraction of available water (X_{irr} where $X_{irr} = 1$ indicates complete availability) from an appropriate distribution of sources. Typical irrigation source water determination is discussed in Grogan *et al.* (2017), and modified here to assign specific supply reservoirs to areas of the simulation (described below). Where water supply is less than complete ($X_{irr} < 1$), all terms above are reduced linearly to utilize available supply via:

$$I^{demand} *= X_{irr} \quad \text{Eq. AII.9}$$

$$I^a *= X_{irr} \quad \text{Eq. AII.10}$$

$$N^a_{mist} *= X_{irr} \quad \text{Eq. AII.11}$$

$$N^d_{evap} *= X_{irr}$$ Eq. AII.12

$$L^d_{perc} *= X_{irr}$$ Eq. AII.13

Actual gross irrigation (G) is calculated following routing later in the time-step, and small deviations between estimated and actual water availability are accounted for in subsequent timesteps.

Following routing through the stream network, the water balance of irrigation retention pool (W_{irr}) is updated using a stable solution and follows a conceptual order of flux priorities. The change in volume of W_{irr} is governed by the differential equation:

$$\frac{dW_{irr}}{dt} = I^{atm} + I^a - N^a_{evap} - L^a_{perc} - L^a_{rnff}$$ Eq. AII.14

where I^{atm} is water incident to irrigated crop fields from natural precipitation or melt, N^a_{evap} is non-beneficial evaporation from saturated soil surface, L^a_{perc} is percolation from saturated soils to groundwater, and L^a_{rnff} is surface runoff from saturated soil. The stock of W_{irr} at the end of the timestep is calculated in four independent steps (denoted by superscripts):

1) $W^1_{irr} = W^0_{irr} + I^{atm} + I^a$
2) $N^a_{evap} = \min\left(A_{irr}\overline{DU} \times E_p, W^1_{irr}\right)$
   $W^2_{irr} = W^1_{irr} - N^a_{evap}$
3) $L^a_{perc} = \min\left(A_{irr}\overline{DU} \times P_{perc}, W^2_{irr}\right)$ Eq. AII.15
   $W^3_{irr} = W^2_{irr} - L^a_{perc}$
4) $L^a_{rnff} = \min\left(A_{irr}\beta_{surf} \times \sqrt{2g} \times \frac{W^3_{irr}}{A_{irr}}, W^3_{irr}\right)$
   $W_{irr} = W^3_{irr} - L^a_{rnff}$

where W^0_{irr} is the stock of the water retention pool at the end of the previous timestep, E_p is the potential evapotranspiration (mm/d), β_{surf} is the parameter describing the rate of leakage from

the irrigation (and surface) retention pools, and g is the constant of gravitational acceleration. The order of updating the irrigation retention pool gives first precedence to non-beneficial evaporation, and lowest precedence to surficial runoff, so non-consumptive losses may biased low. The proportions of delivery technologies were spatially homogenous and reflected the average lengths of technologies in the USGS National Hydrography Dataset (nhd.usgs.gov). The relative proportions of application technology varied by county following USGS surveys (Maupin et al., 2014; Dieter et al., 2018).

Lumped aquifer representation

New functionality was introduced to WBM to account for large aquifers using a lumped aquifer representation with unidirectional vertical movement. Lumped aquifers can be represented over all or portions of the model domain. Recharge percolating through the root zone is proportioned between shallow groundwater (γ_{SGW} between 0.05 and 0.08) and the deeper (lumped) aquifer ($\gamma - \gamma_{SGW}$) at each pixel overlying an identified aquifer. Additionally, inflows from the surface flow network can be specified as point-based losing reaches that infiltrate directly to the aquifer (bypassing the shallow groundwater pool); flows to the ESPA are parameterized as a fraction of daily flow. Outflows from the aquifer occur as springs represented as points with head-dependent conductance similar to drains in MODFLOW (Harbaugh, 2005). Average head within the lumped aquifer head is calculated as:

$$h = \frac{S_A}{C_A} * Z_A + Z_0, \qquad \text{Eq. AII.16}$$

where h is aquifer head (m), S_A is the volume stored within the aquifer (km^3), C_A is the capacity of the aquifer (km^3) (so the ratio of $\frac{S_A}{C_A}$ is the fractional storage), Z_A is the aquifer thickness (m) and Z_0 is the base elevation (m). Drainage through individual springs (Q_{spr}) is calculated as:

$$Q_{spr} = K_{spr}(h - Z_{spr}),$$ Eq. AII.17

where K_{spr} is an individual spring's conductance ($m^2\ d^{-1}$), and Z_{spr} is the elevation of each spring (m). Q_{spr} is then summed for all individual springs. All recharge to and abstractions from the aquifer are summed through the previous day and mass balance of the aquifer is updated at a daily time-step using a Runge-Kutta 3(2) order (Bogacki-Shampine) scheme. Under this split operator solution, water percolating to and pumped from the aquifer is assumed to influence aquifer volume following a one day lag. The single-day lag is expected to underestimate percolation travel-times through the unsaturated zone and the far-field hydrodynamic response of the aquifer to changes in pumping. The volume of water represented by the lumped aquifer model is assumed not to interact with shallow or root zone water (i.e. head is assumed to remain below the base of these zones) and fluxes from the aquifer to these zones are neglected. In the USRB, this is a reasonable assumption over most of the aquifer where vadose zones are fairly thick and dry (Whitehead, 1992).

The extent of the lumped aquifer was the same as that used for the ESPAM2 (IDWR, 2013). We represented the aquifer as two lumped compartments (Figure 11b) to reflect the two types of water identified by Plummer *et al.* (2000), such that the ESPA was disaggregated to upgradient (northeast) and downgradient (southwest) sections just upgradient of Magic Valley irrigated croplands. Inflows into the upgradient portion consisted of natural recharge, percolation as reach gains from six losing rivers of the surface flow network (Big Lost, Little Lost Rivers, Birch, Medicine Lodge, Camas Creeks, and the Snake River), and incidental recharge from irrigation. The downgradient portion received flow from upgradient portion of the ESPA, as well as natural and incidental recharge. Storage parameters were established from several sources (Garabedian, 1992; Whitehead, 1992; IDWR, 2013). For the upgradient section (Group 1 of Plummer), we

selected a specific yield (0.06) and thickness of the aquifer (250 m) comparable to these studies resulting in an average aquifer storage of about ~330 km^3, which is less than half of estimates of the total recoverable water volume of the aquifer (Robertson et al., 1974). The downgradient portion of the ESPA was attributed with a specific yield of 0.05 and thickness of 220 m resulting in an average storage of about 73 km^3.

Springs draining the ESPA came from a detailed study of the Thousand Springs region between Twin Falls and King Hill, ID (Covington and Weaver, 1991). Elevations of the springs decrease linearly along a westward head gradient, and spring elevations input to WBM subtracted out this average gradient such that each spring elevation only reflected deviations from an average head elevation of 828 m. In other examples of spatially lumped aquifers (e.g. Famiglietti and Wood, 1994; van der Velde et al., 2009), statistical or functional accounting for spatial differences in the water table are used adding additional dynamism not simulated here. However, the majority of points of discharge from the ESPA are at known elevations that follow a longitudinal gradient. Therefore, the linear transformation of the outlet elevation of the springs is simplifications appropriate to the ESPA where likely geostatistical or functional methods are more appropriate for spatially distributed water tables in the absence of specific known points of groundwater outflow.

The ESPA is known to be hydraulically connected to the ESPA in the vicinity of the American Falls Reservoir (Garabedian, 1992; IDWR, 2013), which we represent with an additional surface flow sink from the Snake River just upstream of American Falls Reservoir to the upgradient ESPA aquifer, and a spring from the ESPA back to the Snake River at the reservoir. Parameterizing these flow paths was conducted manually primarily by matching the time-series of storage with the American Falls reservoir. Managed aquifer recharge to the ESPA was

parameterized as an increased fraction of flow entering ESPA from the Snake River just upstream of American Falls.

Crop classifications

Crop Data Layer (CDL - Han et al., 2012) provided spatiotemporally explicit data for crop cover in our simulation domain; however, we utilized crop parameterizations from Monthly Irrigated and Rainfed Crop Atlas (MIRCA2000 - Portmann et al., 2010) to simulate crop water use (Grogan et al., 2017). To utilize CDL, we remapped crops according to the groupings in Table AII.1. Moreover, for simulations presented here we pre-processed these crop data to calculate weighted averages of each of the fundamental parameters associated with crop water use for rainfed and irrigated crops. Calculations of irrigation demand, or inefficient irrigation water utilize either crop specific parameters or the average parameters with no fundamental change in calculation method, totals of all fluxes remain the same; however, attribution of water fluxes to specific crops is not possible when averaged inputs are used. Crop parameterization follows Portman et al. (2010) and Siebert and Döll (2010).

Table AII.1: Crop parameters used in study including crop data layer (CDL) crop identifiers as mapped to MIRCA2000 (Monthly Irrigated and Rainfed Crop Atlas) crops and associated planting parameters. Parameters of season length ($\boldsymbol{L_{ini}}$, $\boldsymbol{L_{dev}}$, $\boldsymbol{L_{mid}}$, $\boldsymbol{L_{late}}$) and crop factors ($\boldsymbol{Kc}$) at various stages in the growing season, and crop depletion factor (CDF), are as defined by Grogan et al.(2017) and Siebert and Döll (2010).

Crop	MIRCA ID	CDL ID	L_{ini}	L_{dev}	L_{mid}	L_{late}	Kc_i	Kc_m	Kc_e	Root Depth (mm)	CDF[1]
Wheat	1	22-24,30	0.15	0.25	0.40	0.20	0.40	1.15	0.30	1250	0.55
Maize	2	1,12,13	0.17	0.28	0.33	0.22	0.30	1.20	0.40	1000	0.55
Barley	4	21	0.15	0.25	0.40	0.20	0.30	1.15	0.25	1000	0.55
Potatoes	10	43	0.20	0.25	0.35	0.20	1.15	0.50	0.40	400	0.35
Sugarbeet	13	41	0.20	0.25	0.35	0.20	0.35	1.20	0.80	700	0.55
Canola	15	31,34,38	0.30	0.25	0.30	0.15	0.35	1.10	0.35	1000	0.60
Pulses	17	42,51,52,53	0.18	0.27	0.35	0.20	0.45	1.10	0.60	550	0.45
Other Perennial	24	+	0.00	0.00	1.00	0.00	1.00	1.00	1.00	800	0.50
Fodder Grasses	25	36,37,58-60	0.00	0.00	1.00	0.00	1.00	1.00	1.00	1500	0.55
Other Annual	26	++	0.15	0.25	0.40	0.20	0.40	1.05	0.50	1000	0.55

[1] Crop depletion factor

\+ 55,56,66-68,71,74-77,204,210,211,216-218,220,221,223,242,250

++ 11,14,25,28,32,33,35,39,44,46-50,54,57,205-209,213,214,219,222,227,229,231,243-249

Spring outflow data

Outflow of the Eastern Snake Plain Aquifer (ESPA) to the Snake River along the margins of the Snake River canyon consist predominately of flow out of large springs (Q_{spring}). Spring out flows were provided by J. Sukow (pers. comm., updated from Sukow, 2012) which are calculated as

$$Q_{spring} = Q_{4500} - Q_{0000} - Q_{8150} - Q_{2500} - Q_{ss} - Q_{NRf} + Q_{Div} \quad \text{Eq. AII.18}$$

Where Q_{4500} is the Snake River discharge at King Hill (USGS 13154500), Q_{0000} is Snake River discharge at Kimberly (USGS 13090000), Q_{8150} Salmon Falls Creek discharge (13108150), Q_{2500} is Malad River discharge (USGS 13152500), Q_{SS} is discharge from the South Side canal system, Q_{NRf} is discharge from the North Side canal return flows, and Q_{Div} are diversions from the Snake River between the two reaches. Data provided by J. Sukow contain provisional and interpolated estimates for some flow components.

Validation of Water Balance Model for baseline

County-wide gross and surface irrigation in 2010 and 2015 simulated by WBM (Figure AII.3) was biased low from USGS estimates (-34.9%). The NSE of 0.267 signifies that the spatial variability in irrigation demand was captured despite the low bias. USGS USCO records counties at which diversions were made (Dieter et al., 2018), and WBM tracks counties at which water was used, a discrepancy likely responsible for some of the error. WBM predicted the long-term mean in spring discharge from the ESPA with a percent bias (PBias) of -0.78% but under-predicted seasonal variability (Figure AII.4), leading to a low NSE (0.112). Seasonal storage within the three Snake River reservoirs (Figure AII.5) using observed discharge at the reservoir outflow was accurate at the headwaters of the Snake River in Wyoming (Jackson Lake), though

the representation departed from observations downstream (Palisades, and American Falls), which was attributed to cascading errors in both structure and input precipitation through the network. During the simulation period, overall performance of reservoir volume was characterized with a negative NSE but a PBIAS of only 5.1%. Seasonal discharge in headwaters of the Snake River (13010065, Snake River, Flagg Ranch, Wyoming) was accurately represented [NSE=0.9, PBIAS=5%]; however, simulated discharge in smaller streams in the vicinity of the ESPA (13137500, Trail Creek, Ketchum, ID, and 13039500, Henry's Fork, Lake Idaho) was generally biased high, and exhibited stronger seasonal cycling than observations (Figure AII.5b). The high bias, and exaggerated seasonal cyclicity in discharge is a common observation of WBM's representation of small watersheds; additional damping of discharge occurs through routing through the river network and especially at dams. Both Trail Creek and Henry's Fork have dammed reservoirs upstream of the gaging stations, and simulations may have underestimated the influence of dam operations.

Table AII.2: Summary of observations used for assessing model performance at baseline. Seasonal averages are calculated over meteorological seasons (Winter: DJF, Spring: MAM, Summer: JJA, Autumn: SON).

Metric	Location and timestep of observation	unit	Number of observation	PBIAS	MPE	NSE	RMSE
Discharge from springs	Monthly sum of discharge from springs (2008-2015), missing 15 months	m^3/month	69	-0.78	-0.6	0.112	55,000
Gross and surface irrigation	Annual sums for 24 USRB counties for total and surface supply (2010 and 2015)	km^3/year	96	-34.9	-31.8	0.267	0.369
Headwater discharge	Seasonal means for 3 stations (2008-2015)	m^3/s	84	30.1	354	0.396	9.54
Reservoir storage	Seasonal means for 3 reservoirs (2008-2015)	m^3	84	5.07	13.4	-0.409	4.5E+08

Note: PBIAS: percent bias, MPE: model percent error, NSE: Nash-Sutcliffe efficiency, RMSE: root mean squared error.

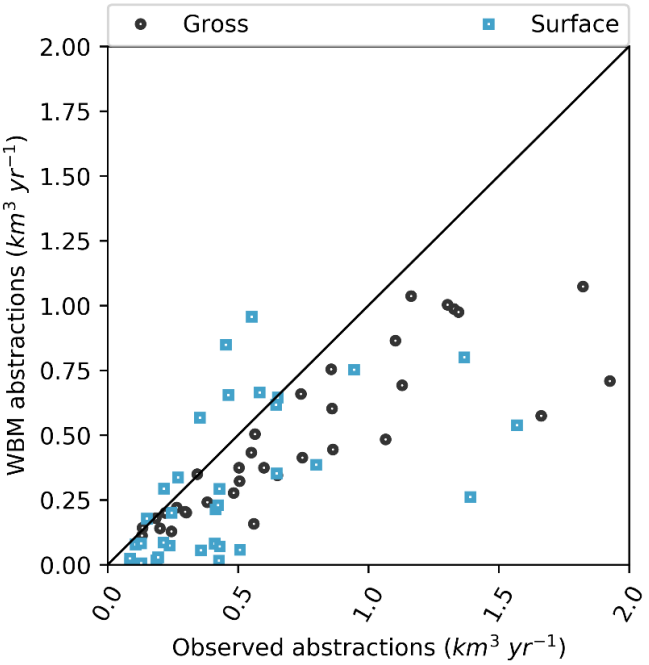


Figure AII.3: Correlation between model and USGS estimated county-wide gross and surface water irrigation water use in 2010 and 2015.

