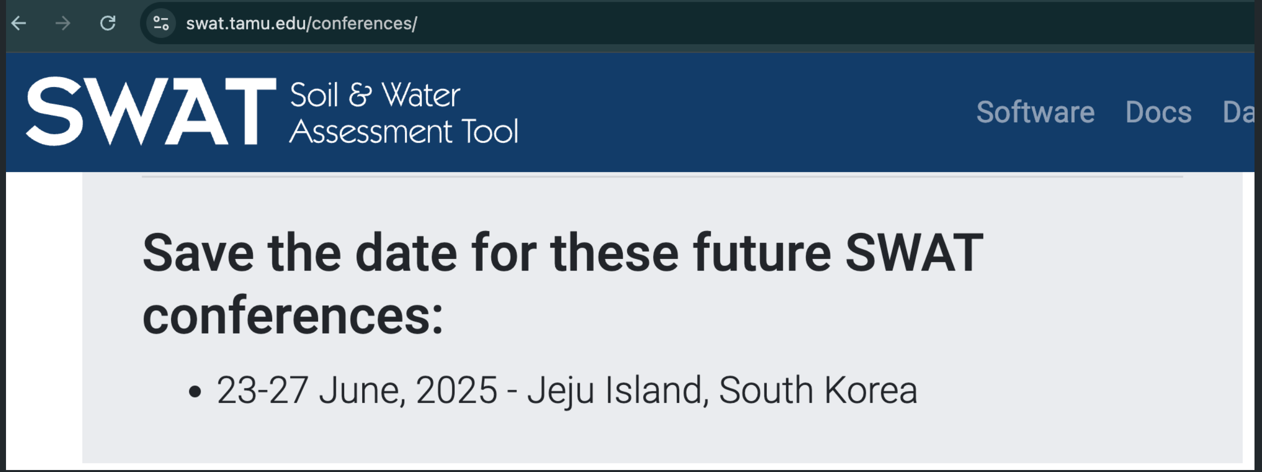
Day 3

Hands-On Practice for Uncertainty Analysis

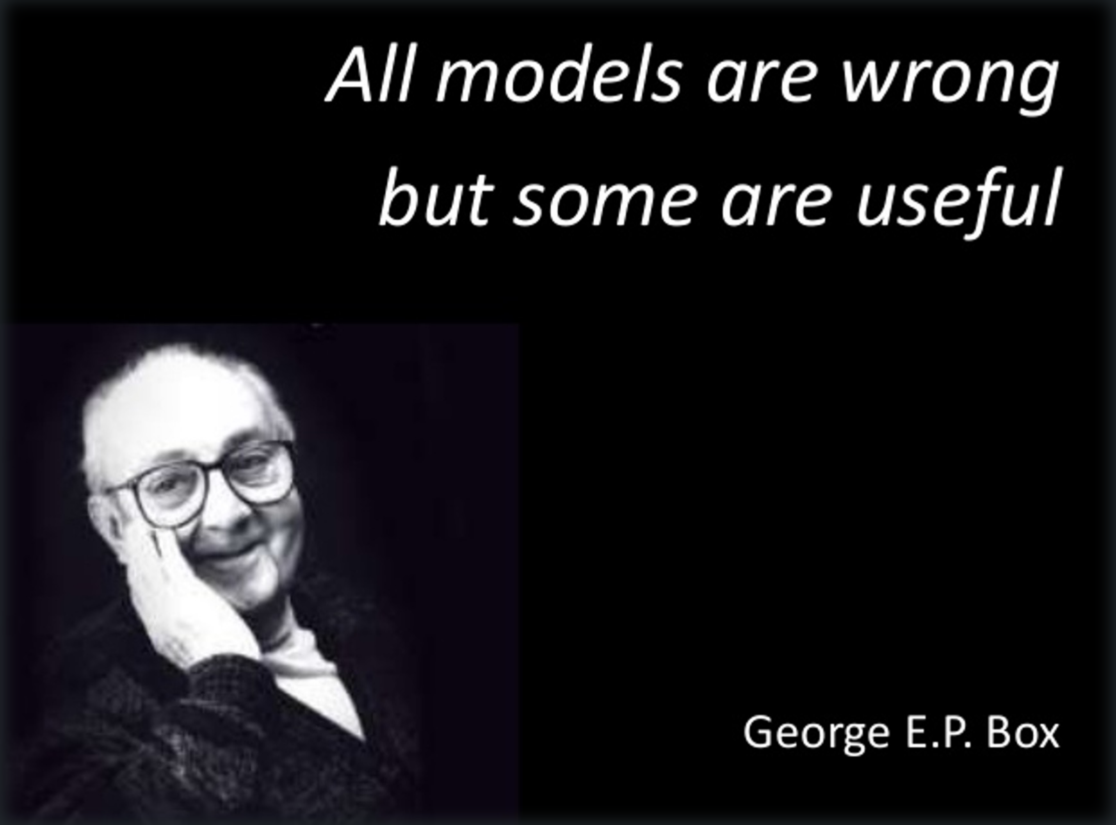


- PEST (Parameter EStimation Tool)
  - PEST official website: https://pesthomepage.org/
  - Tutorial Videos: https://pesthomepage.org/videos
  - Youtube: https://www.youtube.com/@gmdsi
  - Practice with Jupyter notebooks: https://github.com/gmdsi/GMDSI_notebooks
  - Groundwater Modeling Decision Support Initiative https://gmdsi.org/
- swatmf
  - Python Library: https://github.com/spark-brc/swatmf
  - Workflow: https://github.com/spark-brc/swatmf_wf

