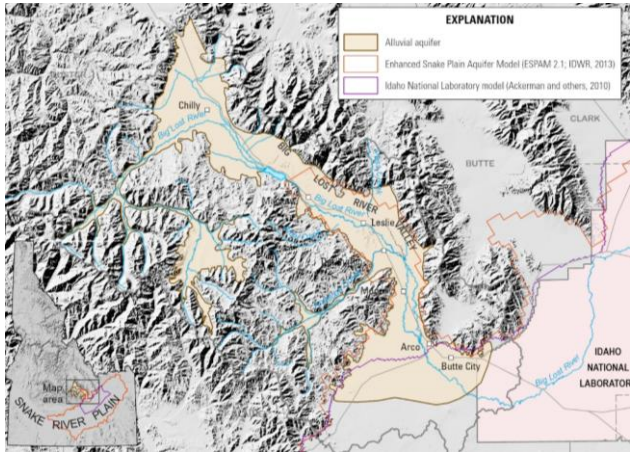
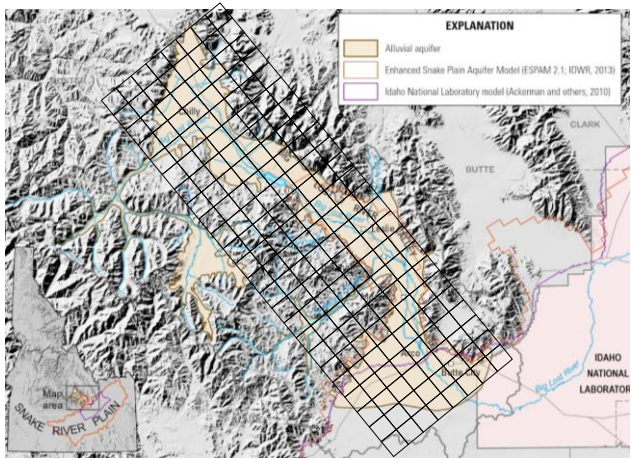




Model Development Process

Stephen Hundt & Jacob Knight
USGS
November 16, 2022

USGS
science for a changing world

Goal of this talk

Explain how we intend to work on the model and how we hope to interact with the MTAC

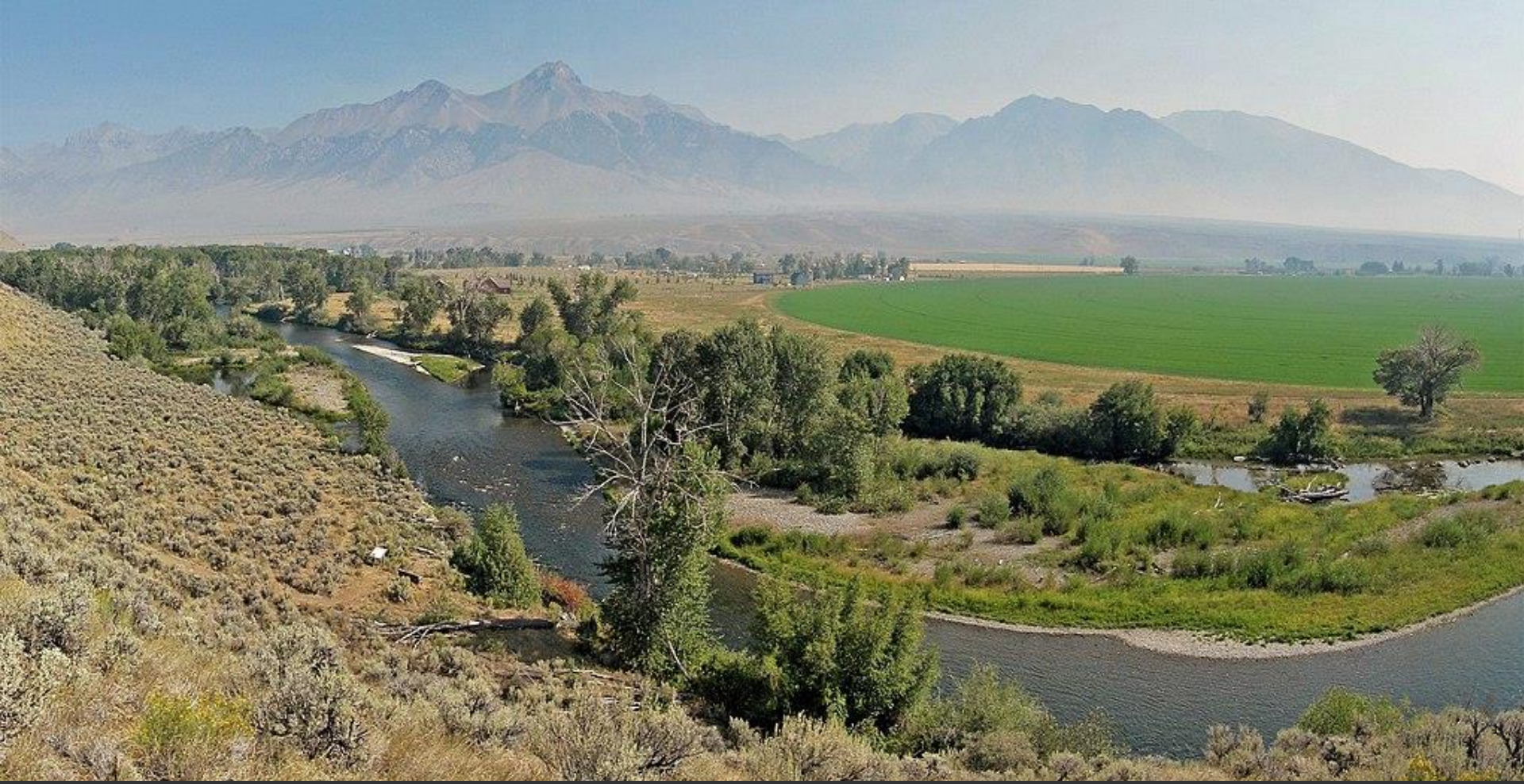
Demonstrate a ‘model status’ presentation that we hope to use in future MTAC meetings