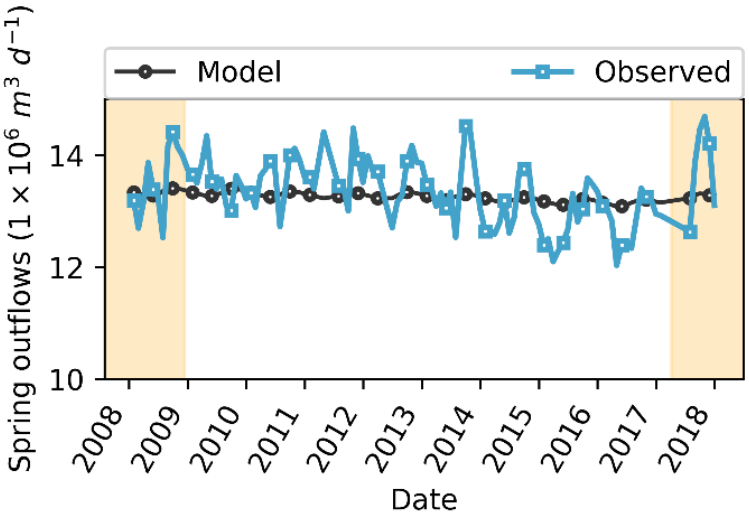


Figure AII.4: Time-series of spring discharges.

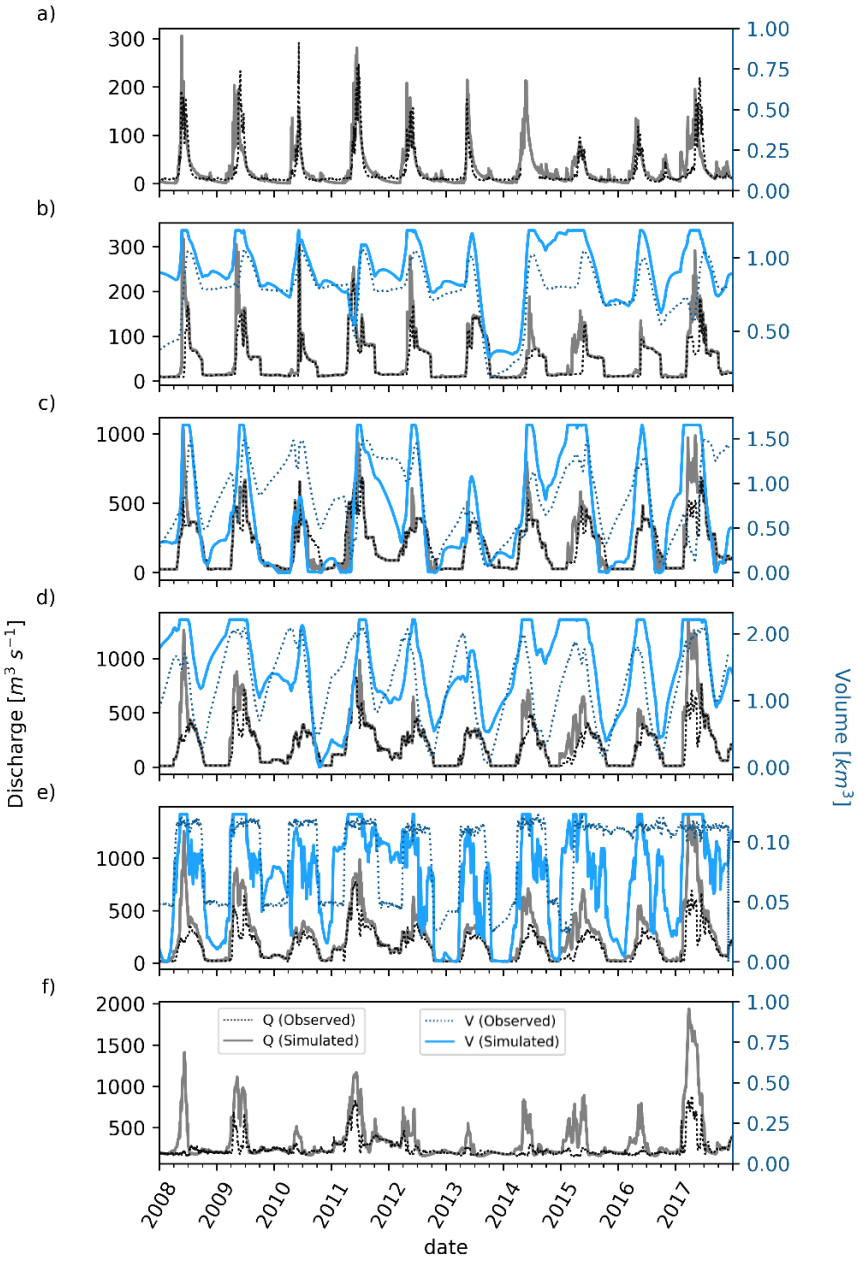


Figure AII.5: Time series of discharge and reservoir volume at six locations on the Snake River. WBM predictions as solid lines, observations as dotted lines. Reservoir volume (blue) observations from the USBR HydroMet network, and discharge (black/gray) from USGS gaging stations. Discharge gaging stations are located as close to immediately downstream of respective reservoirs. The six site locations (all on Snake River) and their respective USGS and USBR (for b-e) station identifiers, Nash-Sutcliffe efficiencies, and percent bias are (a) Flagg Ranch, WY (13010065: 0.72, 13.5%), (b) Moran, WY and Jackson Lake (13011000: 0.63, 23%, JCK: 0.04, 17%), (c) Irwin, ID and Palisades Reservoir (13032500: 0.79, 5.0%, PAL: -1.0, -22%), (d) Neeley, ID and American Falls Reservoir (13077000: 0.42, 25%, AMF: -0.05, 17%), (e) Rupert, ID and Walcott Lake (13081500: -0.26, 52%, MIN: -0.4,-20%), and (f) King Hill, ID (13154500: -3.7, 40%).

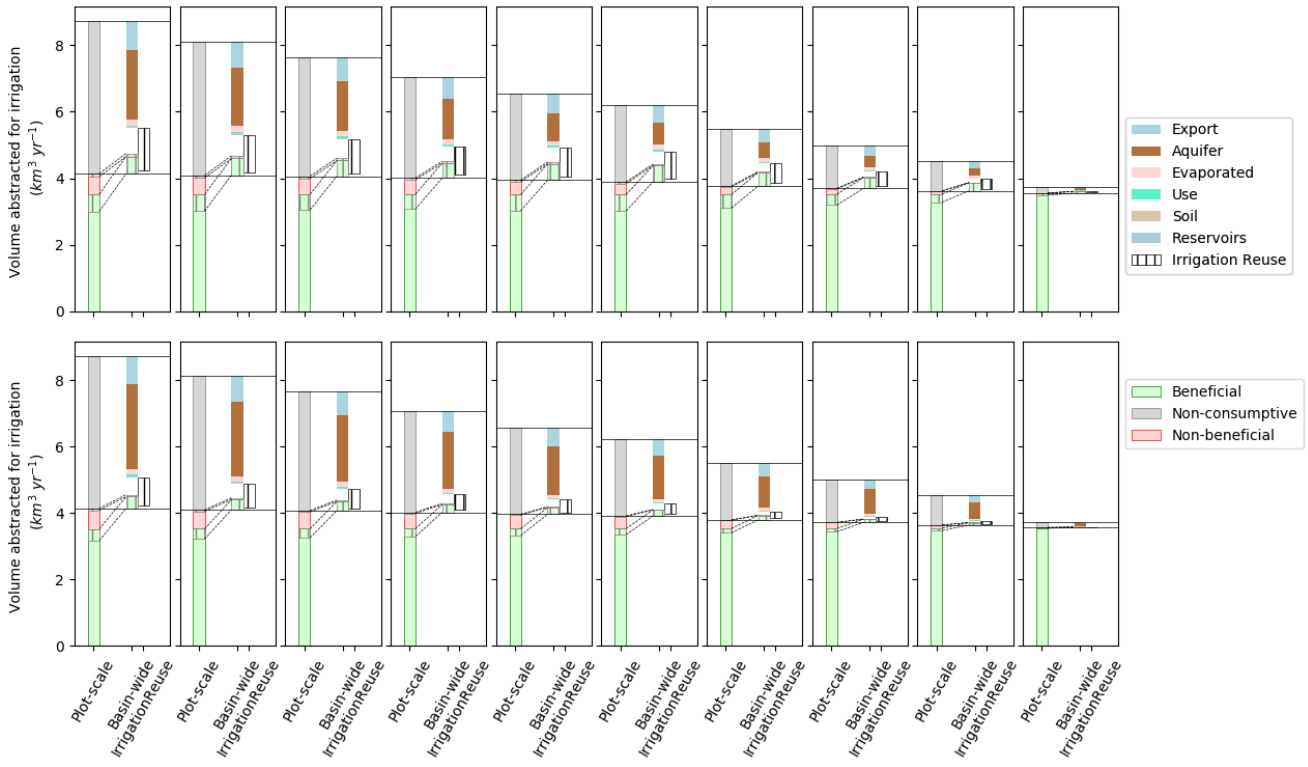


Figure AII.6: Comparison between component fraction (summed for basin) and irrigation return fates across the USRB under the all parameterizations. Top row (a-j) with enhanced aquifer recharge, bottom row without (k-t). Baseline technology in first column (a,k), and increasing modernization scenarios in subsequent columns (Eff.A:b,l, Eff.B:c,m, … , Eff.I:j,t).

APPENDIX III

METHODOLOGICAL DETAILS AND DEMONSTRATION THAT GREATEST NITRATE REMOVAL OCCURS WHEN TREATMENT AREA IS MAXIMIZED

Representing subsurface drains

To represent subsurface drains (SSD), WBM's infiltration fraction (f_{inf}) was lowered according to the fraction of soil areas underlain by subsurface drainage. Variables are defined in Table AIII.1. Surplus water leaving the rooting-zone is directed towards infiltration to shallow groundwater, while the complement runs off as quickflow through the rooting zone. Infiltration fraction was modified by the presence of SSD from a baseline (f_{inf}^*) representing an assumed spatially constant infiltration fraction under natural conditions according to Equation AIII.1.

$$f_{inf} = f_{inf}^* \, (1 - \theta \, f_{ssd}) \qquad \text{Eq. AIII.1}$$

In Equation AIII.1, θ represents an efficiency of SSD to prevent percolation to groundwater which we set to a value between 80% and 90% such that the infiltration fraction never falls below 10% surplus soil water when a pixel is completely covered by tile-drained crops. Moreover, the quickflow time-constant (α_{qf}) in WBM is increased proportionally from a baseline (α_{qf}^*) by the presence of SSD assuming that percolation intercepted by subsurface-drains arrives at treatment wetlands (α_{ssd}) within one day Equation AIII.2.

$$\alpha_{qf} = \alpha_{qf}^* \, (1 - \theta \, f_{ssd}) + \alpha_{ssd} \, \theta \, f_{ssd} \qquad \text{Eq. AIII.2}$$

We assumed that all SSD was emplaced in existing crop land and calculated the intersection of crop limited SSD (A_{cSSD}) by Equation AIII.3:

$$A_{cSSD} = \min(A_{SSD}, A_{crop}) \qquad \text{Eq. AIII.3}$$

where A_{crop} is the pixel area fraction covered by cropland.

Table AIII.1: Variables used to describe fluxes and parameters in the representation of sub-surface drains and treatment wetlands.

Variable	Units	Definition
I_{wl}	$m^3\ d^{-1}$	Water inflow into the wetlands
V_{avail}	m^3	Capacity available within wetland below the maximum flow depth (d_max)
Q_{ro}	$m^3\ d^{-1}$	Runoff from the upland portions of the pixel
dt	d	Duration of timestep
d_{max}	m	Maximum depth of flow treatment wetlands can store
d_{wl}	m	Depth of flow in treatment wetlands at time t
χ_{up}	-	Multiplier dictating the maximum crop area that can drain through a wetland
χ_{buffer}	-	Fraction of total wetland area reserved as buffer
χ_{bf}	-	Fraction of baseflow routed towards wetlands. (Complement is routed to streams)
A_{wl}	-	Total wetland area pixel fraction (active area plus buffer area)
A_{wlA}	-	Active wetland area pixel fraction
A_{wlN}	-	Natural wetland area pixel fraction
A_{wlT}	-	Treatment wetland area pixel fraction
A_{wlR}	-	Riparian marginal wetland area pixel fraction
A_{crop}	-	Total crop area pixel fraction
Q_{wl}	$m^3\ d^{-1}$	Discharge from outlet of the wetland
V_{wl}	m^3	Volume stored in wetland (at time t)
α_{wl}	d^{-1}	Wetland release time-constant
v_f	$m\ d^{-1}$	Uptake velocity describing DIN removal by denitrification (Wollheim et al. 2006)
k_{den}	d^{-1}	Denitrification rate constant
b_{vf}	-	Constant in regression between ln (v_f) and ln (C_din)
a_{vf}	-	Slope in regression between ln (v_f) and ln (C_din)
C_{nitr}	$g\ m^{-3}$	DIN concentration in wetland pool
M_{nitr}	g	DIN mass in wetland pool
$\dot{m}_{in}$	$g\ d^{-1}$	DIN mass input flux to wetland pool
$\dot{m}_{den}$	$g\ d^{-1}$	DIN mass removal flux via denitrification from the wetland pool
t	d	Time available for reaction $t = \min(dt, V\_wl\ /\ Q\_wl)$
T_{wl}	°C	Water temperature in wetland
T_{ref}	°C	Reference water temperature

Representing upland wetlands

Treatment wetlands are represented using a new pool within WBM that receives flow from upland portions of each pixel. To make the model as parsimonious as possible, the treatment

combines riparian marginal processing, natural wetlands, and restored treatment wetlands in a single pool. We acknowledge that engineering constraints imposed on wetland functionality presented below may be inappropriate for riparian areas and natural wetlands. Our formulation should conservatively estimate the amount of natural and riparian wetland denitrification that occurs because our assumptions of buffer area limit available active area for denitrification from existing natural wetlands.

The water and nitrate mass balance of the wetland pool are governed by Equations AIII.4 and AIII.5 respectively.

$$dV / dt = I_{wl} - Q_{wl} \qquad \text{Eq. AIII.4}$$

$$dM_{din} / dt = \dot{m}_{in} - \dot{m}_{den} - \dot{m}_{out} \qquad \text{Eq. AIII.5}$$

The following paragraphs describe our solution of Equations.AIII.4 and AIII.5.