오늘 데이터셋

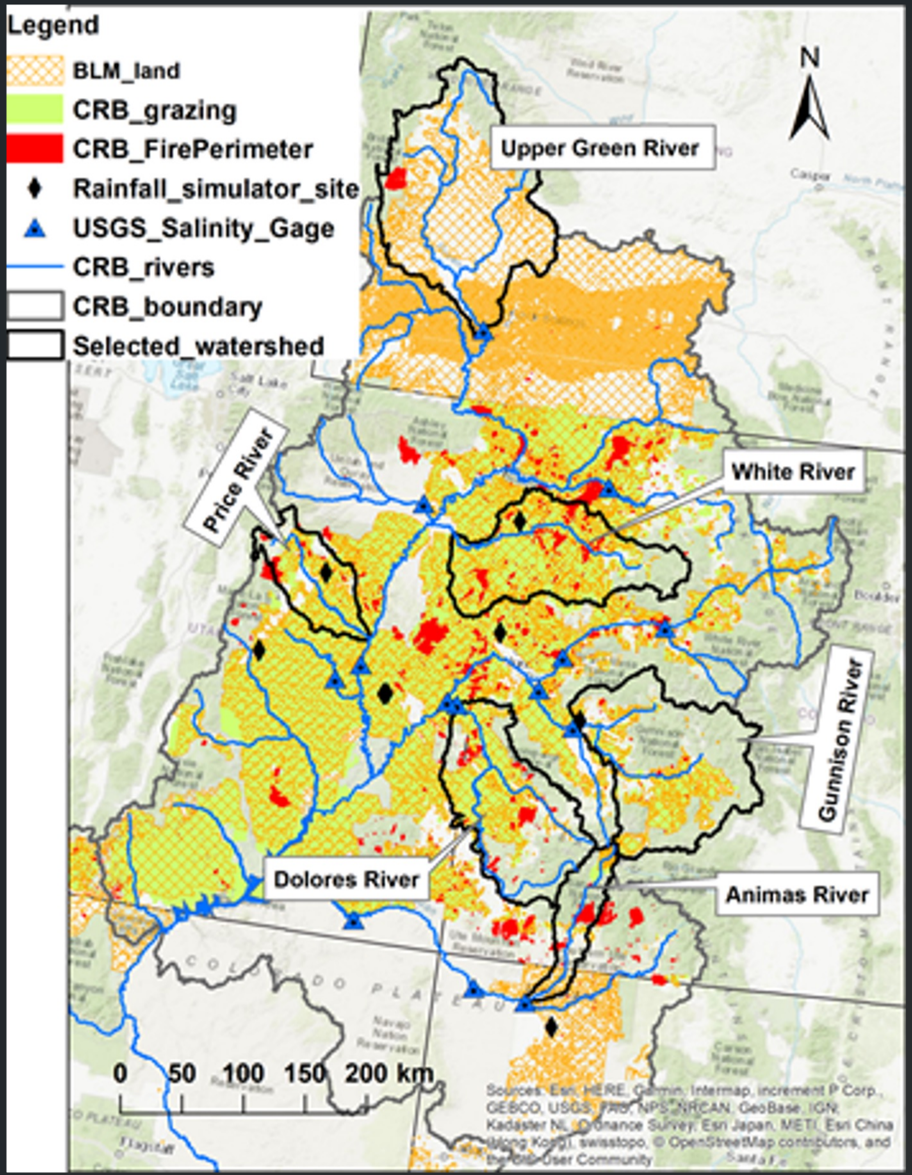
Seonggyu Park: seonggyu.park@brc.tamus.edu

All models are wrong!



All models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind.

Project Background

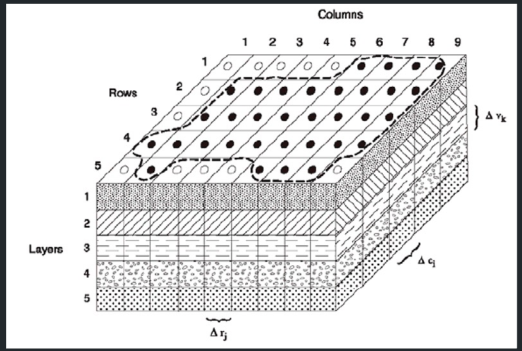


- Colorado river transports **7-9 million tons** of salt annually to the Gulf of California.
- **Irrigated agricultural districts** causes **highly saline conditions.**
- Salinity contributes to more than **$300 million dollars per year** in economic damage.
  - **Assess** the sources of salinity and its transportation in the Colorado River Basin and
  - **Develop** effective management strategies.

Approaches | Model Integration & Development

developed and maintained by

- The Agricultural Policy / Environmental eXtender (APEX) model was developed and included
  - crop growth algorithms
  - nutrient cycling in the soil profile
  - nutrient transport and
  - loading to streams in surface runoff, soil lateral flow, and groundwater flow, and in-stream transport.



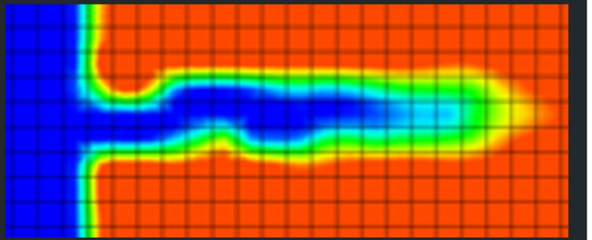
developed and maintained by

Partial Differential Equation: develop water balance for each point (cell) in the aquifer

$$\frac{\partial}{\partial x}\left(hK_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(hK_y\frac{\partial h}{\partial y}\right)+Q_{rech}-Q_{pump}-Q_{ET}=S_y\frac{\partial h}{\partial t}$$

- MODFLOW is the U.S. Geological Survey modular finite-difference flow model
  - solve the groundwater flow equation,
  - simulate the flow of groundwater through aquifers.

Reactive Transport in 3 Dimensions



https://www.pnnl.gov/downloads-rt3d (Pacific Northwest National Laboratory)

NO_3:

$$\frac{\partial C_{NO_3}}{\partial t}=-\overbrace{\frac{\partial}{\partial x_i}\left(v_iC_{NO_3}\right)}^{\text{Advection}}+\overbrace{\frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C_{NO_3}}{\partial x_j}\right)}^{\text{Dispersion}}+\overbrace{\frac{q_s}{\phi}C_{s_{NO_3}}}^{\text{Source/Sink}}\overbrace{-k_{NO_3}C_{NO_3}\left(\frac{C_{NO_3}}{K_{NO_3}+C_{NO_3}}\right)}^{\text{Denitrification}}$$

P:

$$\frac{\partial C_P}{\partial t}\underset{\text{Sorption}}{R_P}=-\overbrace{\frac{\partial}{\partial x_i}\left(v_iC_P\right)}^{\text{Advection}}+\overbrace{\frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C_P}{\partial x_j}\right)}^{\text{Dispersion}}+\overbrace{\frac{q_s}{\phi}C_{s_P}}^{\text{Source/Sink}}$$

\+ Salinity Module 8 major ions (SO_4, Ca, Mg, Na, Cl, K, CO_3, HCO_3)

Approaches | Model Integration & Development

developed and maintained by

USDA United States Department of Agriculture

Texas A&M AgriLife Research

APEX Agricultural Policy Environmental eXtender model

- The Agricultural Policy / Environmental eXtender (APEX) model was developed and included
  - crop growth algorithms
  - nutrient cycling in the soil profile
  - nutrient transport and
  - loading to streams in surface runoff, soil lateral flow, and groundwater flow, and in-stream transport.

