From datasets to decisions

a repeatable workflow for groundwater decision support

Michael Fienen[1], Nicholas Corson-Dosch[1], Jeremy White[2], Andrew Leaf[1], and Randall Hunt[1]

[1]United States Geological Survey Upper Midwest Water Science Center

Madison, Wisconsin, USA

[2]INTERA - Geoscience and Engineering Solutions

Boulder, Colorado USA

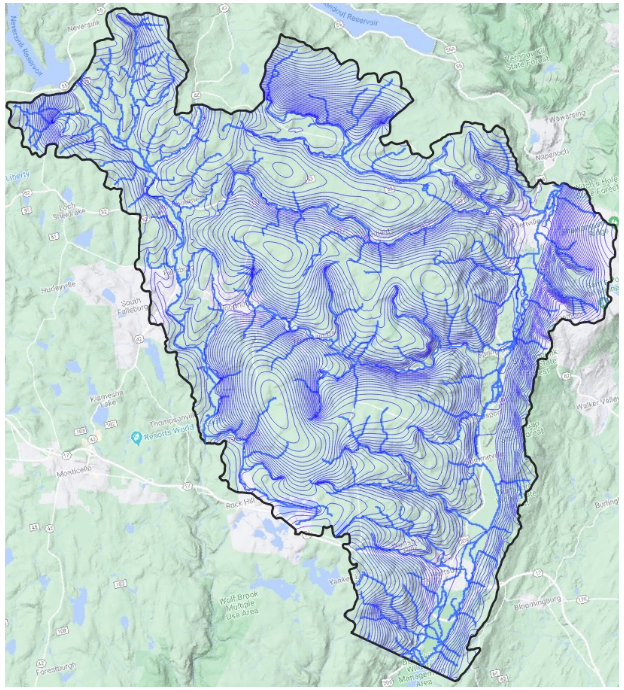
USGS

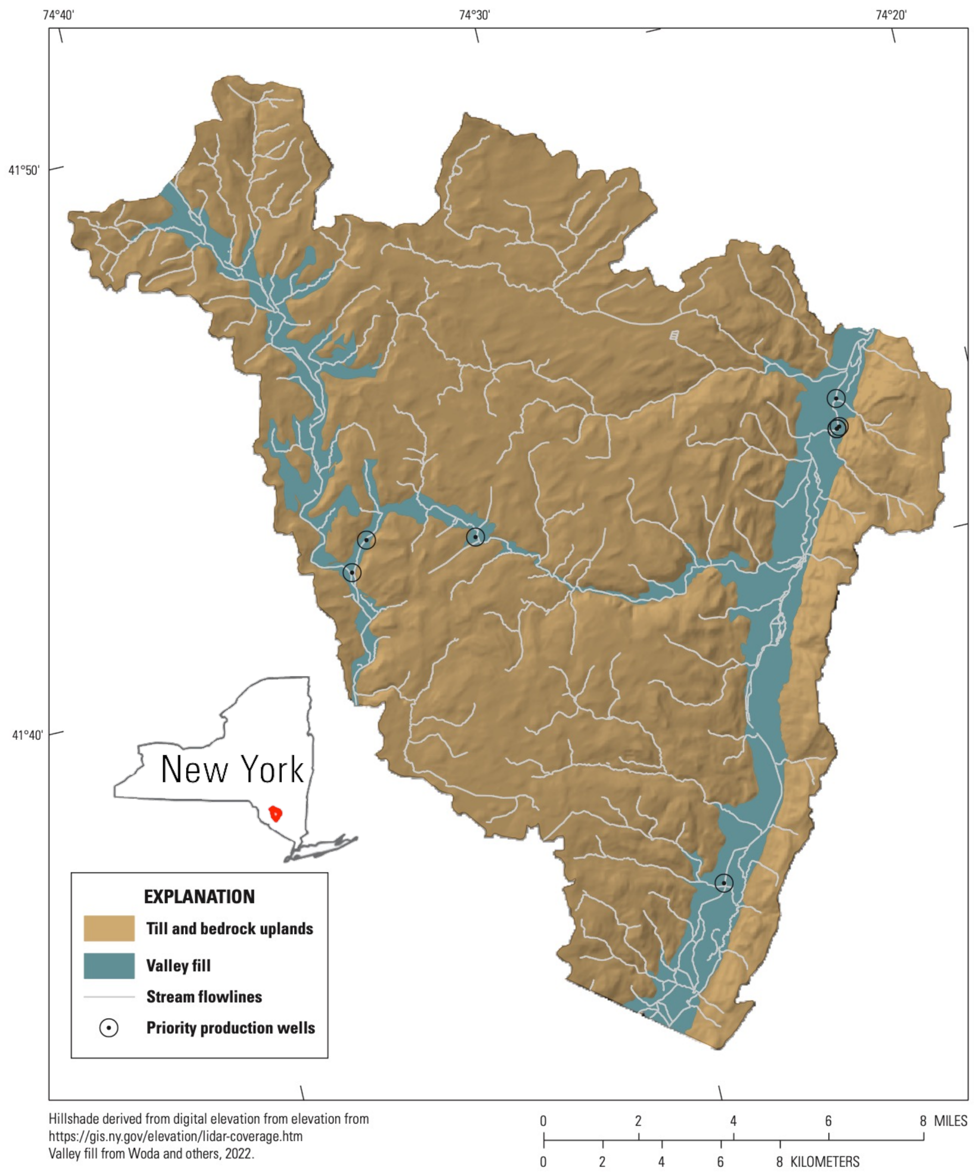
EGU General Assembly

Sharing is encouraged

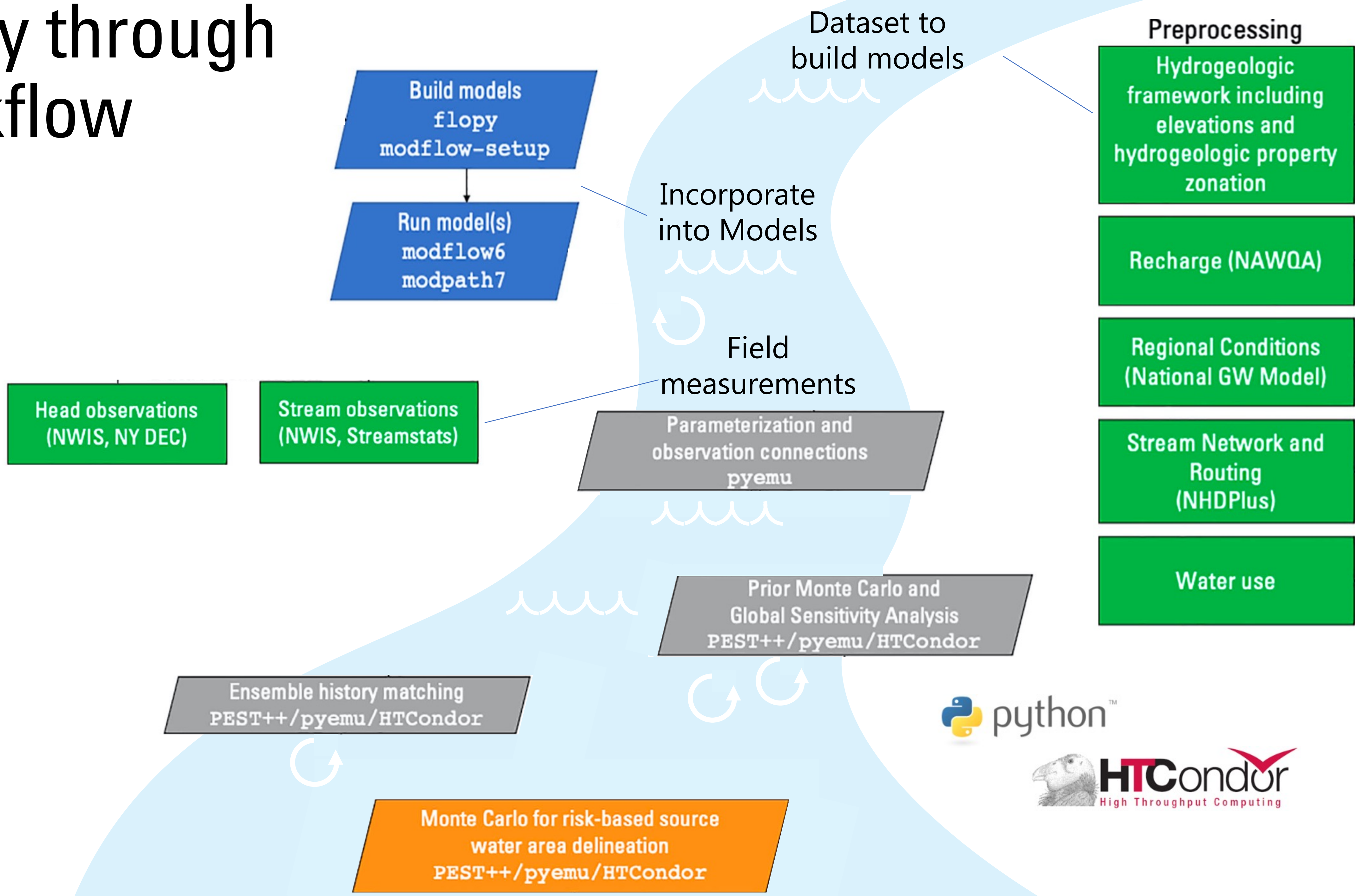
Problem Statement

Rapidly *develop a* ***robust****, risk-based decision-support system for wellhead protection in a* ***reproducible*** *workflow for the State of New York in the US*