Flow through the wetland pool is defined by both a maximum area of upstream crop area draining into a wetland, and a maximum flow-depth (d_{max} [m]). Flow from untreatable crops, and flow exceeding the maximum depth is routed immediately beyond the wetland to the stream (by-passing flow). Flow entering the wetland experiences a detention time specified by the wetland flow time-constant (α_{wl} [d^{-1}]) defined as the inverse of the hydraulic residence time within the pool. Our assumption of a well-mixed wetland system minimizes the removal calculated for a simple reactor (Levenspiel, 1999), but results in the proportion of mass denitrified to be consistent with previous studies of surface flow wetlands (Mitsch and Gosselink, 2007). Terms in the daily water balance are solved sequentially to facilitate an estimation of bypassing flow and to ensure numerical water balance. The three stages of the calculation include estimation of

bypassing flow (Equation AIII.6A), calculation of outflow assuming inflows from the mid-point of the time-step (Equation AIII.6B), and updating storage balance (Equation AIII.6C).

$$I_{wl} = \min (V_{avail} / Q_{ro} dt , (\chi_{up} A_{wl}) / A_{crop}) Q_{ro} \quad \text{Eq. AIII.6A}$$

$$Q_{wl} = (V_{wl}^{t} + 0.5 I_{wl} dt) \alpha_{wl} \quad \text{Eq. AIII.6B}$$

$$V_{wl}^{t+1} = V_{wl}^{t} + I_{wl} dt - Q_{wl} dt \quad \text{Eq. AIII.6C}$$

The split-operator solution for wetland water balance is needed to accommodate WBM's source tracking functionality (Grogan et al., 2022). The volume available (V_{avail}) below the specific maximum depth of flow (d_{max}) is calculated at each time-step (Equation AIII.7) and accounts for short-term storage over the duration of the time-step.

$$V_{avail} = max(0, (1 + \alpha_{wl} dt) A_{wl} A d_{max} - V_{wl}) \quad \text{Eq. AIII.7}$$

Runoff from the upland portions of the pixel arrive via quick-flow (Q_{qf}) in WBM to be typical for treatment wetlands (Tanner et al., 2010). Although WBM assumes a single common water surface and DIN concentration for tractability, in practice multiple separate wetlands would be necessary to intercept leachate prior to loading to streams. The total wetland area is the sum of natural wetlands (A_{wlN}) and treatment wetlands (A_{wlT}) via experimental restoration scenarios. A minimum pixel fraction covered by riparian marginal wetlands (A_{wlR}) is included to capture riparian storage and denitrification processes in the absence of natural or restored wetlands (Equation AIII.8).

$$A_{wl} = min(A_{wlR}, A_{wlN} + A_{wlT}) \quad \text{Eq. AIII.8}$$

A_{wl} represents the total wetland area, but a fraction (χ_{buffer}) of this total represents a buffer area that protects the wetland system. The complement fraction of active wetland area (A_{wlA}) holds

water in storage where denitrification occurs. The active water-holding portion of wetlands (A_{wlA}) are assumed to have vertical banks (benthic area equals the active wetland area). The riparian marginal wetlands are included in our model to account for riparian removal, and is represented as a surface process for simplicity; however, riparian processing also occurs in subsurface environments (Groffman et al., 2000; Lutz et al., 2020; Mayer et al., 2007), but a separate parameterization for such a process was beyond the purpose of this study. We acknowledge that some of the large flux of subsurface nitrate removal estimated by the model may represent a greater amount of riparian removal along subsurface flow-paths, and should be investigated in future work. The riparian marginal area is parameterized as a fraction of the pixel, and is expressed as a buffer width (w_{wlR}) from stream margins (Equation AIII.9) that intercepts runoff prior to entering streams.

$$w_{wlR} = (A_{wl} A) / (2 l_{stream}) \quad \text{Eq. AIII.9}$$

Nitrate balance in wetlands is calculated with advective fluxes in and out of the pool, and a denitrification flux removal calculated using a first order rate constant (k_{den} [d^{-1}]) within the pool. We assume that denitrification occurs along the benthic surface so that the first order process is parameterized with an uptake velocity v_f that is updated daily based on wetland nitrate concentration and temperature according to an efficiency loss parameterization (Wollheim et al., 2014) (Equation AIII.10).

$$v_f = \exp(\ln(b_{vf}) + a_{vf} \ln(C_{nitr})) \times 2^{((Twl - Tref) / 10)} \quad \text{Eq. AIII.10}$$

The relationship between uptake velocity and concentration is a log-linear relationship in both lotic (Mulholland et al., 2008; Wollheim et al., 2008a) and lentic (Wollheim et al., 2014) systems, and is adjusted according to Q10 temperature reactivity following (Kadlec and Reddy,

2001), where water temperature is assumed to be the temperature of soil water at the percolation depth of 1.5 m calculated by AgroIBS. The denitrification rate constant is then calculated from wetland depth (Equation AIII.11).

$$k_{den} = v_f \, d_{wl} \qquad \text{Eq. AIII.11}$$

Nitrate balance within the wetlands is calculated following the calculation of flow and water storage. Nitrate mass balance is updated in a split operation to ensure mass balance. Nitrate concentration (C_{nitr} [mg L^{-1}]) within the pool is estimated assuming input from the mid-point of the time-step (Equation AIII.12A), removal and outflow are estimated (Equations AIII.12B,12C), and then total mass balance is updated Equation (AIII.12D).

$$C_{nitr} = (M_{nitr}^{t} + 0.5 \, m_{ro} \, dt) \, / \, V_{wl}^{t} \qquad \text{Eq. AIII.12A}$$

$$\dot{m}_{den} = (1 - \exp(-k_{den} \, t)) \, (M_{nitr}^{t} + 0.5 \, m_{in} \, dt) \qquad \text{Eq. AIII.12B}$$

$$\dot{m}_{out} = C_{nitr} \, Q_{wl} \qquad \text{Eq. AIII.12C}$$

$$M_{nitr}^{(t+1)} = Mnitr^{t} + m_{in} \, dt - m_{den} \, dt - m_{out} \, dt \qquad \text{Eq. AIII.12D}$$

Mass flux is then routed to the stream network within the pixel. Comparable nitrogen dynamics are formulated within the stream system (Samal et al., 2017; Stewart et al., 2011). The mass flux from soil leachate (m_{ro}) is provided from the crop area weighted average of Agro-IBIS leachate described above. WBM and Agro-IBIS do not presently have coupled soil water pools, so that daily differences in percolation volume can create numerical instability between water volume and nitrate leachate mass. WBM exhibited flashier hydraulic response and greater retention of water within the soil pool between storm events. To ensure smooth functionality we introduce a

holding pool for Agro-IBIS soil leachate and introduce all accumulated mass of nitrate from this holding pool to the wetlands when WBM introduces runoff.

Nitrate leachate that percolates below croplands is detained to account for travel through the vadose zone, and to account for the differences between WBM's hydrodynamic response of the shallow groundwater pool, and the solute travel-time through shallow groundwater. Shallow groundwater hydrodynamic response is controlled by a first-order time-constant $\alpha_{sgw} = 0.025$ d^{-1} that captures typical baseflow recession rates in temperate climates (Grogan et al., 2022). A lag between nitrate percolation and nitrate recharge of shallow groundwater (m_{rech}) is accounted for using an iterative exponential weighting function (Sangrey et al., 1984; Venetis, 1969; Neitsch et al., 2011) given by Equation AIII.13.

$$m_{rech}^{t} = [\ 1 - \exp(\ -\alpha_{vzl}\ )\ ]\ \ m_{leach} + \exp(\ -\alpha_{vzl}\ )\ m_{rech}^{(t-1)} \qquad \text{Eq. AIII.13}$$

We assume α_{vzl} equal to 0.00125 d^{-1} to capture transit time of solutes through typical catchments (Benettin et al., 2015; Berghuijs and Kirchner, 2017). DIN experiences denitrification in the shallow groundwater pool prior to discharge back to the surface. A constant Dämkoholer number of 0.29 is assumed for the shallow groundwater (Green et al., 2008) equating to 22% denitrification of DIN in this pool.

Demonstrating that greatest nitrate removal occurs when treatment area is maximized

Assume a pixel of area (A_p) completely covered in crop such that the area of crop (A_c) initially equals the area of the pixel ($A_c = A_p$). An area of the pixel (A_w) is restored to wetland by replacing crops such that $A_c = A_p - A_w$ and the pixel fraction of wetland area is specified as ($f_w = A_w/(A_p - A_w)$). The wetland has a catchment area proportional to its size, which is imposed as a design constraint of the restoration practice, such that crop leachate from within

this catchment area (A_t) is treated. The wetland is designed such that it occupies a specified fraction of its catchment ($f_c = A_w/A_t$). The fraction of the total crop area treated ($f_t = A_t/A_c$) is ratio between the restored wetland pixel fraction, and the wetlands catchment fraction Equation AIII.14.

$$f_t = \frac{A_t}{A_c} = \left(\frac{A_w}{A_c}\right)\left(\frac{A_w}{A_t}\right)^{-1} = \frac{f_w}{f_c} \qquad \text{Eq. AIII.14}$$

The relationship between the above area fractions is presented schematically in Figure AIII.1.

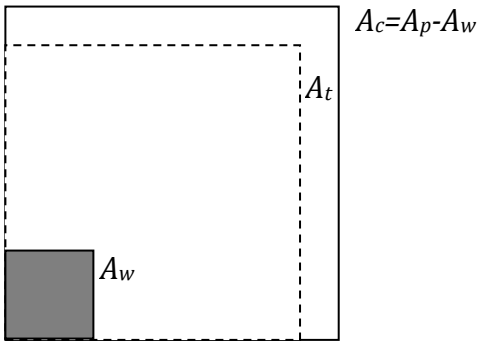


Figure AIII.1: Relationship between restored wetland and crop landscape area.

The total mass denitrified ($\dot{m}_{dnt}$) is proportional to the mass treated and the removal. The solution for removal (r) of a solute mediated by reactions along a benthic surface is dictated by the uptake velocity (v_f $[m\ y^{-1}]$) and hydraulic load (H_l $[m\ y^{-1}]$) given by Equation AIII.15 (Wollheim et al., 2006).

$$r = 1 - \exp\left(-\frac{v_f}{H_l}\right) \qquad \text{Eq. AIII.15}$$

Uptake velocity is dependent on concentration, which we assume is independent of catchment size (all crops in the pixel leach at a uniform rate of $\dot{m}_{lch}$ $[kg\ y^{-1}]$). Hydraulic load is given by

the flow entering the wetland ($Q\ [m^3\ y^{-1}]$) divided by its benthic surface area ($A_w\ [m^2]$) and is a function of the wetland catchment area (Equation AIII.16).

$$H_l = \frac{Q}{A_w} = \frac{q\,A_t}{A_w} = \frac{q}{f_c} \qquad \text{Eq. AIII.16}$$

In A3, q is a uniform runoff rate ($m\ y^{-1}$). Incorporating our definitions in AIII.16 into AIII.15, and relating to total crop leachate, we define the mass flux of denitrification in Equation AIII.17.

$$\dot{m}_{dnt} = \dot{m}_{lch} f_t \left(1 - \exp\left(-\frac{v_f f_c}{q}\right)\right) = \dot{m}_{lch} \frac{f_w}{f_c}\left(1 - \exp\left(-\frac{v_f f_c}{q}\right)\right) \qquad \text{Eq. AIII.17}$$

We define the ratio of total leachate removed by denitrification (R_{dnt}) in Equation AIII.18.

$$R_{dnt} = \frac{\dot{m}_{dnt}}{\dot{m}_{lch}} = \frac{f_w}{f_c}\left(1 - \exp\left(-\frac{v_f f_c}{q}\right)\right) \qquad \text{Eq. AIII.18}$$

The maximum fraction of leachate denitrified occurs when $f_w = f_c$, or when the wetland catchment area covers the entirety of the pixel. As the area fraction of the wetland inside its catchment (f_c) increases, the negative term becomes smaller, and more mass is removed. However, for any given increase in f_w up to f_c there is no mechanism by which denitrified fraction could increase that would make treating less mass more favorable. Any change in f_c to increase removal has a commensurate and necessarily larger decrease in the treated mass.

LIST OF REFERENCES

Abatzoglou, J. T.: Development of gridded surface meteorological data for ecological applications and modelling, Int. J. Climatol., 33, 121–131, https://doi.org/10.1002/joc.3413, 2013.

Adam, J. C., Haddeland, I., Su, F., and Lettenmaier, D. P.: Simulation of reservoir influences on annual and seasonal streamflow changes for the Lena, Yenisei, and Ob' rivers, 112, https://doi.org/10.1029/2007JD008525, 2007.

Akaike, H.: A new look at the statistical model identification, 19, 716–723, https://doi.org/10.1109/TAC.1974.1100705, 1974.

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., and others: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 300, 1998.

Allison, I., Intergovernmental Panel on Climate Change, and Climate Change Research Centre: The Copenhagen diagnosis, 2009: updating the world on the latest climate science, University of New South Wales, Climate Change Research Centre, Sydney, Australia, 2009.

Almada, M. and Attux, R.: Ethical design of social simulations, in: Anais do Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software (WASHES), Anais do III Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software, https://doi.org/10.5753/washes.2018.3471, 2018.

Ameli, A. A. and Creed, I. F.: Does Wetland Location Matter When Managing Wetlands for Watershed-Scale Flood and Drought Resilience?, 55, 529–542, https://doi.org/10.1111/1752-1688.12737, 2019.

Andino, D. F., Beven, K., Kauffeldt, A., Xu, C.-Y., Halldin, S., and Baldassarre, G. D.: Event and model dependent rainfall adjustments to improve discharge predictions, 0, null, https://doi.org/10.1080/02626667.2016.1183775, 2016.

Anning, D. W. and Flynn, M. E.: Dissolved-solids sources, loads, yields, and concentrations in streams of the conterminous United States, U.S. Geological Survey, Reston VA, 2014.

Arshad, M., Guillaume, J. H. A., and Ross, A.: Assessing the Feasibility of Managed Aquifer Recharge for Irrigation under Uncertainty, 6, 2748–2769, https://doi.org/10.3390/w6092748, 2014.

Arvidsson, A., Blomqvist, G., and Öberg, G.: Impact of climate change on use of anti-icing and deicing salt in Sweden, 3–10, 2012.

Ascott, M. J., Gooddy, D. C., Wang, L., Stuart, M. E., Lewis, M. A., Ward, R. S., and Binley, A. M.: Global patterns of nitrate storage in the vadose zone, 8, 1416, https://doi.org/10.1038/s41467-017-01321-w, 2017.

Bachand, P. A. and Horne, A. J.: Denitrification in constructed free-water surface wetlands: I. Very high nitrate removal rates in a macrocosm study, 14, 9–15, 1999.

Baker, B. H., Czarnecki, J. M. P., Omer, A. R., Aldridge, C. A., Kröger, R., and Prevost, J. D.: Nutrient and sediment runoff from agricultural landscapes with varying suites of conservation practices in the Mississippi Alluvial Valley, 73, 75–85, https://doi.org/10.2489/jswc.73.1.75, 2018.

Banin, A. and Fish, A.: Secondary desertification due to salinization of intensively irrigated lands: The Israeli experience, Environ Monit Assess, 37, 17–37, https://doi.org/10.1007/BF00546878, 1995.

Barlas, Y. and Carpenter, S.: Philosophical roots of model validation: Two paradigms, 6, 148–166, https://doi.org/10.1002/sdr.4260060203, 1990.