Columns

Environmental Modelling and Software 143 (2021) 105093

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft

ELSEVIER

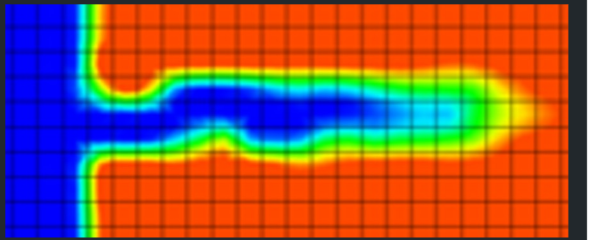
APEX-MODFLOW: A New integrated model to simulate hydrological processes in watershed systems

Ryan T. Bailey[a,*], Ali Tasdighi[a], Seonggyu Park[b], Saman Tavakoli-Kivi, PhD[a,d], Tadesse Abitew[b], Jaehak Jeong[b], Colleen H.M. Green[c], Abeyou W. Worqlul[b]

difference flow model

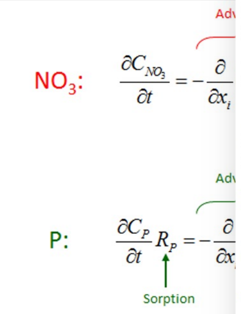
- solve the groundwater flow equation,
- simulate the flow of groundwater through aquifers.

Reactive Transport in 3 Dimensions



Pacific Northwest NATIONAL LABORATORY

https://www.pnnl.gov/downloads-rt3d (Pacific Northwest National Laboratory)



Journal of Hydrology 610 (2022) 127873

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

ELSEVIER

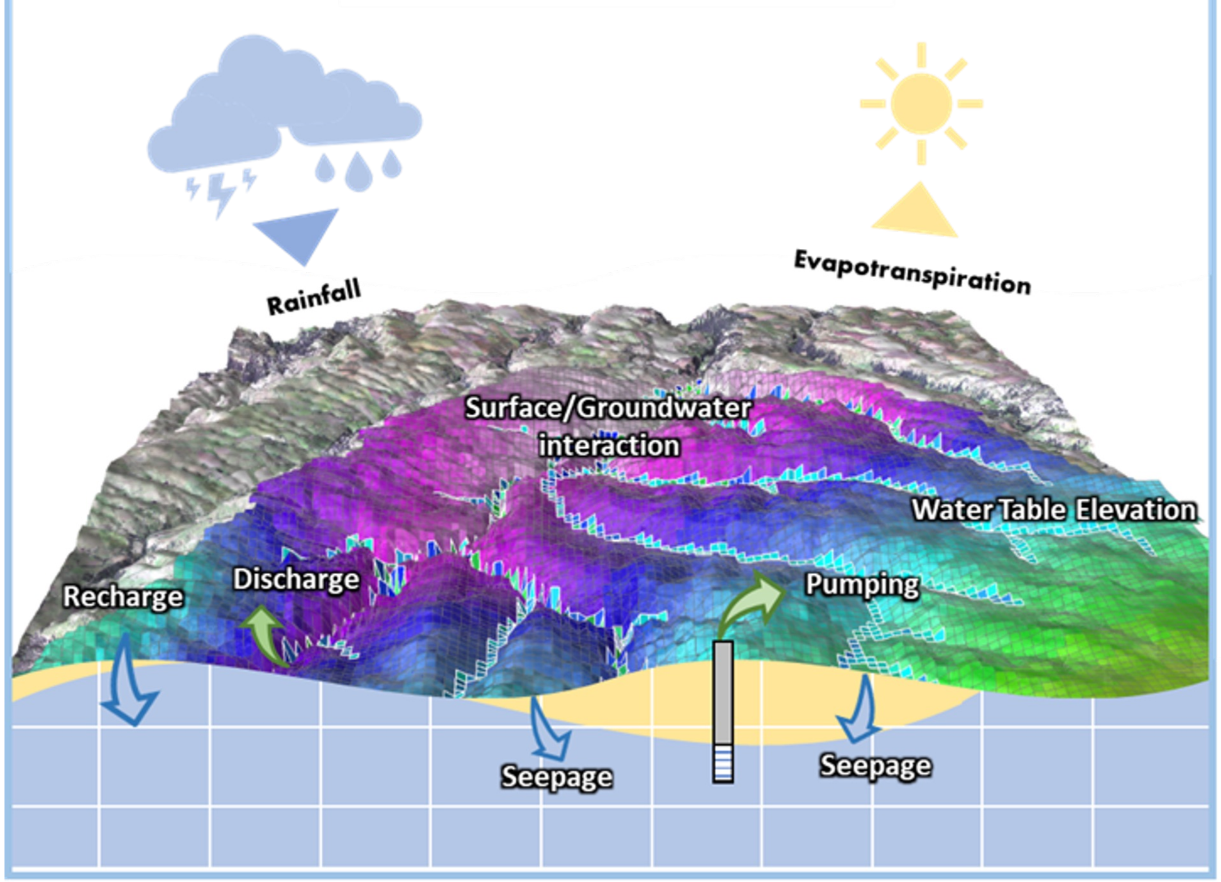
Research papers

Simulating salinity transport in High-Desert landscapes using APEX-MODFLOW-Salt

Ryan T. Bailey[a,*], Jaehak Jeong[b], Seonggyu Park[b], Colleen H.M. Green[c]