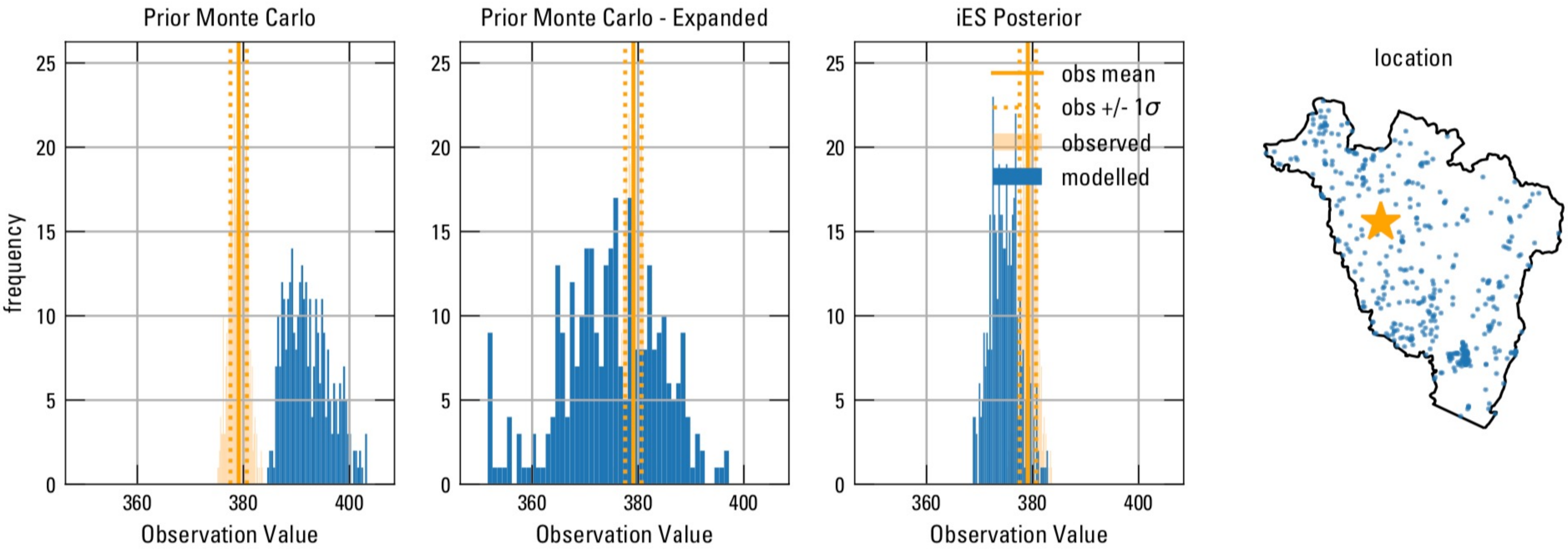


A journey through the workflow



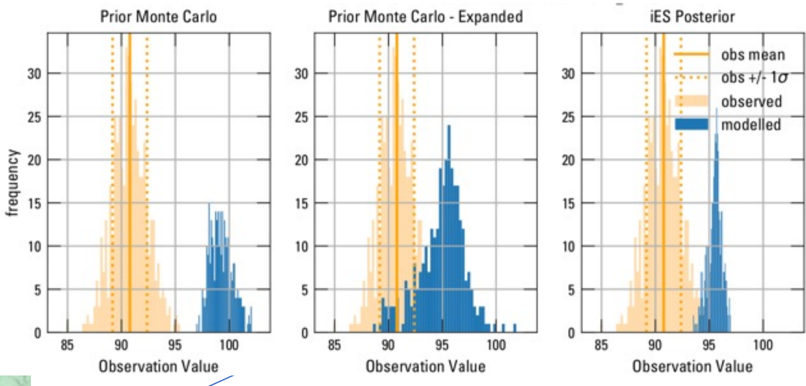
USGS

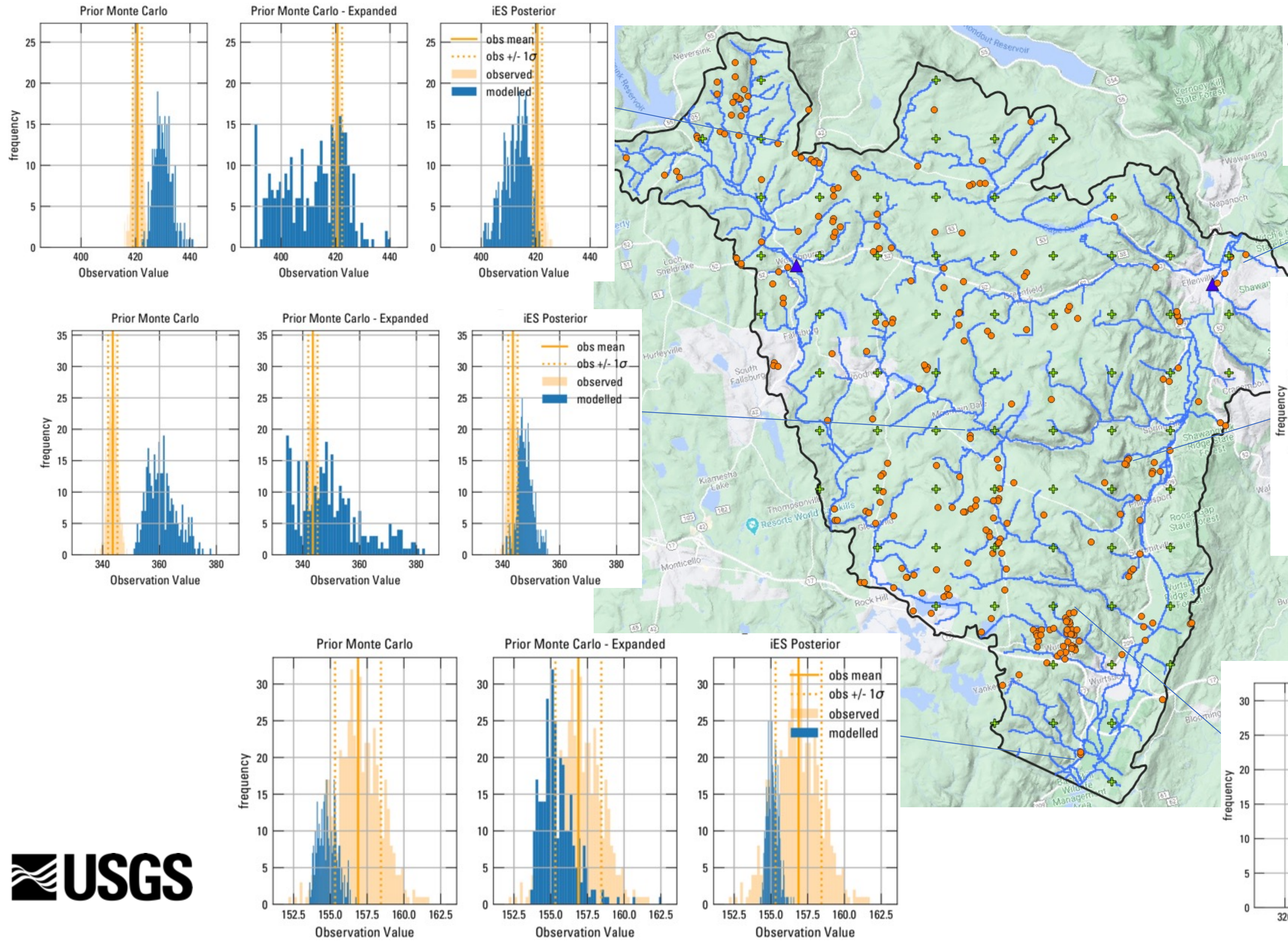
Prior Monte Carlo Analysis