Bastviken, D., Thomsen, F., Svensson, T., Karlsson, S., Sandén, P., Shaw, G., Matucha, M., and Öberg, G.: Chloride retention in forest soil by microbial uptake and by natural chlorination of organic matter, Geochimica et Cosmochimica Acta, 71, 3182–3192, https://doi.org/10.1016/j.gca.2007.04.028, 2007.

Basu, N. B., Van Meter, K. J., Byrnes, D. K., Van Cappellen, P., Brouwer, R., Jacobsen, B. H., Jarsjö, J., Rudolph, D. L., Cunha, M. C., Nelson, N., Bhattacharya, R., Destouni, G., and Olsen, S. B.: Managing nitrogen legacies to accelerate water quality improvement, Nat. Geosci., 15, 97–105, https://doi.org/10.1038/s41561-021-00889-9, 2022.

Bavi, A., Kashuli, H. A., Boroomand, S., Naseri, A., and Albaji, M.: Evaporation Losses from Sprinkler Irrigation Systems under Various Operating Conditions, 9, 597–600, https://doi.org/10.3923/jas.2009.597.600, 2009.

Benettin, P., Bailey, S. W., Campbell, J. L., Green, M. B., Rinaldo, A., Likens, G. E., McGuire, K. J., and Botter, G.: Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH, USA, Water Resour. Res., 51, 9256–9272, https://doi.org/10.1002/2015WR017552, 2015.

Berghuijs, W. R. and Kirchner, J. W.: The relationship between contrasting ages of groundwater and streamflow, 44, 8925–8935, https://doi.org/10.1002/2017GL074962, 2017.

Beven, K.: On subsurface stormflow: Predictions with simple kinematic theory for saturated and unsaturated flows, Water Resour. Res., 18, 1627–1633, https://doi.org/10.1029/WR018i006p01627, 1982.

Beven, K.: Towards a coherent philosophy for modelling the environment, Proc. R. Soc. Lond. A, 458, 2465–2484, https://doi.org/10.1098/rspa.2002.0986, 2002.

Beven, K.: A manifesto for the equifinality thesis, 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, 6, 279–298, https://doi.org/10.1002/hyp.3360060305, 1992.

Bierkens, M. F. P. and Wada, Y.: Non-renewable groundwater use and groundwater depletion: a review, 14, 063002, https://doi.org/10.1088/1748-9326/ab1a5f, 2019.

Bierwagen, B. G., Theobald, D. M., Pyke, C. R., Choate, A., Groth, P., Thomas, J. V., and Morefield, P.: National housing and impervious surface scenarios for integrated climate impact assessments, PNAS, 107, 20887–20892, https://doi.org/10.1073/pnas.1002096107, 2010.

Blasius, B. J. and Merritt, R. W.: Field and laboratory investigations on the effects of road salt (NaCl) on stream macroinvertebrate communities, Environ. Pollut., 120, 219–231, 2002.

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., Amorim, P. B. de, Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chifflard, P., Claps, P., Clark, M., Collins, A., Croke, B., Dathe, A., David, P. C., Barros, F. P. J. de, Rooij, G. de, Baldassarre, G. D., Driscoll, J. M., Düthmann, D., Dwivedi, R., Eris, E., Farmer, W. H., Feiccabrino, J., Ferguson, G., Ferrari, E., Ferraris, S., Fersch, B., Finger, D., Foglia, L., Fowler, K., Gartsman, B., Gascoin, S., Gaume, E., Gelfan, A., Geris, J., Gharari, S., Gleeson, T., Glendell, M., Bevacqua, A. G., González-Dugo, M. P., Grimaldi, S., Gupta, A. B., Guse, B., Han, D., Hannah, D., Harpold, A., Haun, S., Heal, K., Helfricht, K., Herrnegger, M., Hipsey, M., Hlaváčiková, H., Hohmann, C., Holko, L., Hopkinson, C., Hrachowitz, M., Illangasekare, T. H., Inam, A., Innocente, C., Istanbulluoglu, E., Jarihani, B., et al.: Twenty-three Unsolved Problems in Hydrology (UPH) – a community perspective, 0, null, https://doi.org/10.1080/02626667.2019.1620507, 2019.

Bogacki, P. and Shampine, L. F.: A 3(2) pair of Runge - Kutta formulas, Applied Mathematics Letters, 2, 321–325, https://doi.org/10.1016/0893-9659(89)90079-7, 1989.

Boggs, K. G., Kirk, R. W. V., Johnson, G. S., Fairley, J. P., and Porter, P. S.: Analytical Solutions to the Linearized Boussinesq Equation for Assessing the Effects of Recharge on Aquifer Discharge1, 46, 1116–1132, https://doi.org/10.1111/j.1752-1688.2010.00479.x, 2010.

Booth, C., Rowlinson, M., Clark, P., Delahaye, A., and Procter, S.: Scenarios and counterfactuals as modal narratives, Futures, 41, 87–95, https://doi.org/10.1016/j.futures.2008.07.037, 2009.

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., and Finnveden, G.: Scenario types and techniques: Towards a user's guide, Futures, 38, 723–739, https://doi.org/10.1016/j.futures.2005.12.002, 2006.

Borsuk, M., Mavrommati, G., Samal, N., Zuidema, S., Wollheim, W., Rogers, S., Thorn, A., Lutz, D., Mineau, M., Grimm, C., Wake, C., Howarth, R., and Gardner, K.: Deliberative

multiattribute valuation of ecosystem services across a range of regional land-use, socioeconomic, and climate scenarios for the upper Merrimack River watershed, New Hampshire, USA, 24, https://doi.org/10.5751/ES-10806-240211, 2019.

Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, 359, https://doi.org/10.1126/science.aam7240, 2018.

Burakowski, E. A., Wake, C. P., Braswell, B., and Brown, D. P.: Trends in wintertime climate in the northeastern United States: 1965–2005, 113, D20114, 2008.

Burt, C. M., A. J. Clemmens, T. S. Strelkoff, K. H. Solomon, R. D. Bliesner, L. A. Hardy, T. A. Howell, and D. E. Eisenhauer: Irrigation Performance Measures: Efficiency and Uniformity, 123, 423–442, https://doi.org/10.1061/(ASCE)0733-9437(1997)123:6(423), 1997.

Campbell, J. L., Ollinger, S. V., Flerchinger, G. N., Wicklein, H., Hayhoe, K., and Bailey, A. S.: Past and projected future changes in snowpack and soil frost at the Hubbard Brook Experimental Forest, New Hampshire, USA, Hydrol. Process., 24, 2465–2480, https://doi.org/10.1002/hyp.7666, 2010.

Campolongo, F., Cariboni, J., and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environmental Modelling & Software, 22, 1509–1518, https://doi.org/10.1016/j.envsoft.2006.10.004, 2007.

Cañedo-Argüelles, M., Kefford, B. J., Piscart, C., Prat, N., Schäfer, R. B., and Schulz, C.-J.: Salinisation of rivers: An urgent ecological issue, 173, 157–167, https://doi.org/10.1016/j.envpol.2012.10.011, 2013.

Cañedo-Argüelles, M., Hawkins, C. P., Kefford, B. J., Schäfer, R. B., Dyack, B. J., Brucet, S., Buchwalter, D., Dunlop, J., Frör, O., Lazorchak, J., Coring, E., Fernandez, H. R., Goodfellow, W., Achem, A. L. G., Hatfield-Dodds, S., Karimov, B. K., Mensah, P., Olson, J. R., Piscart, C., Prat, N., Ponsá, S., Schulz, C.-J., and Timpano, A. J.: Saving freshwater from salts, 351, 914–916, https://doi.org/10.1126/science.aad3488, 2016.

Cappiella, K. and Brown, K.: Impervious Cover and Land Use in the Chesapeake Bay Watershed, Center for Watershed Protection, Ellicott City, MD, 2001.

Carleton, J. N. and Montas, H. J.: An analysis of performance models for free water surface wetlands, 44, 3595–3606, https://doi.org/10.1016/j.watres.2010.04.008, 2010.

Carr, G., Nortcliff S., and Potter R. B.: Water reuse for irrigated agriculture in Jordan: challenges of soil sustainability and the role of management strategies, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368, 5315–5321, https://doi.org/10.1098/rsta.2010.0181, 2010.

Casey, R. E. and Klaine, S. J.: Nutrient Attenuation by a Riparian Wetland during Natural and Artificial Runoff Events, J. Environ. Qual., 30, 1720–1731, https://doi.org/10.2134/jeq2001.3051720x, 2001.

Chang, H., Thiers, P., Netusil, N. R., Yeakley, J. A., Rollwagen-Bollens, G., Bollens, S. M., and Singh, S.: Relationships between environmental governance and water quality in a growing metropolitan area of the Pacific Northwest, USA, 18, 1383–1395, https://doi.org/10.5194/hess-18-1383-2014, 2014.

Cheng, F. Y. and Basu, N. B.: Biogeochemical hotspots: Role of small water bodies in landscape nutrient processing, 53, 5038–5056, https://doi.org/10.1002/2016WR020102, 2017.

Cheng, F. Y., Van Meter, K. J., Byrnes, D. K., and Basu, N. B.: Maximizing US nitrate removal through wetland protection and restoration, Nature, 588, 625–630, https://doi.org/10.1038/s41586-020-03042-5, 2020.

Cho, E., Jacobs, J. M., Jia, X., and Kraatz, S.: Identifying Subsurface Drainage using Satellite Big Data and Machine Learning via Google Earth Engine, 55, 8028–8045, https://doi.org/10.1029/2019WR024892, 2019.

Christianson, L., Tyndall, J., and Helmers, M.: Financial comparison of seven nitrate reduction strategies for Midwestern agricultural drainage, Water Resources and Economics, 2–3, 30–56, https://doi.org/10.1016/j.wre.2013.09.001, 2013.

Christianson, R., Christianson, L., Wong, C., Helmers, M., McIsaac, G., Mulla, D., and McDonald, M.: Beyond the nutrient strategies: Common ground to accelerate agricultural water quality improvement in the upper Midwest, Journal of Environmental Management, 206, 1072–1080, https://doi.org/10.1016/j.jenvman.2017.11.051, 2018.

CIESIN, Center For International Earth Science Information Network, and Socioeconomic Data and Applications Center, Columbia University: Gridded Population of the World, Version 4 (GPWv4): Population Density, Socioeconomic Data and Applications Center, Columbia University, Palisades, NY, 2016.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, 44, W00B02, 2008.

Coles, J. F., Cuffney, T. F., McMahon, G., and Beaulieu, K. M.: The Effects of Urbanization on the Biological, Physical, and Chemical Characteristics of Coastal New England Streams, U.S. Department of Interior, U.S. Geological Survey, 47 pp., 2004.

Conservation Reserve Program: Farmable Wetlands Program, Constructed Wetlands, USDA Farm Service Agency, 2015.

Contor, B. A. and Taylor, R. G.: Why improving irrigation efficiency increases total volume of consumptive use: irrigation efficiency increases consumptive use, 62, 273–280, https://doi.org/10.1002/ird.1717, 2013.

Contosta, A. R., Adolph, A., Burchsted, D., Burakowski, E., Green, M., Guerra, D., Albert, M., Dibb, J., Martin, M., McDowell, W. H., Routhier, M., Wake, C., Whitaker, R., and Wollheim, W.: A longer vernal window: the role of winter coldness and snowpack in driving spring transitions and lags, Glob Change Biol, 23, 1610–1625, https://doi.org/10.1111/gcb.13517, 2017.

Cooper, C. A., Mayer, P. M., and Faulkner, B. R.: Effects of road salts on groundwater and surface water dynamics of sodium and chloride in an urban restored stream, Biogeochemistry, 121, 149–166, https://doi.org/10.1007/s10533-014-9968-z, 2014.

Corsi, S. R., Graczyk, D. J., Geis, S. W., Booth, N. L., and Richards, K. D.: A Fresh Look at Road Salt: Aquatic Toxicity and Water-Quality Impacts on Local, Regional, and National Scales, 44, 7376–7382, https://doi.org/10.1021/es101333u, 2010.

Corsi, S. R., De Cicco, L. A., Lutz, M. A., and Hirsch, R. M.: River chloride trends in snow-affected urban watersheds: increasing concentrations outpace urban growth rate and are common among all seasons, Science of The Total Environment, 508, 488–497, https://doi.org/10.1016/j.scitotenv.2014.12.012, 2015.

Covington, H. R. and Weaver, J. N.: Geologic map and profiles of the north wall of the Snake River Canyon, Thousand Springs and Niagara quadrangles, Idaho, U.S. Geological Survey, Reston, VA, 1991.

Cronan, D., Trammel, J., Center for Resilient Communities (University of Idaho), and Water Systems Analysis Group (University of New Hampshire): Idaho Agricultural Futures, Center for Resilient Communities, University of Idaho, Moscow, ID, 2021.

Crumpton, W. G.: Using wetlands for water quality improvement in agricultural watersheds; the importance of a watershed scale approach, 44, 559–564, https://doi.org/10.2166/wst.2001.0880, 2001.

Crumpton, W. G., Stenback, G. A., Miller, B. A., and Helmers, M.: Potential Benefits of Wetland Filters for Tile Drainage Systems: Impact on Nitrate Loads to Mississippi River Subbasins, 36, 2006.

Czuba, J. A., Hansen, A. T., Foufoula-Georgiou, E., and Finlay, J. C.: Contextualizing Wetlands Within a River Network to Assess Nitrate Removal and Inform Watershed Management, 54, 1312–1337, https://doi.org/10.1002/2017WR021859, 2018.

Dale, V.: Hypoxia in the Northern Gulf of Mexico., Springer New York, 2010.

Daley, M. L., Potter, J. D., and McDowell, W. H.: Salinization of urbanizing New Hampshire streams and groundwater: effects of road salt and hydrologic variability, Journal of the North American Benthological Society, 28, 929–940, https://doi.org/10.1899/09-052.1, 2009.

De Jouvenel, H.: A Brief Methodological Guide to Scenario Building, Technological Forecasting and Social Change, 65, 37–48, https://doi.org/10.1016/S0040-1625(99)00123-7, 2000.

Demars, B. O. l., Russell Manson, J., Ólafsson, J. S., Gíslason, G. M., Gudmundsdóttir, R., Woodward, G., Reiss, J., Pichler, D. E., Rasmussen, J. J., and Friberg, N.: Temperature and the metabolic balance of streams, 56, 1106–1121, https://doi.org/10.1111/j.1365-2427.2010.02554.x, 2011.

Dewandel, B., Gandolfi, J.-M., Condappa, D. de, and Ahmed, S.: An efficient methodology for estimating irrigation return flow coefficients of irrigated crops at watershed and seasonal scale, 22, 1700–1712, https://doi.org/10.1002/hyp.6738, 2008.

Di Baldassarre, G., Kooy, M., Kemerink, J. S., and Brandimarte, L.: Towards understanding the dynamic behaviour of floodplains as human-water systems, 17, 3235–3244, https://doi.org/10.5194/hess-17-3235-2013, 2013.

Di Baldassarre, G., Viglione, A., Carr, G., Kuil, L., Yan, K., Brandimarte, L., and Blöschl, G.: Debates—Perspectives on socio-hydrology: Capturing feedbacks between physical and social processes, 51, 4770–4781, https://doi.org/10.1002/2014WR016416, 2015.

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., Barber, N. L., and Linsey, K. S.: Estimated use of water in the United States in 2015, U.S. Geological Survey, Reston, VA, 2018.