Challenges | Source of uncertainty

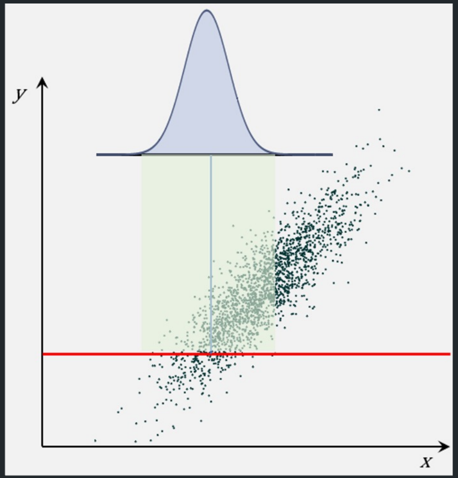
Hydrological Model Structure



Parameters for APEX-MODFLOW-RT3D-Salt (Source of Uncertainty)

- Field Parameters: Available water capacity of the soil layer,
  Runoff curve number
  Soil evaporation coefficient
  Transmissivity (T)
  Hydraulic Conductivity (K)
  Specific Storage (Ss)
  Specific Yield (Sy)
- Boundary Parameters: Initial Head Boundary
  Riverbed Conductance, Thickness
- Decision Parameters: Rainfall, Recharge, Evapotranspiration,
  Pumping and Injection Rates
- Numerical Algorithm: Spatio-Temporal Variation
- Salinity: Initial Salt Ion Concentrations / fractions

Uncertainty Analysis | Bayes Theorem



$$P(\theta|X) = P(X|\theta)P(\theta)$$

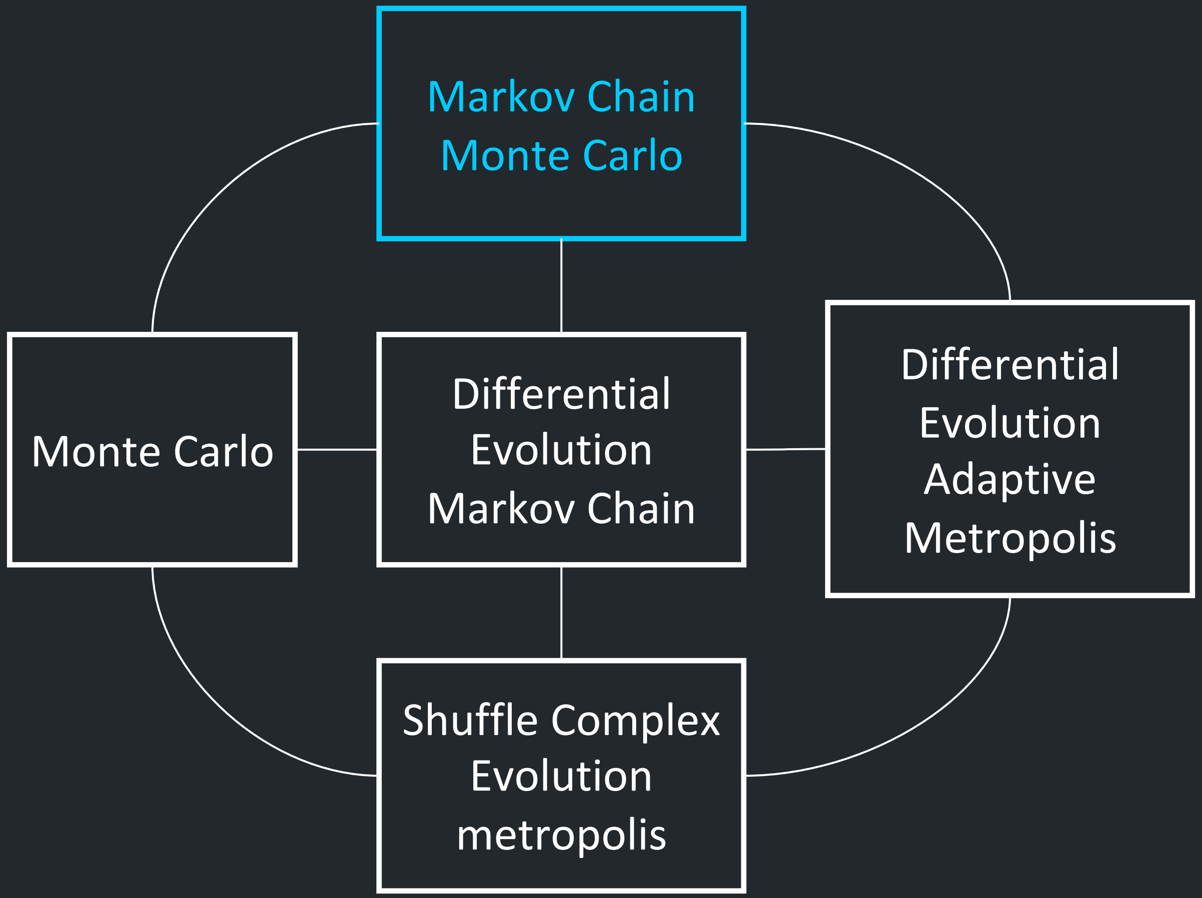
- Conditional Probability
- Prior
- Posterior
- Likelihood

Uncertainty Analysis | Sampling the posterior parameter probability distribution

- Populate a model with realizations of parameter fields
- Filter realizations based a metric of likelihood function

Uncertainty Analysis | Sampling the posterior parameter probability distribution

- Populate a model with realizations of parameter fields
- Filter realizations based a metric of likelihood function



- Less than 1000 parameters
- Short model run time
- Long convergence time

Uncertainty Analysis | Ensemble Method

- Fundamental Concept:

  approximate model-based relations with ensembles

- Ensemble methods: optimization (minimization) algorithms that use ensembles to approximate first-order (partial first derivatives) between parameters and outputs to enable (very) high-dimensional data assimilation (White et al, 2018)

Ensembles: a collection of randomly-sampled model inputs and the corresponding collection of model outputs

Ensemble Methods: Monte Carlo + Linear Algebra

Uncertainty Analysis | Ensemble Method

- Model-based relations of interest:
  - Covariance between measured and unmeasured model outputs
    - State estimation
  - Cross-covariance (e.g. gradient) between model inputs and outputs
    - Parameter estimation
- What's Great: uncertainty analysis and history matching in a single algorithmic workflow
- What's Great: forces you to think about model input uncertainty from the very beginning!