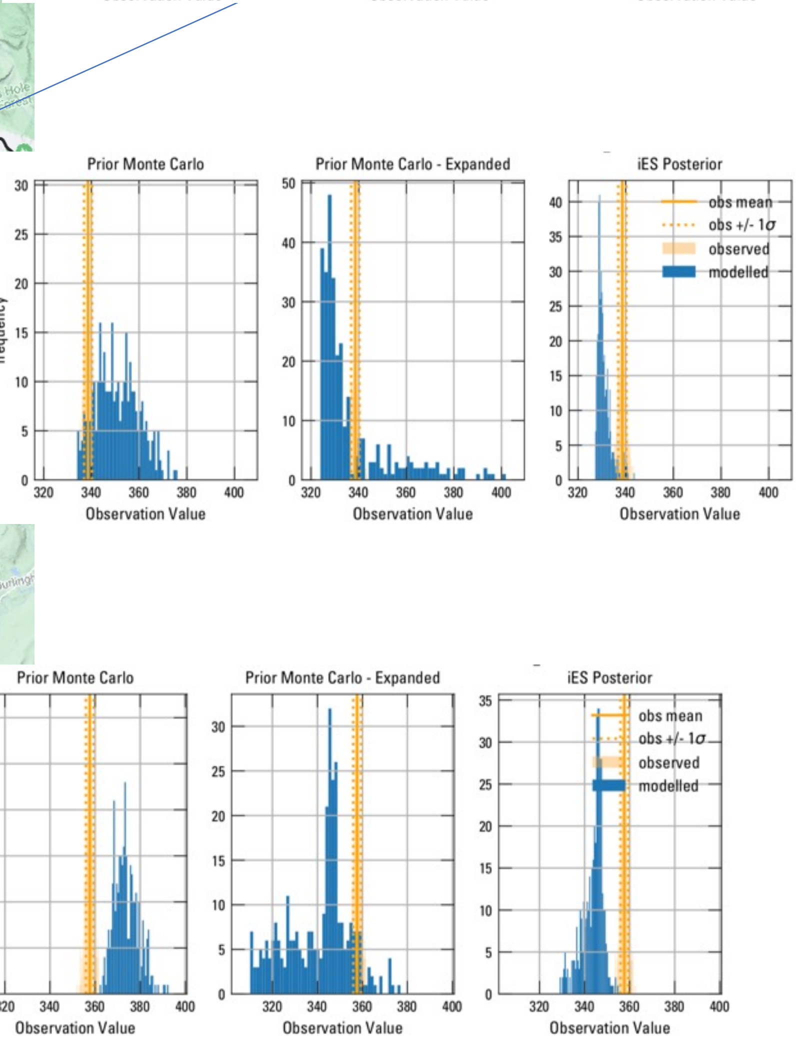


USGS

PEST++ iES Fit to Data – water levels



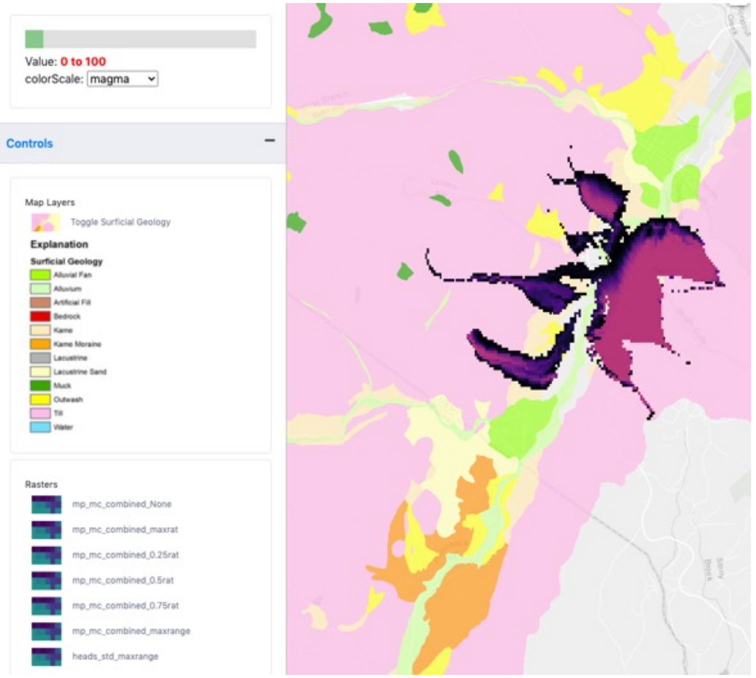




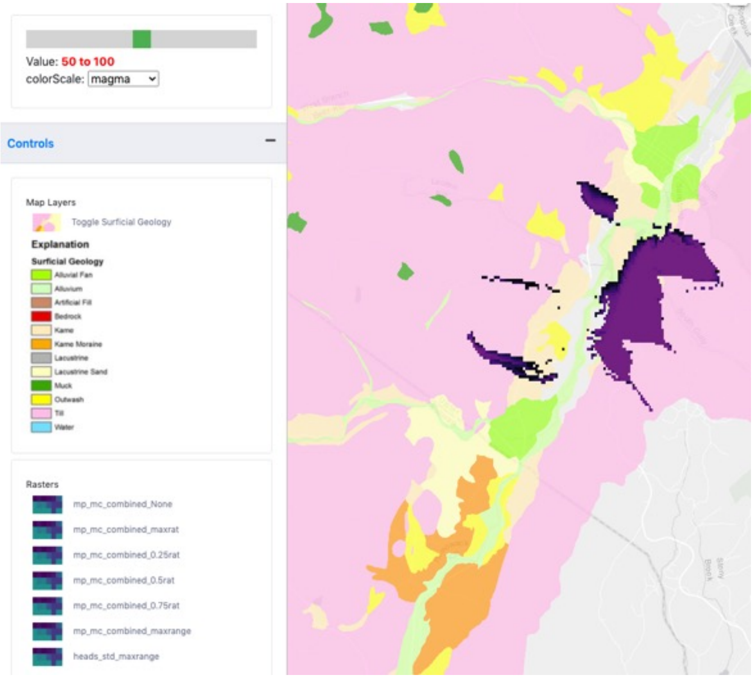