Dillon, P.: Future management of aquifer recharge, Hydrogeol J, 13, 313–316, https://doi.org/10.1007/s10040-004-0413-6, 2005.

Dillon, P., Stuyfzand, P., Grischek, T., Lluria, M., Pyne, R. D. G., Jain, R. C., Bear, J., Schwarz, J., Wang, W., Fernandez, E., Stefan, C., Pettenati, M., van der Gun, J., Sprenger, C., Massmann, G., Scanlon, B. R., Xanke, J., Jokela, P., Zheng, Y., Rossetto, R., Shamrukh, M., Pavelic, P., Murray, E., Ross, A., Bonilla Valverde, J. P., Palma Nava, A., Ansems, N., Posavec, K., Ha, K., Martin, R., and Sapiano, M.: Sixty years of global progress in managed aquifer recharge, 27, 1–30, https://doi.org/10.1007/s10040-018-1841-z, 2019.

Dillon, P., Fernández Escalante, E., Megdal, S. B., and Massmann, G.: Managed Aquifer Recharge for Water Resilience, 12, 1846, https://doi.org/10.3390/w12071846, 2020.

Donner, S. D. and Kucharik, C. J.: Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River, PNAS, 105, 4513–4518, https://doi.org/10.1073/pnas.0708300105, 2008.

Dowling, D.: Experimenting on Theories, 12, 261–273, https://doi.org/10.1017/S0269889700003410, 1999.

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., Stouffer, R. J., Cooke, W., Dunne, K. A., Harrison, M. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Phillipps, P. J., Sentman, L. T., Samuels, B. L., Spelman, M. J., Winton, M., Wittenberg, A. T., and Zadeh, N.: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics, 25, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1, 2012.

Dupas, R., Ehrhardt, S., Musolff, A., Fovet, O., and Durand, P.: Long-term nitrogen retention and transit time distribution in agricultural catchments in western France, Environ. Res. Lett., 15, 115011, https://doi.org/10.1088/1748-9326/abbe47, 2020.

Eaton, J. G. and Scheller, R. M.: Effects of climate warming on fish thermal habitat in streams of the United States, Limnol. Oceanogr., 41, 1109–1115, https://doi.org/10.4319/lo.1996.41.5.1109, 1996.

Ejsmond, M. J., Czarnołęski, M., Kapustka, F., and Kozłowski, J.: How to Time Growth and Reproduction during the Vegetative Season: An Evolutionary Choice for Indeterminate Growers in Seasonal Environments, 175, 551–563, https://doi.org/10.1086/651589, 2010.

Eker, S.: Practice and perspectives in the validation of resource management models, 10, 2018.

Ellsworth, J.: Saline and Sodic Soils in Idaho, in: Proceedings of the Idaho Nutrient Management Conference, Idaho Nutrient Management Conference, Twin Falls, ID, 2004.

van Emmerik, T. H. M., Li, Z., Sivapalan, M., Pande, S., Kandasamy, J., Savenije, H. H. G., Chanan, A., and Vigneswaran, S.: Socio-hydrologic modeling to understand and mediate the competition for water between agriculture development and environmental health: Murrumbidgee River basin, Australia, 18, 4239–4259, https://doi.org/10.5194/hess-18-4239-2014, 2014.

Environment Canada: The Environmental Management of Road Salts, 2004.

Evenson, G. R., Golden, H. E., Christensen, J. R., Lane, C. R., Rajib, A., D'Amico, E., Mahoney, D. T., White, E., and Wu, Q.: Wetland restoration yields dynamic nitrate responses across the Upper Mississippi river basin, Environ. Res. Commun., 3, 095002, https://doi.org/10.1088/2515-7620/ac2125, 2021.

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, 30, 3061–3078, https://doi.org/10.1029/94WR01498, 1994.

Fekete, B. M., Vörösmarty, C. J., and Lammers, R. B.: Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error, 37, 1955–1967, 2001.

Fennel, K. and Testa, J. M.: Biogeochemical Controls on Coastal Hypoxia, Annu. Rev. Mar. Sci., 11, 105–130, https://doi.org/10.1146/annurev-marine-010318-095138, 2019.

Ferraro, P. J., Sanchirico, J. N., and Smith, M. D.: Causal inference in coupled human and natural systems, PNAS, 116, 5311–5318, https://doi.org/10.1073/pnas.1805563115, 2019.

Ferrington JR., L. C.: Endangered rivers: A case history of the Arkansas River in the Central Plains, 3, 305–316, https://doi.org/10.1002/aqc.3270030405, 1993.

Findlay, S. E. G. and Kelly, V. R.: Emerging indirect and long-term road salt effects on ecosystems, 1223, 58–68, https://doi.org/10.1111/j.1749-6632.2010.05942.x, 2011a.

Findlay, S. E. G. and Kelly, V. R.: Emerging indirect and long-term road salt effects on ecosystems: Findlay & Kelly, 1223, 58–68, https://doi.org/10.1111/j.1749-6632.2010.05942.x, 2011b.

Fisher, J. and Acreman, M. C.: Wetland nutrient removal: a review of the evidence, 8, 673–685, https://doi.org/10.5194/hess-8-673-2004, 2004.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., and Zaks, D. P. M.: Solutions for a cultivated planet, 478, 337–342, https://doi.org/10.1038/nature10452, 2011.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: emcee: The MCMC Hammer, 125, 306–312, https://doi.org/10.1086/670067, 2013.

Foster, S. and van Steenbergen, F.: Conjunctive groundwater use: a "lost opportunity" for water management in the developing world?, 19, 959–962, http://dx.doi.org.unh.idm.oclc.org/10.1007/s10040-011-0734-1, 2011.

Frans, L., Rupert, M., Hunt, C., Jr, and Skinner, K.: Groundwater quality in the Columbia Plateau, Snake River Plain, and Oahu basaltic-rock and basin-fill aquifers in the Northwestern United States and Hawaii, 1992–2010, U.S. Geological Survey, Reston VA, 2012.

Frederiksen, H. D. and Allen, R. G.: A common basis for analysis, evaluation and comparison of offstream water uses, 36, 266–282, https://doi.org/10.1080/02508060.2011.580449, 2011.

Fu, L. and Usman, T.: Safety Impacts of Using Deicing Salt, 2014.

Garabedian, S. P.: Hydrology and digital simulation of the regional aquifer system, eastern Snake River Plain, Idaho, U.S. Government Printing Office, 1992.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Ghassemi, F., Jakeman, A. J., Nix, H. A., Australian National University., and Centre for Resource and Environmental Studies.: Salinisation of land and water resources: human causes, extent, management, and case studies, NSW University Press, Sydney, N.S.W., 1995.

Gilmore, S.: Assessing the Adaptive Capacity of Idaho's Magic Valley As a Complex Social-Ecological System, MS, University of Idaho, Moscow, ID, 115 pp., 2019.

Giuliani, M., Li, Y., Castelletti, A., and Gandolfi, C.: A coupled human-natural systems analysis of irrigated agriculture under changing climate: CHNS ANALYSIS OF IRRIGATED AGRICULTURE UNDER CLIMATE CHANGE, 52, 6928–6947, https://doi.org/10.1002/2016WR019363, 2016.

Gleick, P. H., Christian-Smith, J., and Cooley, H.: Water-use efficiency and productivity: rethinking the basin approach, 36, 784–798, https://doi.org/10.1080/02508060.2011.631873, 2011.

Godwin, K. S., Hafner, S. D., and Buff, M. F.: Long-term trends in sodium and chloride in the Mohawk River, New York: the effect of fifty years of road-salt application, 124, 273–281, https://doi.org/10.1016/S0269-7491(02)00481-5, 2003.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, 5, 65–80, 2010.

Goolsby, D. A., Battaglin, W. A., Aulenbach, B. T., and Hooper, R. P.: Nitrogen flux and sources in the Mississippi River Basin, Science of The Total Environment, 248, 75–86, https://doi.org/10.1016/S0048-9697(99)00532-X, 2000.

Goolsby, D. A., Battaglin, W. A., Aulenbach, B. T., and Hooper, R. P.: Nitrogen Input to the Gulf of Mexico, 30, 329–336, https://doi.org/10.2134/jeq2001.302329x, 2001.

Gourevitch, J. D., Singh, N. K., Minot, J., Raub, K. B., Rizzo, D. M., Wemple, B. C., and Ricketts, T. H.: Spatial targeting of floodplain restoration to equitably mitigate flood risk, Global Environmental Change, 61, 102050, https://doi.org/10.1016/j.gloenvcha.2020.102050, 2020.

Grafton, R. Q., Williams, J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., and Allen, R. G.: The paradox of irrigation efficiency, 361, 748–750, 2018.

Granato, G. E., DeSimone, L. A., Barbaro, J. R., and Jeznach, L. C.: Methods for evaluating potential sources of chloride in surface waters and groundwaters of the conterminous United States, U.S. Geological Survey, Reston, VA, 2015.

Green, C. T., Puckett, L. J., Böhlke, J. K., Bekins, B. A., Phillips, S. P., Kauffman, L. J., Denver, J. M., and Johnson, H. M.: Limited occurrence of denitrification in four shallow aquifers in agricultural areas of the United States, Journal of Environmental Quality, 37, 9941009, https://doi.org/10.2134/jeq2006.0419, 2008.

Gregersen, H. M., Ffolliott, P. F., and Brooks, K. N.: Integrated watershed management: connecting people to their land and water, CABI, Wallingford, Oxforshire, UK ; Cambridge, MA, 201 pp., 2007.

Groffman, P. M., Gold, A. J., and Addy, K.: Nitrous oxide production in riparian zones and its importance to national emission inventories, Chemosphere - Global Change Science, 2, 291–299, https://doi.org/10.1016/S1465-9972(00)00018-0, 2000.

Grogan, D. S.: Global and regional assessments of unsustainable groundwater use in irrigated agriculture, PhD, University of New Hampshire, Durham, NH, 222 pp., 2016.

Grogan, D. S., Wisser, D., Prusevich, A., Lammers, R. B., and Frolking, S.: The use and re-use of unsustainable groundwater for irrigation: a global budget, Environ. Res. Lett., 12, 034017, https://doi.org/10.1088/1748-9326/aa5fb2, 2017.

Grogan, D. S., Zuidema, S., Prusevich, A., Wollheim, W. M., Glidden, S., and Lammers, R. B.: WBM: A scalable gridded global hydrologic model with water tracking functionality, https://doi.org/10.5194/gmd-2022-59, 2022.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Guyennon, N., Salerno, F., Portoghese, I., and Romano, E.: Climate Change Adaptation in a Mediterranean Semi-Arid Catchment: Testing Managed Aquifer Recharge and Increased Surface Reservoir Capacity, 9, 689, https://doi.org/10.3390/w9090689, 2017.

Haggerty, R. and Gorelick, S. M.: Multiple-Rate Mass Transfer for Modeling Diffusion and Surface Reactions in Media with Pore-Scale Heterogeneity, 31, 2383–2400, https://doi.org/10.1029/95WR10583, 1995.

Haie, N. and Keller, A. A.: Effective Efficiency as a Tool for Sustainable Water Resources Management [1], 44, 961–968, https://doi.org/10.1111/j.1752-1688.2008.00194.x, 2008.

Hale, I. L., Wollheim, W. M., Smith, R. G., Asbjornsen, H., Brito, A. F., Broders, K., Grandy, A. S., and Rowe, R.: A Scale-Explicit Framework for Conceptualizing the Environmental Impacts of Agricultural Land Use Changes, 6, 8432–8451, https://doi.org/10.3390/su6128432, 2014.

Hamon, W. R.: Computation of direct runoff amounts from storm rainfall, in: International Association of Hydrological Sciences Publications, vol. 63, International Association of Hydrological Sciences, Oxford, MS, 52–62, 1963.

Han, W., Yang, Z., Di, L., and Mueller, R.: CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support, Computers and Electronics in Agriculture, 84, 111–123, https://doi.org/10.1016/j.compag.2012.03.005, 2012.

Hansen, A. T., Dolph, C. L., Foufoula-Georgiou, E., and Finlay, J. C.: Contribution of wetlands to nitrate removal at the watershed scale, Nature Geosci, 11, 127–132, https://doi.org/10.1038/s41561-017-0056-6, 2018.

Hansen, A. T., Campbell, T., Cho, S. J., Czuba, J. A., Dalzell, B. J., Dolph, C. L., Hawthorne, P. L., Rabotyagov, S., Lang, Z., Kumarasamy, K., Belmont, P., Finlay, J. C., Foufoula-Georgiou, E., Gran, K. B., Kling, C. L., and Wilcock, P.: Integrated assessment modeling reveals near-channel management as cost-effective to improve water quality in agricultural watersheds, Proc Natl Acad Sci USA, 118, e2024912118, https://doi.org/10.1073/pnas.2024912118, 2021.

Harbaugh, A.: MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model—the Ground-Water Flow Process, in: Book 6. Modeling techniques, Section A. Ground Water, vol. 6, U.S. Geological Survey, Reston VA, 253, 2005.

Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., Wood, E., Anderson, B., Bradbury, J., DeGaetano, A., Troy, T. J., and Wolfe, D.: Past and future changes in climate and hydrological indicators in the US Northeast, Clim Dyn, 28, 381–407, https://doi.org/10.1007/s00382-006-0187-8, 2006.

Hébert, M.-P., Symons, C. C., Cañedo-Argüelles, M., Arnott, S. E., Derry, A. M., Fugère, V., Hintz, W. D., Melles, S. J., Astorg, L., Baker, H. K., Brentrup, J. A., Downing, A. L., Ersoy, Z., Espinosa, C., Franceschini, J. M., Giorgio, A. T., Göbeler, N., Gray, D. K., Greco, D., Hassal, E., Huynh, M., Hylander, S., Jonasen, K. L., Kirkwood, A., Langenheder, S., Langvall, O., Laudon, H., Lind, L., Lundgren, M., McClymont, A., Proia, L., Relyea, R. A., Rusak, J. A., Schuler, M. S., Searle, C. L., Shurin, J. B., Steiner, C. F., Striebel, M., Thibodeau, S., Urrutia Cordero, P., Vendrell-Puigmitja, L., Weyhenmeyer, G. A., and Beisner, B. E.: Lake salinization drives consistent losses of zooplankton abundance and diversity across coordinated mesocosm experiments, n/a, https://doi.org/10.1002/lol2.10239, 2022.

Higginson: Moratorium order, Idaho Department of Water Resources, Idaho Department of Water Resources, 2, 1992.

Horvath, E. K., Christensen, J. R., Mehaffey, M. H., and Neale, A. C.: Building a potential wetland restoration indicator for the contiguous United States, Ecol Indic, 83, 462–473, https://doi.org/10.1016/j.ecolind.2017.07.026, 2017.

Hrachowitz, M., Soulsby, C., Tetzlaff, D., Malcolm, I. A., and Schoups, G.: Gamma distribution models for transit time estimation in catchments: Physical interpretation of parameters and implications for time-variant transit time assessment, Water Resour. Res., 46, W10536, https://doi.org/10.1029/2010WR009148, 2010.

IDWR: Enhanced Snake Plain Aquifer Model Version 2.1: Final Report, Idaho Department of Water Resources, Boise ID, 2013.

Administrative Basins: https://data-idwr.opendata.arcgis.com/datasets/fb0df7d688a04074bad92ca8ef74cc26_4, last access: 1 June 2018.

IDWR: Settlement Agreement Entered into June 30, 2015 Between Participating Members of the Surface Water Coalition and Participating Members of the Idaho Ground Water Appropriators, Inc., 2015.