Uncertainty Analysis | Iterative Ensemble Smoother (IES)

P (ne x npar)

	hk	mineral fraction	Initial sconc
real1	0.123	0.9	12.3
real2	0.321	0.0001	32.1
real3	0.987	0.2	69.0
real N	0.789	0.2345	54.1

model runs = ne (number reals)
(e.g. Monte Carlo)

$\mathbf{D}_{sim}$ (ne x nobs)

	head	sload instream	sconc inaqufier
real1	1.11	0.0001	3.21
real2	2.22	0.003	1.23
real3	3.33	0.0003	9.6
realN	4.44	0.001	1.45

Uncertainty Analysis | Iterative Ensemble Smoother (IES)

$$J \propto \Delta_{D_{sim}}^{T} \Delta_{P}^{-1}$$

$$\Delta_P \propto (P - \bar{P})$$

$$\Delta_{D_{sim}} \propto (D_{sim} - \bar{D}_{sim})$$

P (ne x npar)

	hk	mineral fraction	Initial sconc
real1	0.123	0.9	12.3
real2	0.321	0.0001	32.1
real3	0.987	0.2	69.0
real N	0.789	0.2345	54.1

model runs = ne (number reals)
(e.g. Monte Carlo)

$\mathbf{D_{sim}}$ (ne x nobs)

	head	sload instream	sconc inaqufier
real1	1.11	0.0001	3.21
real2	2.22	0.003	1.23
real3	3.33	0.0003	9.6
realN	4.44	0.001	1.45

Uncertainty Analysis | Iterative Ensemble Smoother (IES)

$$P_\Delta = -((J^T\Sigma_\varepsilon^{-1}J) + (1+\lambda)\Sigma_P^{-1})^{-1}(\Sigma_P^{-1}(P - P_0) + J^T\Sigma_\varepsilon^{-1}(D_{sim} - D_{obs}))$$

upgraded parameter matrix | Approx Hessian | dampening | parameter change matrix | residual vector

$$\Delta_P \propto (P - \bar{P})$$

$$J \propto \Delta_{D_{sim}}^T \Delta_P^{-1}$$

$$\Delta_{D_{sim}} \propto (D_{sim} - \bar{D}_{sim})$$

P (ne x npar)

	hk	mineral fraction	Initial sconc
real1	0.123	0.9	12.3
real2	0.321	0.0001	32.1
real3	0.987	0.2	69.0
real N	0.789	0.2345	54.1

update via GLM
(e.g. maths)

model runs = ne (number reals)
(e.g. Monte Carlo)

D_{sim} (ne x nobs)

	head	sload instream	sconc inaqufier
real1	1.11	0.0001	3.21
real2	2.22	0.003	1.23
real3	3.33	0.0003	9.6
realN	4.44	0.001	1.45

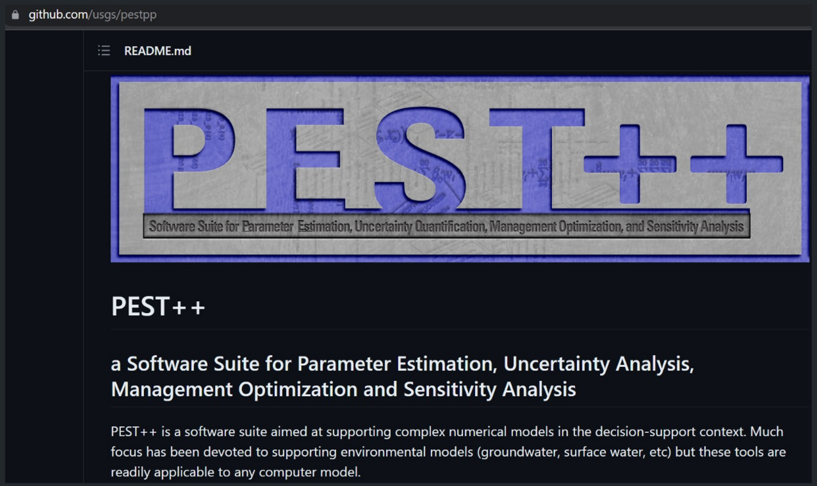
Uncertainty Analysis | Iterative Ensemble Smoother (IES)

- Number of forward model runs required is controlled by the number of realizations.

Uncertainty Analysis | Iterative Ensemble Smoother (IES)

- Number of forward model runs required is controlled by the number of realizations.
- The parameter ensemble is a sample of the posterior parameter distribution which we can use for uncertainty analysis.

Uncertainty Analysis | Framework


github.com/usgs/pestpp

README.md

PEST++

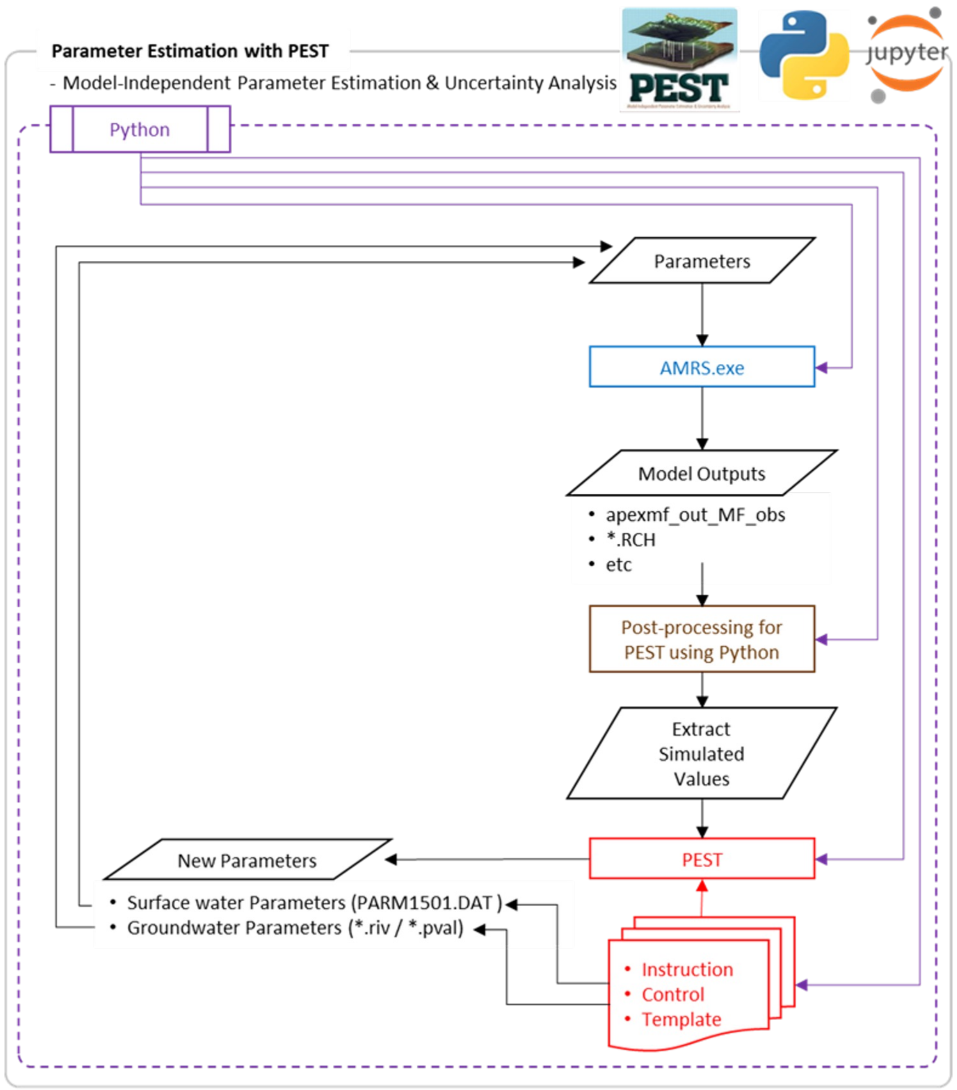
Software Suite for Parameter Estimation, Uncertainty Quantification, Management Optimization, and Sensitivity Analysis