USGS

Risk-based decision support tool

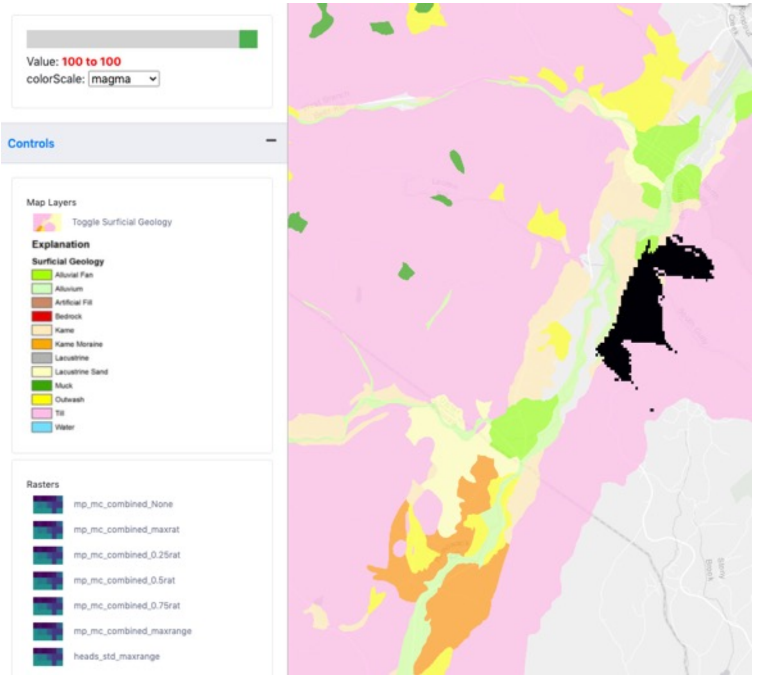
Risk averse



Risk neutral

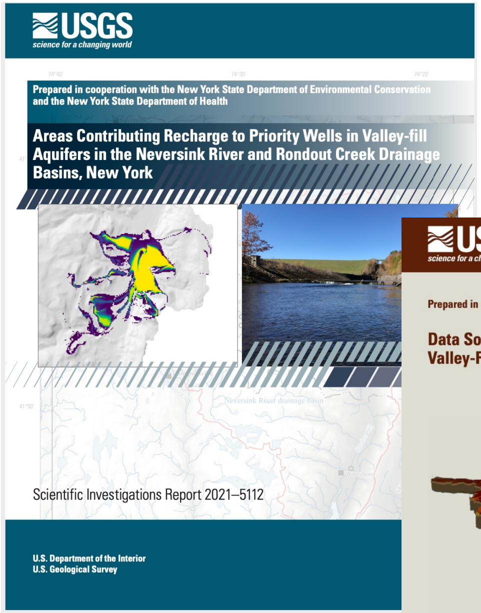


Risk tolerant