Inserillo, E. A., Green, M. B., Shanley, J. B., and Boyer, J. N.: Comparing catchment hydrologic response to a regional storm using specific conductivity sensors, Hydrol. Process., 31, 1074–1085, https://doi.org/10.1002/hyp.11091, 2017.

Iovanna, R., Hyberg, S., and Crumpton, W.: Treatment wetlands: Cost-effective practice for intercepting nitrate before it reaches and adversely impacts surface waters, 63, 14A-15A, https://doi.org/10.2489/jswc.63.1.14A, 2008.

IPCC (Ed.): Emissions scenarios: summary for policymakers;a special report of IPCC Working Group III$Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, 20 pp., 2000.

Ippolito, A., Carolli, M., Varolo, E., Villa, S., and Vighi, M.: Evaluating pesticide effects on freshwater invertebrate communities in alpine environment: a model ecosystem experiment, 21, 2051–2067, https://doi.org/10.1007/s10646-012-0957-5, 2012.

IWRB: Resolution Adopting Swan Falls Amendments to Policy 32 of State Water Plan -, , Policy 32 (Amendment), 9, 1985.

IWRB: Eastern Snake Plain Aquifer (ESPA) Comprehensive Aquifer Management Plan, Idaho Water Resources Board, Boise, ID, 2009.

IWRB: Final Report: Eastern Snake Plain Aquifer (ESPA) Review of Comprehensive Managed Aquifer Recharge Program, Idaho Water Resources Board, Boise, ID, 2016.

Jackson, R. B. and Jobbagy, E. G.: From icy roads to salty streams, 102, 14487–14488, 2005.

Jägermeyr, J., Gerten, D., Heinke, J., Schaphoff, S., Kummu, M., and Lucht, W.: Water savings potentials of irrigation systems: global simulation of processes and linkages, Hydrol. Earth Syst. Sci., 19, 3073–3091, https://doi.org/10.5194/hess-19-3073-2015, 2015.

Jägermeyr, J., Gerten, D., Schaphoff, S., Heinke, J., Lucht, W., and Rockström, J.: Integrated crop water management might sustainably halve the global food gap, Environ. Res. Lett., 11, 025002, https://doi.org/10.1088/1748-9326/11/2/025002, 2016.

Jaynes, D. B. and Isenhart, T. M.: Reconnecting Tile Drainage to Riparian Buffer Hydrology for Enhanced Nitrate Removal, J. Environ. Qual., 43, 631–638, https://doi.org/10.2134/jeq2013.08.0331, 2014.

Jaynes, D. B., Kaspar, T. C., Moorman, T. B., and Parkin, T. B.: In Situ Bioreactors and Deep Drain-Pipe Installation to Reduce Nitrate Losses in Artificially Drained Fields, 37, 429–436, https://doi.org/10.2134/jeq2007.0279, 2008.

Johnston, C. A.: Sediment and nutrient retention by freshwater wetlands: Effects on surface water quality, 21, 491–565, https://doi.org/10.1080/10643389109388425, 1991.

Kadlec, R. H. and Knight, R. L.: Treatment Wetlands, CRC Press, Boca Raton FL, 1996.

Kadlec, R. H. and Reddy, K. R.: Temperature Effects in Treatment Wetlands, 73, 543–557, 2001.

Kaushal, S. S., Groffman, P. M., Likens, G. E., Belt, K. T., Stack, W. P., Kelly, V. R., Band, L. E., and Fisher, G. T.: Increased salinization of fresh water in the northeastern United States, PNAS, 102, 13517–13520, https://doi.org/10.1073/pnas.0506414102, 2005.

Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, Water Resour. Res., 42, W03407, https://doi.org/10.1029/2005WR004368, 2006a.

Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, Water Resour. Res., 42, W03408, https://doi.org/10.1029/2005WR004376, 2006b.

Keller, A. and Keller, J.: Effective Efficiency: A Water Use Efficiency Concept for Allocating Freshwater Resources, in: Discussion Paper 22, Arlington, VA, 20, 1995.

Keller, A. A., Keller, J., Seckler, D. W., Garces-Restrepo, C., and International Water Management Institute: Integrated water resource systems: theory and policy implications, International Water Management Institute, Colombo, Sri Lanka, 1996.

Kelly, V. R., Lovett, G. M., Weathers, K. C., Findlay, S. E. G., Strayer, D. L., Burns, D. J., and Likens, G. E.: Long-Term Sodium Chloride Retention in a Rural Watershed: Legacy Effects of Road Salt on Streamwater Concentration, 42, 410–415, https://doi.org/10.1021/es071391l, 2008.

Kelly, W. R., Panno, S. V., and Hackley, K. C.: Impacts of Road Salt Runoff on Water Quality of the Chicago, Illinois, Region, Environmental & Engineering Geoscience, 18, 65–81, https://doi.org/10.2113/gseegeosci.18.1.65, 2012.

Kelting, D. L., Laxson, C. L., and Yerger, E. C.: Regional analysis of the effect of paved roads on sodium and chloride in lakes, 46, 2749–2758, https://doi.org/10.1016/j.watres.2012.02.032, 2012.

Kincaid, D. W. and Findlay, S. E. G.: Sources of Elevated Chloride in Local Streams: Groundwater and Soils as Potential Reservoirs, Water Air Soil Pollut, 203, 335–342, https://doi.org/10.1007/s11270-009-0016-x, 2009.

Kirchner, J. W., Feng, X., and Neal, C.: Fractal stream chemistry and its implications for contaminant transport in catchments, Nature, 403, 524–527, https://doi.org/10.1038/35000537, 2000.

Kirchner, J. W., Tetzlaff, D., and Soulsby, C.: Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments, Hydrol. Process., 24, 1631–1645, https://doi.org/10.1002/hyp.7676, 2010.

Kjelstrom, L. C.: Streamflow gains and losses in the Snake River and ground-water budgets for the Snake River plain, Idaho and eastern Oregon, 1995.

Kleindorfer, G. B., O'Neill, L., and Ganeshan, R.: Validation in Simulation: Various Positions in the Philosophy of Science, 44, 1087–1099, https://doi.org/10.1287/mnsc.44.8.1087, 1998.

Konikow, L. F. and Bredehoeft, J. D.: Ground-water models cannot be validated, Advances in Water Resources, 15, 75–83, https://doi.org/10.1016/0309-1708(92)90033-X, 1992.

Konikow, L. F. and Leake, S. A.: Depletion and Capture: Revisiting "The Source of Water Derived from Wells," 52, 100–111, https://doi.org/10.1111/gwat.12204, 2014.

Kovacic, D. A., David, M. B., Gentry, L. E., Starks, K. M., and Cooke, R. A.: Effectiveness of Constructed Wetlands in Reducing Nitrogen and Phosphorus Export from Agricultural Tile Drainage, 29, 1262–1274, 2000.

Kraus, J. M., Schmidt, T. S., Walters, D. M., Wanty, R. B., Zuellig, R. E., and Wolf, R. E.: Cross-ecosystem impacts of stream pollution reduce resource and contaminant flux to riparian food webs, Ecological Applications, 24, 235–243, https://doi.org/10.1890/13-0252.1, 2013.

Kroes, D. E., Schenk, E. R., Noe, G. B., and Benthem, A. J.: Sediment and nutrient trapping as a result of a temporary Mississippi River floodplain restoration: The Morganza Spillway during the 2011 Mississippi River Flood, Ecological Engineering, 82, 91–102, https://doi.org/10.1016/j.ecoleng.2015.04.056, 2015.

Kucharik, C. J.: Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield, Earth Interact., 7, 1–33, https://doi.org/10.1175/1087-3562(2003)007<0001:EOAPAM>2.0.CO;2, 2003.

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young-Molling, C., Ramankutty, N., Norman, J. M., and Gower, S. T.: Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure, 14, 795–825, https://doi.org/10.1029/1999GB001138, 2000.

Lankford, B.: Fictions, fractions, factorials and fractures; on the framing of irrigation efficiency, 108, 27–38, https://doi.org/10.1016/j.agwat.2011.08.010, 2012.

Lehner, B., Verdin, K., and Jarvis, A.: New Global Hydrography Derived From Spaceborne Elevation Data, Eos Trans. AGU, 89, 93–94, https://doi.org/10.1029/2008EO100001, 2008.

Leopold, L. B.: Fluvial processes in geomorphology., San Francisco : W. H. Freeman, [1964], 1964.

Levenspiel, O.: Chemical reaction engineering, Wiley, New York, 1999.

Leytem, A. B., Williams, P., Zuidema, S., Martinez, A., Chong, Y. L., Vincent, A., Vincent, A., Cronan, D., Kliskey, A., Wulfhorst, J. D., Alessa, L., and Bjorneberg, D.: Cycling Phosphorus and Nitrogen through Cropping Systems in an Intensive Dairy Production Region, 11, 1005, https://doi.org/10.3390/agronomy11051005, 2021.

Lightbody, A. F., Avener, M. E., and Nepf, H. M.: Observations of short-circuiting flow paths within a free-surface wetland in Augusta, Georgia, USA, 53, 1040, 2008.

Lin Y. and Garcia L. A.: Assessing the Impact of Irrigation Return Flow on River Salinity for Colorado's Arkansas River Valley, Journal of Irrigation and Drainage Engineering, 138, 406–415, https://doi.org/10.1061/(ASCE)IR.1943-4774.0000410, 2012.

Lindholm, G. F.: Summary of the Snake River plain Regional Aquifer-System Analysis in Idaho and eastern Oregon, U.S. Government Printing Office, 1996.

Liu, J., Hertel, T. W., Bowling, L., Jame, S., Kucharik, C., and Ramankutty, N.: Evaluating Alternative Options for Managing Nitrogen Losses from Corn Production, 4, 5, 2018.

Lovin, H. T.: Dreamers, Schemers, and Doers of Idaho Irrigation, 76, 232–243, 2002.

Lutz, S. R., Trauth, N., Musolff, A., Breukelen, B. M. V., Knöller, K., and Fleckenstein, J. H.: How Important is Denitrification in Riparian Zones? Combining End-Member Mixing and Isotope Modeling to Quantify Nitrate Removal from Riparian Groundwater, 56, e2019WR025528, https://doi.org/10.1029/2019WR025528, 2020.

Maliva, R. G.: Economics of Managed Aquifer Recharge, 6, 1257–1279, https://doi.org/10.3390/w6051257, 2014.

Manzato, A.: A Note On the Maximum Peirce Skill Score, Wea. Forecasting, 22, 1148–1154, https://doi.org/10.1175/WAF1041.1, 2007.

Marshall, E.: Reducing Nutrient Losses From Cropland in the Mississippi/Atchafalaya River Basin: Cost Efficiency and Regional Distribution, US Department of Agiculture, Economic Research Service, Washington, D. C., 2018.

Masaki, Y., Hanasaki, N., Biemans, H., Schmied, H. M., Tang, Q., Wada, Y., Gosling, S. N., Takahashi, K., and Hijioka, Y.: Intercomparison of global river discharge simulations focusing on dam operation—multiple models analysis in two case-study river basins, Missouri–Mississippi and Green–Colorado, Environ. Res. Lett., 12, 055002, https://doi.org/10.1088/1748-9326/aa57a8, 2017.

Maupin, M. A., Kenny, J. F., Hutson, S. S., Lovelace, J. K., Barber, N. L., and Linsey, K. S.: Estimated use of water in the United States in 2010, U.S. Geological Survey, Reston, Virginia, 56 pp., 2014.

Mavrommati, G., Borsuk, M., and Howarth, R.: A novel deliberative multicriteria evaluation approach to ecosystem service valuation, 22, https://doi.org/10.5751/ES-09105-220239, 2017.

Mayer, P. M., Reynolds, S. K., McCutchen, M. D., and Canfield, T. J.: Meta-Analysis of Nitrogen Removal in Riparian Buffers, 36, 1172–1180, https://doi.org/10.2134/jeq2006.0462, 2007.

McDonnell, J. J.: A Rationale for Old Water Discharge Through Macropores in a Steep, Humid Catchment, Water Resour. Res., 26, 2821–2832, https://doi.org/10.1029/WR026i011p02821, 1990.

McDonnell, J. J.: Are all runoff processes the same?, Hydrol. Process., 27, 4103–4111, https://doi.org/10.1002/hyp.10076, 2013.

McDonnell, J. J., McGuire, K., Aggarwal, P., Beven, K. J., Biondi, D., Destouni, G., Dunn, S., James, A., Kirchner, J., Kraft, P., Lyon, S., Maloszewski, P., Newman, B., Pfister, L., Rinaldo, A., Rodhe, A., Sayama, T., Seibert, J., Solomon, K., Soulsby, C., Stewart, M., Tetzlaff, D., Tobin, C., Troch, P., Weiler, M., Western, A., Wörman, A., and Wrede, S.: How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis, 24, 1745–1754, https://doi.org/10.1002/hyp.7796, 2010.

McLean, R. K., Ranjan, R. S., and Klassen, G.: Spray evaporation losses from sprinkler irrigation systems, 42, 1–8, 2000.

McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality, Hydrological Processes, 4078, 2012.

McVay, M.: Incorporating Recharge Limitations into the Prioritization of Aquifer Recharge Sites Based on Hydrologic Benefits Using ESPAM2.1, IDWR, 2015.

Meter, K. J. V., Basu, N. B., Veenstra, J. J., and Burras, C. L.: The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes, Environ. Res. Lett., 11, 035014, https://doi.org/10.1088/1748-9326/11/3/035014, 2016.

Mississippi River/Gulf of Mexico Watershed Nutrient Task Force: Gulf Hypoxia Action Plan 2008 for Reducing, Mitigating, and Controlling Hypoxia in the Northern Gulf of Mexico and Improving Water Quality in the Mississippi River Basin, 2008.

Mitsch, W. J. and Day, J. W.: Restoration of wetlands in the Mississippi–Ohio–Missouri (MOM) River Basin: Experience and needed research, Ecological Engineering, 26, 55–69, https://doi.org/10.1016/j.ecoleng.2005.09.005, 2006.

Mitsch, W. J. and Gosselink, J. G.: Wetlands, 4th ed., Wiley, Hoboken, N.J, 582 pp., 2007.

Mitsch, W. J., Day, J. W., Gilliam, J. W., Groffman, P. M., Hey, D. L., Randall, G. W., and Wang, N.: Reducing Nitrogen Loading to the Gulf of Mexico from the Mississippi River

Basin: Strategies to Counter a Persistent Ecological ProblemEcotechnology—the use of natural ecosystems to solve environmental problems—should be a part of efforts to shrink the zone of hypoxia in the Gulf of Mexico, BioScience, 51, 373–388, https://doi.org/10.1641/0006-3568(2001)051[0373:RNLTTG]2.0.CO;2, 2001.

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S. J., Baldassarre, G. D., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D. A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z., and Belyaev, V.: "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, 58, 1256–1275, https://doi.org/10.1080/02626667.2013.809088, 2013.

Monteith, J. L.: Evaporation and environment: the state and movement of water in living organisms, 19, 205–224, 1965.

Moore, A., Zglobicki, S., and Olsen, N.: Manure Management in Potatoes: Salt Accumulations in Idaho Soils, in: Proceedings of the Idaho Potato Conference, Idaho Potato Conference, Twin Falls, ID, 2011.

Moreland, J. A.: Digital-model analysis of the effects of water-use alternatives on spring discharges Gooding and Jerome Counties, Idaho, Idaho Department of Water Resources, Boise, ID, 1976.