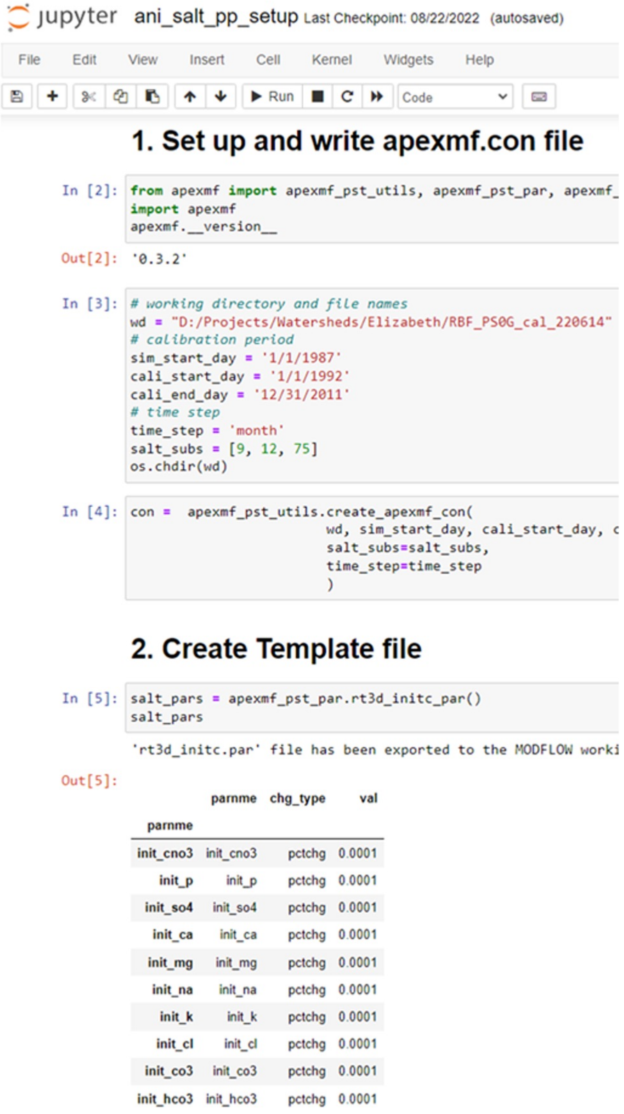
PEST++

a Software Suite for Parameter Estimation, Uncertainty Analysis, Management Optimization and Sensitivity Analysis

PEST++ is a software suite aimed at supporting complex numerical models in the decision-support context. Much focus has been devoted to supporting environmental models (groundwater, surface water, etc) but these tools are readily applicable to any computer model.

pyEMU + FloPy + apexmf




jupyter ani_salt_pp_setup Last Checkpoint: 08/22/2022 (autosaved)

File Edit View Insert Cell Kernel Widgets Help

1. Set up and write apexmf.con file

In [2]:

```
from apexmf import apexmf_pst_utils, apexmf_pst_par, apexmf_
import apexmf
apexmf.__version__
```

Out[2]: '0.3.2'

In [3]:

```
# working directory and file names
wd = "D:/Projects/Watersheds/Elizabeth/RBF_PS0G_cal_220614"
# calibration period
sim_start_day = '1/1/1987'
cali_start_day = '1/1/1992'
cali_end_day = '12/31/2011'
# time step
time_step = 'month'
salt_subs = [9, 12, 75]
os.chdir(wd)
```

In [4]:

```
con =  apexmf_pst_utils.create_apexmf_con(
                    wd, sim_start_day, cali_start_day, c
                    salt_subs=salt_subs,
                    time_step=time_step
                    )
```

2. Create Template file

In [5]:

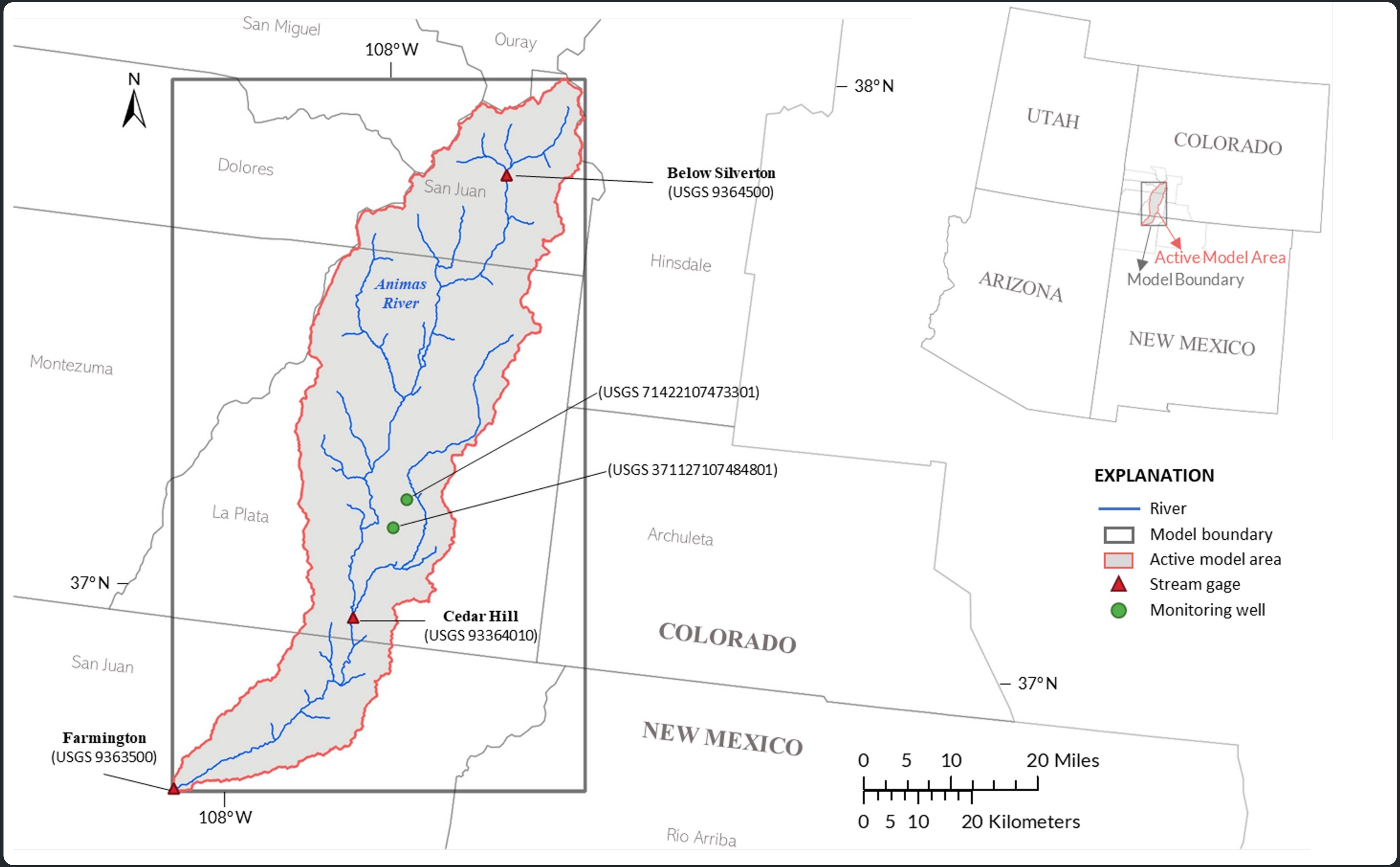
```
salt_pars = apexmf_pst_par.rt3d_initc_par()
salt_pars
```

'rt3d_initc.par' file has been exported to the MODFLOW worki

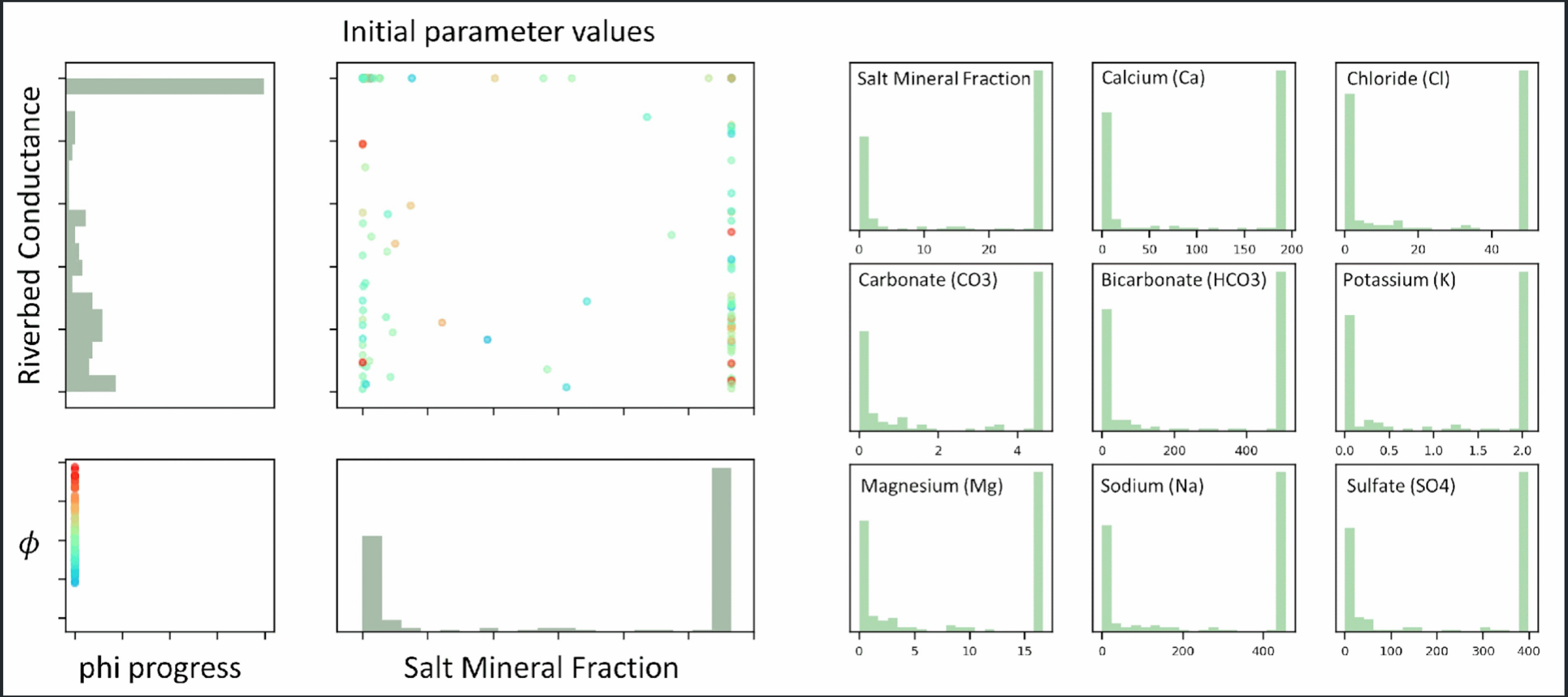
Out[5]:

parnme	parnme	chg_type	val
init_cno3	init_cno3	pctchg	0.0001
init_p	init_p	pctchg	0.0001
init_so4	init_so4	pctchg	0.0001
init_ca	init_ca	pctchg	0.0001
init_mg	init_mg	pctchg	0.0001
init_na	init_na	pctchg	0.0001
init_k	init_k	pctchg	0.0001
init_cl	init_cl	pctchg	0.0001
init_co3	init_co3	pctchg	0.0001
init_hco3	init_hco3	pctchg	0.0001