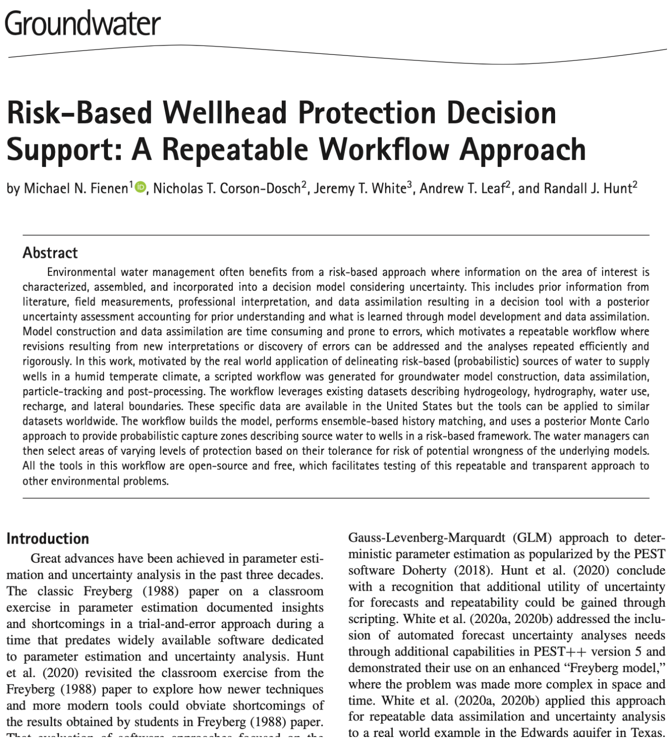
USGS

Resources for the Workflow



https://doi.org/10.3133/sir20215112

https://doi.org/10.3133/sir20225024



Groundwater

Risk-Based Wellhead Protection Decision Support: A Repeatable Workflow Approach

by Michael N. Fienen[1], Nicholas T. Corson-Dosch[2], Jeremy T. White[3], Andrew T. Leaf[2], and Randall J. Hunt[2]