Morgan, R. P., Kline, K. M., Kline, M. J., Cushman, S. F., Sell, M. T., Weitzell, R. E., and Churchill, J. B.: Stream Conductivity: Relationships to Land Use, Chloride, and Fishes in Maryland Streams, 32, 941–952, https://doi.org/10.1080/02755947.2012.703159, 2012.

Morgenstern, U., Stewart, M. K., and Stenger, R.: Dating of streamwater using tritium in a post nuclear bomb pulse world: continuous variation of mean transit time with streamflow, Hydrol. Earth Syst. Sci., 14, 2289–2301, https://doi.org/10.5194/hess-14-2289-2010, 2010.

Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, 33, 161–174, https://doi.org/10.2307/1269043, 1991.

Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., Tank, J. L., Ashkenas, L. R., Cooper, L. W., Dahm, C. N., Dodds, W. K., Findlay, S. E. G., Gregory, S. V., Grimm, N. B., Johnson, S. L., McDowell, W. H., Meyer, J. L., Valett, H. M., Webster, J. R., Arango, C. P., Beaulieu, J. J., Bernot, M. J., Burgin, A. J., Crenshaw, C. L., Johnson, L. T., Niederlehner, B. R., O'Brien, J. M., Potter, J. D., Sheibley, R. W., Sobota, D. J., and Thomas, S. M.: Stream denitrification across biomes and its response to anthropogenic nitrate loading, 452, 202–205, https://doi.org/10.1038/nature06686, 2008.

Murray, D. M.: Economic analysis of the environmental impact of highway deicing salts, in: Transportation Research Record, 56th Annual Meeting of the Transportation Research Board, 1977.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I–A discussion of principles, 10, 282–290, 1970.

Natural Resources Conservation Service: Agricultural Conservation Easement Program (ACEP): Is ACEP Right for Me?, 2021.

Nazemi, A. and Wheater, H. S.: On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand, 19, 33–61, https://doi.org/10.5194/hess-19-33-2015, 2015.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R.: Soil and Water Assessment Tool Theoretical Documentation Version 2009, Texas A&M University System, College Station TX, 647 pp., 2011.

NH Fish and Game Department: NH WIldlife Action Plan 2015: Aquatic Habitat, Concord, NH, 2015.

2010 Census | State Data Center | NH Office of Energy and Planning: http://www.nh.gov/oep/data-center/census/, last access: 24 June 2015.

Nimiroski, M. T. and Waldron, M. C.: Sources of sodium and chloride in the Scituate Reservoir drainage basin, Rhode Island, US Department of the Interior, US Geological Survey, 2002.

Niswonger, R. G., Morway, E. D., Triana, E., and Huntington, J. L.: Managed aquifer recharge through off-season irrigation in agricultural regions, 53, 6970–6992, https://doi.org/10.1002/2017WR020458, 2017.

NOAA, NASA, and USAF: U.S. Standard Atmosphere, 1976, NOAA, Washington, D.C., 1976.

Novotny, E. V., Murphy, D., and Stefan, H. G.: Increase of urban lake salinity by road deicing salt, 406, 131–144, https://doi.org/10.1016/j.scitotenv.2008.07.037, 2008.

Novotny, E. V., Sander, A. R., Mohseni, O., and Stefan, H. G.: Chloride ion transport and mass balance in a metropolitan area using road salt, 45, https://doi.org/10.1029/2009WR008141, 2009.

Öberg, G. and Bastviken, D.: Transformation of Chloride to Organic Chlorine in Terrestrial Environments: Variability, Extent, and Implications, 42, 2526–2545, https://doi.org/10.1080/10643389.2011.592753, 2012.

Oreskes, N., Shrader-Frechette, K., and Belitz, K.: Verification, validation, and confirmation of numerical mode, 263, 641, 1994.

Ostendorf, D. W., Peeling, D. C., Mitchell, T. J., and Pollock, S. J.: Chloride persistence in a deiced access road drainage system, 30, 1756–1770, 2001.

Ostendorf, D. W., Hinlein, E. S., Ahlfeld, D. P., and DeJong, J. T.: Calibrated models of deicing agent solids, pavement texture, and specific conductivity of highway runoff, 132, 1562–1571, 2006.

Ostrom, E.: COPING WITH TRAGEDIES OF THE COMMONS, Annu. Rev. Polit. Sci., 2, 493–535, https://doi.org/10.1146/annurev.polisci.2.1.493, 1999.

Peck, S. L.: Simulation as experiment: a philosophical reassessment for biological modeling, Trends in Ecology & Evolution, 19, 530–534, https://doi.org/10.1016/j.tree.2004.07.019, 2004.

Penny, G., Mondal, M. S., Biswas, S., Bolster, D., Tank, J. L., and Müller, M. F.: Using Natural Experiments and Counterfactuals for Causal Assessment: River Salinity and the Ganges Water Agreement, 56, e2019WR026166, https://doi.org/10.1029/2019WR026166, 2020.

Perry, C.: Accounting for water use: Terminology and implications for saving water and increasing production, 98, 1840–1846, https://doi.org/10.1016/j.agwat.2010.10.002, 2011.

Perry, G. L. W., Wainwright, J., Etherington, T. R., and Wilmshurst, J. M.: Experimental Simulation: Using Generative Modeling and Palaeoecological Data to Understand Human-Environment Interactions, 4, 2016.

Pfeiffer, L. and Lin, C.-Y. C.: Does efficient irrigation technology lead to reduced groundwater extraction? Empirical evidence, Journal of Environmental Economics and Management, 67, 189–208, https://doi.org/10.1016/j.jeem.2013.12.002, 2014.

Pierce, D. W., Cayan, D. R., and Thrasher, B. L.: Statistical Downscaling Using Localized Constructed Analogs (LOCA), J. Hydrometeor., 15, 2558–2585, https://doi.org/10.1175/JHM-D-14-0082.1, 2014.

Plummer, L. n., Rupert, M. g., Busenberg, E., and Schlosser, P.: Age of Irrigation Water in Ground Water from the Eastern Snake River Plain Aquifer, South-Central Idaho, 38, 264–283, https://doi.org/10.1111/j.1745-6584.2000.tb00338.x, 2000.

Portmann, F. T., Siebert Stefan, and Döll Petra: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochemical Cycles, 24, https://doi.org/10.1029/2008GB003435, 2010.

Puckett, L. J., Tesoriero, A. J., and Dubrovsky, N. M.: Nitrogen Contamination of Surficial Aquifers—A Growing Legacy, Environ. Sci. Technol., 45, 839–844, https://doi.org/10.1021/es1038358, 2011.

Qadir, M.: Policy Note: Reversing Salt-Induced Land Degradation Requires Integrated Measures, 02, 1671001, https://doi.org/10.1142/S2382624X16710016, 2016.

Racchetti, E., Bartoli, M., Soana, E., Longhi, D., Christian, R. R., Pinardi, M., and Viaroli, P.: Influence of hydrological connectivity of riverine wetlands on nitrogen removal via denitrification, 103, 335–354, https://doi.org/10.1007/s10533-010-9477-7, 2010.

Rajib, A., Golden, H. E., Lane, C. R., and Wu, Q.: Surface Depression and Wetland Water Storage Improves Major River Basin Hydrologic Predictions, 56, e2019WR026561, https://doi.org/10.1029/2019WR026561, 2020.

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A.: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, 22, https://doi.org/10.1029/2007GB002952, 2008.

Redon, P.-O., Abdelouas, A., Bastviken, D., Cecchini, S., Nicolas, M., and Thiry, Y.: Chloride and Organic Chlorine in Forest Soils: Storage, Residence Times, And Influence of Ecological Conditions, 2011.

Redon, P.-O., Jolivet, C., Saby, N. P. A., Abdelouas, A., and Thiry, Y.: Occurrence of natural organic chlorine in soils for different land uses, Biogeochemistry, 114, 413–419, https://doi.org/10.1007/s10533-012-9771-7, 2012.

Reed, P. M., Hadjimichael, Antonia, Malek, Keyvan, Karimi, Tina, Vernon, Chris R., Srikrishnan, Vivek, Gupta, Rohini S., Gold, David F., Lee, Ben, Keller, Klaus, Thurber, Travis B., and Rice, Jennie S.: Addressing Uncertainty in Multisector Dynamics Research, Zenodo, https://doi.org/10.5281/ZENODO.6347960, 2022.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, 46, https://doi.org/10.1029/2009WR008328, 2010.

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., and Franks, S. W.: Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation, Water Resour. Res., 47, W11516, https://doi.org/10.1029/2011WR010643, 2011.

Ribaudo, M. O., Heimlich, R., Claassen, R., and Peters, M.: Least-cost management of nonpoint source pollution: source reduction versus interception strategies for controlling nitrogen loss in the Mississippi Basin, Ecological Economics, 37, 183–197, https://doi.org/10.1016/S0921-8009(00)00273-1, 2001.

Richburg, J. A., Patterson, W. A., and Lowenstein, F.: Effects of road salt and Phragmites australis invasion on the vegetation of a western Massachusetts calcareous lake-basin fen, 21, 247–255, 2001.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A.,

da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Robertson, J. B., Schoen, R., and Barraclough, J. T.: The influence of liquid waste disposal on the geochemistry of water at the National Reactor Testing Station, Idaho, 1952-1970, 1974.

Rosa, W. (Ed.): Transforming Our World: The 2030 Agenda for Sustainable Development, in: A New Era in Global Health, Springer Publishing Company, New York, NY, https://doi.org/10.1891/9780826190123.ap02, 2017.

Roy, E. D., Wagner, C. R. H., and Niles, M. T.: Hot spots of opportunity for improved cropland nitrogen management across the United States, Environ. Res. Lett., 16, 035004, https://doi.org/10.1088/1748-9326/abd662, 2021.

Ruano, M. V., Ribes, J., Seco, A., and Ferrer, J.: An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors, Environmental Modelling & Software, 37, 103–109, https://doi.org/10.1016/j.envsoft.2012.03.008, 2012.

Runkel, R. L., Crawford, C. G., and Cohn, T. A.: Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers, in: Techniques and Methods Book 4, US Geological Survey, Reston VA, 75, 2004.

Sadegh, M. and Vrugt, J. A.: Bridging the gap between GLUE and formal statistical approaches: approximate Bayesian computation, Hydrol. Earth Syst. Sci., 17, 4831–4850, https://doi.org/10.5194/hess-17-4831-2013, 2013.

Salt Institute: The Snowfighters Handbook: 40th Year Edition, A practical guide for snow and ice control, Salt Institute, Alexandria Virginia, 2007.

Samal, N. R., Wollheim, W., Zuidema, S., Stewart, R., Zhou, Z., Mineau, M., Borsuk, M., Gardner, K., Glidden, S., Huang, T., Lutz, D., Mavrommati, G., Thorn, A., Wake, C., and Huber, M.: A coupled terrestrial and aquatic biogeophysical model of the Upper Merrimack River watershed, New Hampshire, to inform ecosystem services evaluation and management under climate and land-cover change, 22, 18, https://doi.org/10.5751/ES-09662-220418, 2017.

Sander, A., Novotny, E., Mohseni, O., and Stefan, H.: Inventory of Road Salt Use in the Minneapolis/St. Paul Metropolitan Area, 2007.

Sangrey, D. A., Harrop-Williams, K. O., and Klaiber, J. A.: Predicting Ground-Water Response to Precipitation, 110, 957–975, https://doi.org/10.1061/(ASCE)0733-9410(1984)110:7(957), 1984.

Santhi, C., Kannan, N., White, M., Di Luzio, M., Arnold, J. G., Wang, X., and Williams, J. R.: An integrated modeling approach for estimating the water quality benefits of conservation practices at the river basin scale., 43, 177–198, http://dx.doi.org.unh.idm.oclc.org/10.2134/jeq2011.0460, 2014.

Sauer, T., Havlík, P., Schneider, U. A., Schmid, E., Kindermann, G., and Obersteiner, M.: Agriculture and resource availability in a changing world: The role of irrigation, 46, W06503, https://doi.org/10.1029/2009WR007729, 2010.

Scanlon, B. R., Reedy, R. C., Faunt, C. C., Pool, D., and Uhlman, K.: Enhancing drought resilience with conjunctive use and managed aquifer recharge in California and Arizona, 11, 035013, https://doi.org/10.1088/1748-9326/11/3/035013, 2016.

Scavia, D., Bertani, I., Obenour, D. R., Turner, R. E., Forrest, D. R., and Katin, A.: Ensemble modeling informs hypoxia management in the northern Gulf of Mexico, PNAS, https://doi.org/10.1073/pnas.1705293114, 2017.

Scherberg, J., Baker, T., Selker, J. S., and Henry, R.: Design of Managed Aquifer Recharge for Agricultural and Ecological Water Supply Assessed Through Numerical Modeling, Water Resour Manage, 28, 4971–4984, https://doi.org/10.1007/s11269-014-0780-2, 2014.

Schipper, L. A., Robertson, W. D., Gold, A. J., Jaynes, D. B., and Cameron, S. C.: Denitrifying bioreactors—An approach for reducing nitrate loads to receiving waters, Ecological Engineering, 36, 1532–1543, https://doi.org/10.1016/j.ecoleng.2010.04.008, 2010.

Shaw, S. B., Marjerison, R. D., Bouldin, D. R., Parlange, J.-Y., and Walter, M. T.: Simple Model of Changes in Stream Chloride Levels Attributable to Road Salt Applications, 138, 112–118, https://doi.org/10.1061/(ASCE)EE.1943-7870.0000458, 2012.

Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., and Smith, D. R.: Impacts of impervious surface on watershed hydrology: A review, 2, 263–275, https://doi.org/10.1080/15730620500386529, 2005.

Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, Journal of Hydrology, 384, 198–217, https://doi.org/10.1016/j.jhydrol.2009.07.031, 2010.

Silva, O., Carrera, J., Dentz, M., Kumar, S., Alcolea, A., and Willmann, M.: A general real-time formulation for multi-rate mass transfer problems, 13, 1399–1411, 2009.

Simons, G. W. H. (Gijs), Bastiaanssen, W. G. M. (Wim), and Immerzeel, W. W. (Walter): Water reuse in river basins with multiple users: A literature review, 522, 558–571, https://doi.org/10.1016/j.jhydrol.2015.01.016, 2015.

Sivapalan, M., Savenije, H. H. G., and Blöschl, G.: Socio-hydrology: A new science of people and water: INVITED COMMENTARY, Hydrol. Process., 26, 1270–1276, https://doi.org/10.1002/hyp.8426, 2012.

Sokal, A.: Monte Carlo methods in statistical mechanics: foundations and new algorithms, Springer, 1997.

Soule, M. J., Nimon, R. W., and Mullarkey, D. J. (Eds.): RISK MANAGEMENT AND THE ENVIRONMENT: IMPACTS AT THE INTENSIVE AND EXTENSIVE MARGINS, 36 pp., https://doi.org/10.22004/ag.econ.20670, 2001.

Southworth, M.: Streets and the shaping of towns and cities, Island Press, Washington, DC, 197 pp., 2003.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou: Frequency Analysis of Extreme Events, in: Handbook of Hydrology, edited by: Maidment, D. R., 1992.

Stedinger, J. R., Vogel, R. M., Lee, S. U., and Batchelder, R.: Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, Water Resour. Res., 44, W00B06, https://doi.org/10.1029/2008WR006822, 2008.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., and de Haan, C.: Livestock's long shadow, 2006.