A Worked Example | Animas River Watershed

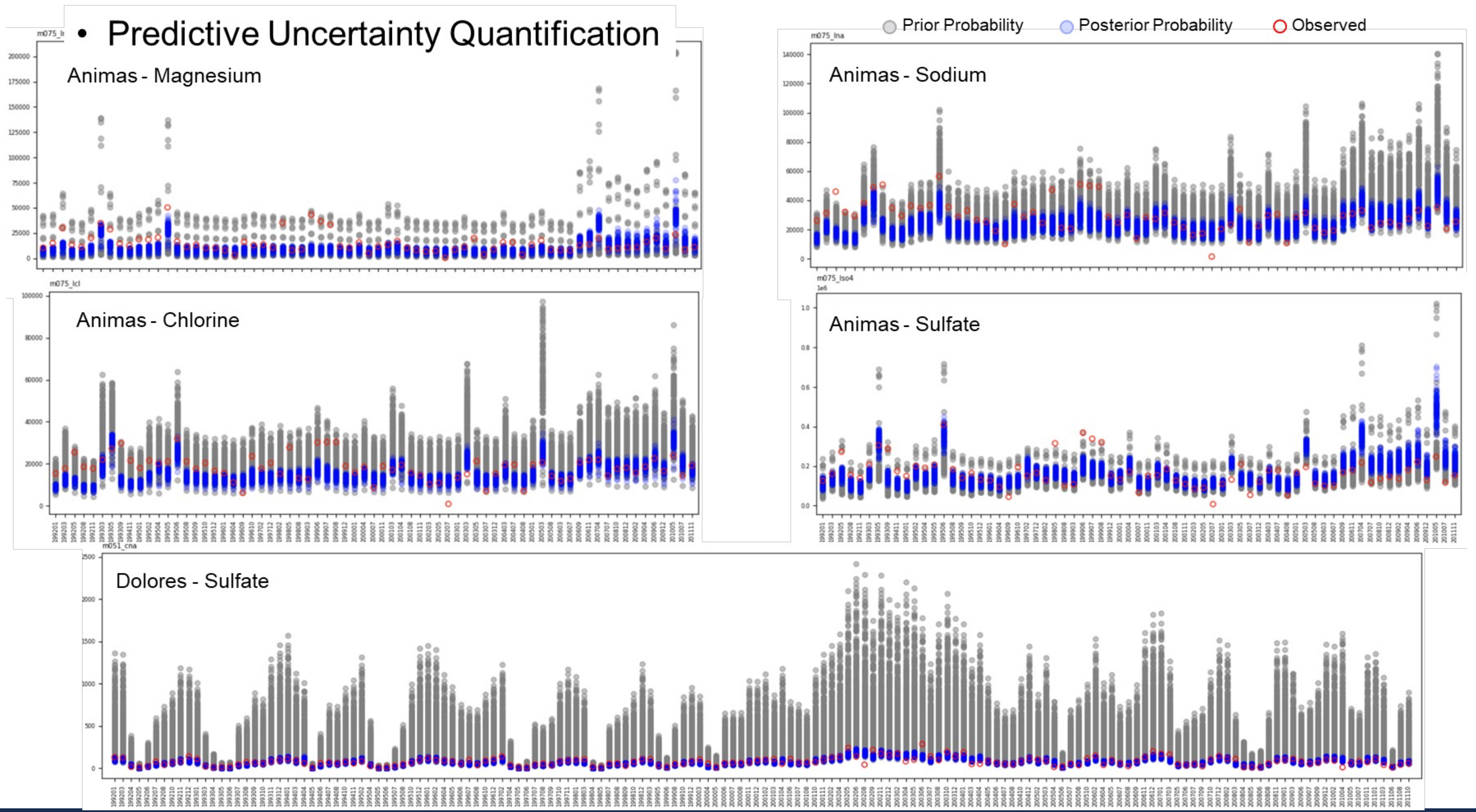


A Worked Example | Animas River Watershed



A Worked Example | Animas River Watershed

- Predictive Uncertainty Quantification



References

- Algorithm Guide:
  - SPOTPY: https://spotpy.readthedocs.io/en/latest/Algorithm_guide/
- Bayesian Theorem:
  - Probabilistic-Programming-and-Bayesian-Methods-for-Hackers: https://github.com/CamDavidsonPilon/Probabilistic-Programming-and-Bayesian-Methods-for-Hackers
  - Bayesian Modeling and Computation in Python: https://bayesiancomputationbook.com/welcome.html
  - PyMC: https://www.pymc.io/welcome.html
- DREAM:
  - Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environmental Modelling & Software 75, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

References

- DREAM:

    - Tasdighi, A., Arabi, M., Harmel, D., Line, D., 2018. A Bayesian total uncertainty analysis framework for assessment of management practices using watershed models. Environmental Modelling & Software 108, 240–252. https://doi.org/10.1016/j.envsoft.2018.08.006
    - Sadegh, M., Vrugt, J.A., 2014. Approximate Bayesian Computation using Markov Chain Monte Carlo simulation: DREAM(ABC). Water Resources Research 50, 6767–6787. https://doi.org/10.1002/2014WR015386
    - Laloy, E., Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM $_{(ZS)}$ and high-performance computing: EFFICIENT MCMC FOR HIGH-DIMENSIONAL PROBLEMS. Water Resour. Res. 48. https://doi.org/10.1029/2011WR010608