Abstract

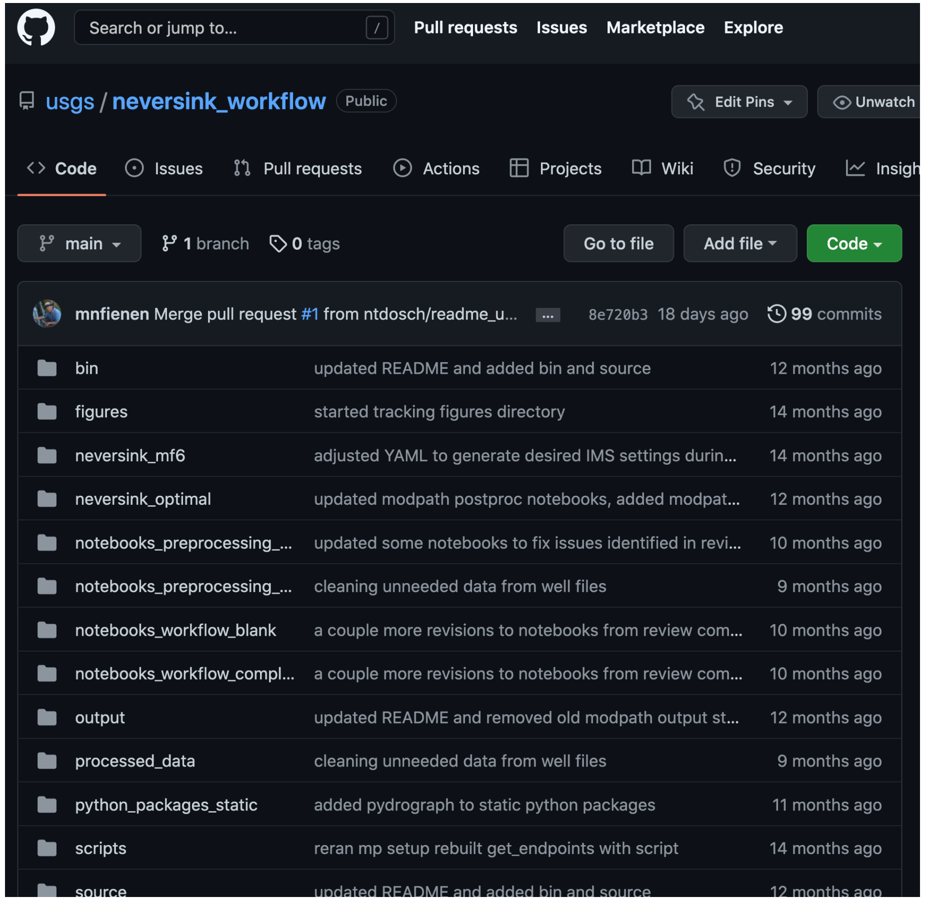
Environmental water management often benefits from a risk-based approach where information on the area of interest is characterized, assembled, and incorporated into a decision model considering uncertainty. This includes prior information from literature, field measurements, professional interpretation, and data assimilation resulting in a decision tool with a posterior uncertainty assessment accounting for prior understanding and what is learned through model development and data assimilation. Model construction and data assimilation are time consuming and prone to errors, which motivates a repeatable workflow where revisions resulting from new interpretations or discovery of errors can be addressed and the analyses repeated efficiently and rigorously. In this work, motivated by the real world application of delineating risk-based (probabilistic) sources of water to supply wells in a humid temperate climate, a scripted workflow was generated for groundwater model construction, data assimilation, particle-tracking and post-processing. The workflow leverages existing datasets describing hydrogeology, hydrography, water use, recharge, and lateral boundaries. These specific data are available in the United States but the tools can be applied to similar datasets worldwide. The workflow builds the model, performs ensemble-based history matching, and uses a posterior Monte Carlo approach to provide probabilistic capture zones describing source water to wells in a risk-based framework. The water managers can then select areas of varying levels of protection based on their tolerance for risk of potential wrongness of the underlying models. All the tools in this workflow are open-source and free, which facilitates testing of this repeatable and transparent approach to other environmental problems.

Introduction

Great advances have been achieved in parameter estimation and uncertainty analysis in the past three decades. The classic Freyberg (1988) paper on a classroom exercise in parameter estimation documented insights and shortcomings in a trial-and-error approach during a time that predates widely available software dedicated to parameter estimation and uncertainty analysis. Hunt et al. (2020) revisited the classroom exercise from the Freyberg (1988) paper to explore how newer techniques and more modern tools could obviate shortcomings of the results obtained by students in Freyberg (1988) paper.

Gauss-Levenberg-Marquardt (GLM) approach to deterministic parameter estimation as popularized by the PEST software Doherty (2018). Hunt et al. (2020) conclude with a recognition that additional utility of uncertainty for forecasts and repeatability could be gained through scripting. White et al. (2020a, 2020b) addressed the inclusion of automated forecast uncertainty analyses needs through additional capabilities in PEST++ version 5 and demonstrated their use on an enhanced "Freyberg model," where the problem was made more complex in space and time. White et al. (2020a, 2020b) applied this approach for repeatable data assimilation and uncertainty analysis to a real world example in the Edwards aquifer in Texas.

https://doi.org/10.1111/gwat.13129



https://github.com/usgs/neversink_workflow

"…an article about a computational result is advertising, not scholarship.
The actual scholarship is the full software environment, code and data, that produced the result."

Paraphrase of John Claerbout, recounted in
Donoho, David L. "An invitation to reproducible computational research."
Biostatistics 11, no. 3 (2010): 385-388.

USGS