Stern, I.: Dataset: project=CMIP5, model=NCAR Community Climate System Model, CCSM version 4, experiment=1 percent per year CO2, time_frequency=mon, modeling realm=atmos, ensemble=r1i1p1, version=20121031, 2012.

Stewart, M. K., Morgenstern, U., and McDonnell, J. J.: Truncation of stream residence time: how the use of stable isotopes has skewed our concept of streamwater age and origin, Hydrol. Process., 24, 1646–1659, https://doi.org/10.1002/hyp.7576, 2010.

Stewart, R. J., Wollheim, W. M., Gooseff, M. N., Briggs, M. A., Jacobs, J. M., Peterson, B. J., and Hopkinson, C. S.: Separation of river network–scale nitrogen removal among the main channel and two transient storage compartments, 47, https://doi.org/10.1029/2010WR009896, 2011.

Stewart, R. J., Wollheim, W. M., Miara, A., Vörösmarty, C. J., Fekete, B., Lammers, R. B., and Rosenzweig, B.: Horizontal cooling towers: riverine ecosystem services and the fate of thermoelectric heat in the contemporary Northeast US, Environ. Res. Lett., 8, 025010, https://doi.org/10.1088/1748-9326/8/2/025010, 2013.

Stites, W. and Kraft, G. J.: Nitrate and chloride loading to groundwater from an irrigated north-central U.S. sand-plain vegatable field, J. Environ. Qual., 30, 1176–1184, 2001.

Struzeski, E.: Environmental Impact of Highway Deicing, U.S. Environmental Protection Agency, Edison, NJ, 1971.

Sugg, Z.: Assessing U.S. Farm Drainage:, 8, 2007.

Sukow, J.: COMPARISON OF ENHANCED SNAKE PLAIN AQUIFER MODEL VERSION 2.1 WITH VERSION 1.1 VIA THE CURTAILMENT SCENARIO, 2012.

Svensson, T., Lovett, G. M., and Likens, G. E.: Is chloride a conservative ion in forest ecosystems?, Biogeochemistry, 107, 125–134, https://doi.org/10.1007/s10533-010-9538-y, 2010.

Tal, A.: Rethinking the sustainability of Israel's irrigation practices in the Drylands, Water Research, 90, 387–394, https://doi.org/10.1016/j.watres.2015.12.016, 2016.

Tanner, C. C., Sukias, J. P. S., and Yates, C. R.: Constructed Wetland Treatment of Tile Drainage, National Institute of Water & Atmospheric Research Ltd, Hamilton, NZ, 2010.

Thorn, A. M., Wake, C. P., Grimm, C. D., Mitchell, C. R., Mineau, M. M., and Ollinger, S. V.: Development of scenarios for land cover, population density, impervious cover, and conservation in New Hampshire, 2010–2100, 22, https://doi.org/10.5751/ES-09733-220419, 2017.

Tian, H., Xu, R., Pan, S., Yao, Y., Bian, Z., Cai, W.-J., Hopkinson, C. S., Justic, D., Lohrenz, S., Lu, C., Ren, W., and Yang, J.: Long-Term Trajectory of Nitrogen Loading and Delivery From Mississippi River Basin to the Gulf of Mexico, 34, e2019GB006475, https://doi.org/10.1029/2019GB006475, 2020.

Tidwell, V. C., Moreland, B. D., Zemlick, K. M., Roberts, B. L., Passell, H. D., Jensen, D., Forsgren, C., Sehlke, G., Cook, M. A., King, C. W., and Larsen, S.: Mapping water availability, projected use and cost in the western United States, Environ. Res. Lett., 9, 064009, https://doi.org/10.1088/1748-9326/9/6/064009, 2014.

Trabucco, A. and Zomer, R.: Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2, , https://doi.org/10.6084/m9.figshare.7504448.v3, 2019.

Tran, D., Kovacs, K., and Wallander, S.: Long run optimization of landscape level irrigation through managed aquifer recharge or expanded surface reservoirs, Journal of Hydrology, 579, 124220, https://doi.org/10.1016/j.jhydrol.2019.124220, 2019.

Tran, D. Q., Kovacs, K. F., and West, G. H.: Spatial economic predictions of managed aquifer recharge for an agricultural landscape, Agricultural Water Management, 241, 106337, https://doi.org/10.1016/j.agwat.2020.106337, 2020.

Trowbridge, P. R., Kahl, J. S., Sassan, D. A., Heath, D. L., and Walsh, E. M.: Relating road salt to exceedances of the water quality standard for chloride in New Hampshire streams, 44, 4903–4909, 2010.

Troy, T. J., Konar, M., Srinivasan, V., and Thompson, S.: Moving sociohydrology forward: a synthesis across studies, 19, 3667–3679, https://doi.org/10.5194/hess-19-3667-2015, 2015.

Tukey, J. W.: Comparing Individual Means in the Analysis of Variance, Biometrics, 5, 99, https://doi.org/10.2307/3001913, 1949.

Uddin, J., Smith, R., Hancock, N., and Foley, J. P.: Droplet evaporation losses during sprinkler irrigation: an overview, in: Australian Irrigation Conference and Exibition 2010: Proceedings, Australian Irrigation Conference and Exibition 2010: One Water Many Futures, Sydney, Australia, 1–10, 2010.

U.S. Census Bureau: County Census Statistics, Massachusetts, 2015.

U.S. Geological Survey: National Water Information System data available on the World Wide Web (USGS Water Data for the Nation), https://doi.org/dx.doi.org/10.5066/F7P55KJN, 2016.

National inventory of dams: nid.usace.army.mil, last access: 17 January 2020.

USDA: Agricultural Conservation Easement Program, , 7 CFR Part 1468, 19, 2021.

USDA NASS: 2012 Census of Agriculture, Ag Census Web Maps, 2014.

USEPA: 2012 Edition of the Drinking Water Standards and Health Advisories, U.S. Environmental Protection Agency, Washington, D. C., 2012.

USEPA: National Rivers and Streams Assessment 2008-2009, A collaborative survey, DRAFT, U.S. Environmental Protection Agency, Washington, D. C., 2013.

National Hydrography Dataset (NHD): https://www2.usgs.gov/science/cite-view.php?cite=250, last access: 16 April 2019.

Usher, W., Herman, J., Iwanaga, T., Teixeira, L., CELLIER, N., Whealton, C., Hadka, D., Xantares, Bernardoct, Rios, F., Mutel, C., Cederstrand, E., TobiasKAndersen, Engelen, J. V., ANtlord, Kranas, H., Vaibhav Kumar Dixit, and Bsteubing: SALib/SALib: SALib v1.3.12, Zenodo, https://doi.org/10.5281/ZENODO.4270041, 2020.

Van Kirk, R. W., Contor, B. A., Morrisett, C. N., Null, S. E., and Loibman, A. S.: Potential for Managed Aquifer Recharge to Enhance Fish Habitat in a Regulated River, 12, 673, https://doi.org/10.3390/w12030673, 2020.

Van Meter, K. J., Van Cappellen, P., and Basu, N. B.: Legacy nitrogen may prevent achievement of water quality goals in the Gulf of Mexico, Science, 360, 427–430, https://doi.org/10.1126/science.aar4462, 2018.

Vanham, D., Hoekstra, A. Y., Wada, Y., Bouraoui, F., de Roo, A., Mekonnen, M. M., van de Bund, W. J., Batelaan, O., Pavelic, P., Bastiaanssen, W. G. M., Kummu, M., Rockström, J., Liu, J., Bisselink, B., Ronco, P., Pistocchi, A., and Bidoglio, G.: Physical water scarcity metrics for monitoring progress towards SDG target 6.4: An evaluation of indicator 6.4.2 “Level of water stress,” Science of The Total Environment, 613–614, 218–232, https://doi.org/10.1016/j.scitotenv.2017.09.056, 2018.

van der Velde, Y., de Rooij, G. H., and Torfs, P. J. J. F.: Catchment-scale non-linear groundwater-surface water interactions in densely drained lowland catchments, Hydrol. Earth Syst. Sci., 13, 1867–1885, https://doi.org/10.5194/hess-13-1867-2009, 2009.

Venetis, C.: A STUDY ON THE RECESSION OF UNCONFINED ACQUIFERS, International Association of Scientific Hydrology. Bulletin, 14, 119–125, https://doi.org/10.1080/02626666909493759, 1969.

Vogelmann, J. E., Sohl, T. L., Campbell, P. V., and Shaw, D. M.: Regional Land Cover Characterization Using Landsat Thematic Mapper Data and Ancillary Data Sources, Environ Monit Assess, 51, 415–428, https://doi.org/10.1023/A:1005996900217, 1998.

Vörösmarty, C. J., Moore, B., Grace, A. L., Gildea, M. P., Melillo, J. M., Peterson, B. J., Rastetter, E. B., and Steudler, P. A.: Continental scale models of water balance and fluvial transport: An application to South America, Global Biogeochem. Cycles, 3, 241–265, https://doi.org/10.1029/GB003i003p00241, 1989.

Vörösmarty, C. J., Federer, C. A., and Schloss, A. L.: Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling, Journal of Hydrology, 207, 147–169, https://doi.org/10.1016/S0022-1694(98)00109-7, 1998.

Vrugt, J. A. and Sadegh, M.: Toward diagnostic model calibration and evaluation: Approximate Bayesian computation, Water Resour. Res., 49, 4335–4345, https://doi.org/10.1002/wrcr.20354, 2013.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44, W00B09, https://doi.org/10.1029/2007WR006720, 2008.

Wake, C., Burakowski, E., Wilkinson, P., Hayhoe, K., Stoner, A., Keeley, C., and LaBranche, J.: Climate Change in Northern New Hampshire Past, Present, and Future, Climate Solutions New England, Durham NH, 2014a.

Wake, C., Burakowski, E., Wilkinson, P., Hayhoe, K., Stoner, A., Keeley, C., and LaBranche, J.: Climate Change in Southern New Hampshire Past, Present, and Future, Climate Solutions New England, Durham NH, 2014b.

Wales, D. J. and Doye, J. P. K.: Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms, J. Phys. Chem. A, 101, 5111–5116, https://doi.org/10.1021/jp970984n, 1997.

Werner, T. M. and Kadlec, R. H.: Wetland residence time distribution modeling, 15, 77–90, 2000.

Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C.-Y.: Calibration of hydrological models using flow-duration curves,

Hydrol. Earth Syst. Sci., 15, 2205–2227, https://doi.org/10.5194/hess-15-2205-2011, 2011.

Westerberg, I. K., Gong, L., Beven, K. J., Seibert, J., Semedo, A., Xu, C.-Y., and Halldin, S.: Regional water balance modelling using flow-duration curves with observational uncertainties, Hydrol. Earth Syst. Sci., 18, 2993–3013, https://doi.org/10.5194/hess-18-2993-2014, 2014.

Whitehead, R. L.: Geohydrologic framework of the Snake River plain regional aquifer system, Idaho and eastern Oregon, U.S. Government Printing Office, 1992.

Williams, A. P., Cook, B. I., and Smerdon, J. E.: Rapid intensification of the emerging southwestern North American megadrought in 2020–2021, Nat. Clim. Chang., 12, 232–234, https://doi.org/10.1038/s41558-022-01290-z, 2022.

Willmott, C. J., Rowe, C. M., and Mintz, Y.: Climatology of the terrestrial seasonal water cycle, 5, 589–606, https://doi.org/10.1002/joc.3370050602, 1985.

Wisser, D., Frolking, S., Douglas, E. M., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H.: Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, 35, L24408, https://doi.org/10.1029/2008GL035296, 2008.

Wisser, D., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H.: Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), 14, 1–24, https://doi.org/10.5194/hess-14-1-2010, 2010.

Wollheim, W., Vörösmarty, C. J., Peterson, B. J., Seitzinger, S. P., and Hopkinson, C. S.: Relationship between river size and nutrient removal, 33, https://doi.org/10.1029/2006GL025845, 2006.

Wollheim, W., Peterson, B. J., Thomas, S. M., Hopkinson, C. S., and Vorosmarty, C. J.: Dynamics of N removal over annual time periods in a suburban river network, 2008a.

Wollheim, W., Vörösmarty, C. J., Bouwman, A. F., Green, P., Harrison, J., Linder, E., Peterson, B. J., Seitzinger, S. P., and Syvitski, J. P. M.: Global N removal by freshwater aquatic systems using a spatially distributed, within-basin approach, Global Biogeochem. Cycles, 22, https://doi.org/10.1029/2007GB002963, 2008b.

Wollheim, W., Harms, T. K., Peterson, B. J., Morkeski, K., Hopkinson, C. S., Stewart, R. J., Gooseff, M. N., and Briggs, M. A.: Nitrate uptake dynamics of surface transient storage in stream channels and fluvial wetlands, Biogeochemistry, 120, 239–257, https://doi.org/10.1007/s10533-014-9993-y, 2014.

Wu, Q. and Lane, C. R.: Delineating wetland catchments and modeling hydrologic connectivity using lidar data and aerial imagery, Hydrol Earth Syst Sci, 21, 3579–3595, https://doi.org/10.5194/hess-21-3579-2017, 2017.

Wulfhorst, J. D. and Glenn, E.: Irrigation, community, and historical development along the upper Snake River, 76, 434–447, 2002.

Xian, G., Homer, C., Dewitz, J., Fry, J., Hossain, N., and Wickham, J.: Change of Impervious Surface Area Between 2001 and 2006 in the Conterminous United States, 77, 758–762, 2011.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resour. Res., 55, 5053–5073, https://doi.org/10.1029/2019WR024873, 2019.

Yaraghi, N., Ronkanen, A., Darabi, H., Kløve, B., and Torabi Haghighi, A.: Impact of managed aquifer recharge structure on river flow regimes in arid and semi-arid climates, Science of The Total Environment, 675, 429–438, https://doi.org/10.1016/j.scitotenv.2019.04.253, 2019.

Yu, P.-S. and Yang, T.-C.: Using synthetic flow duration curves for rainfall–runoff model calibration at ungauged sites, Hydrol. Process., 14, 117–133, https://doi.org/10.1002/(SICI)1099-1085(200001)14:1<117::AID-HYP914>3.0.CO;2-Q, 2000.

Zimmerman, E. K., Tyndall, J. C., and Schulte, L. A.: Using Spatially Targeted Conservation to Evaluate Nitrogen Reduction and Economic Opportunities for Best Management Practice Placement in Agricultural Landscapes, Environmental Management, 64, 313–328, https://doi.org/10.1007/s00267-019-01190-7, 2019.

Zuidema, S., Wollheim, W., Mineau, M. M., Green, M. B., and Stewart, R. J.: Controls of Chloride Loading and Impairment at the River Network Scale in New England, 47, 839–847, https://doi.org/10.2134/jeq2017.11.0418, 2018.

Zuidema, S., Grogan, D., Prusevich, A., Lammers, R., Gilmore, S., and Williams, P.: Interplay of changing irrigation technologies and water reuse: Example from the Upper Snake River Basin, Idaho, USA, 24, 5231–5249, https://doi.org/10.5194/hess-24-5231-2020, 2020.

de Zwart, D., Dyer, S. D., Posthuma, L., and Hawkins, C. P.: Predictive models attribute effects on fish assemblages to toxicity and habitat alteration, Ecological Applications, 16, 1295–1310, https://doi.org/10.1890/1051-0761(2006)016[1295:PMAEOF]2.0.CO;2, 2006.