Get your feedback

Contents

USGS science for a changing world

Model Development Approaches



USGS science for a changing world

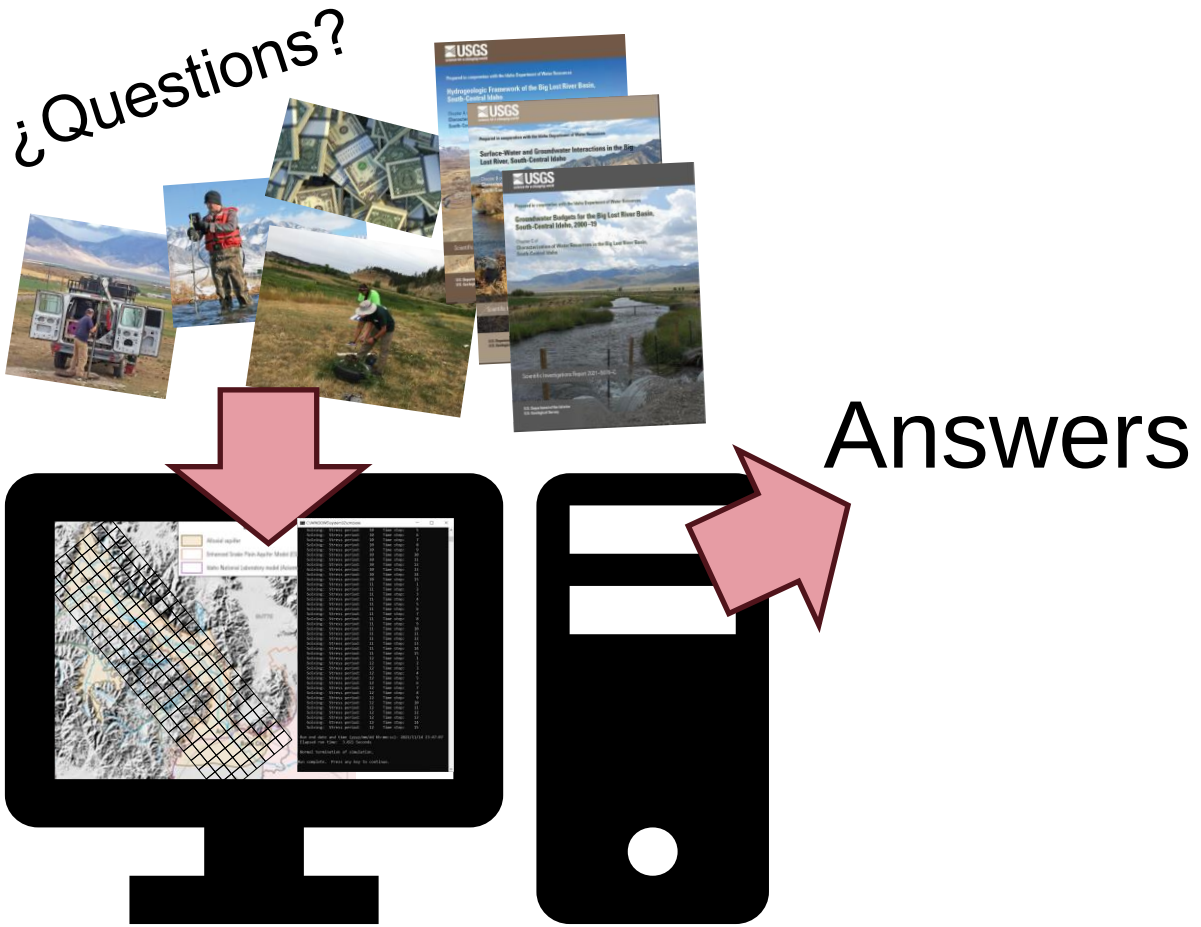
Questions — **Model** → Answers

We want to build a model, but how?

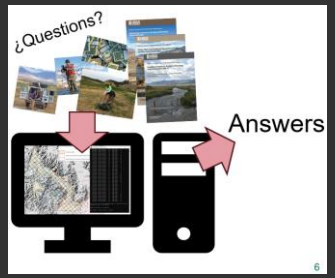
We have a strategy

For now: how you come in because, surprise, you're helping.

Future meetings: hydrology and model details.



The modeling process



Define problem
- Talk to stake-holders
- Literature review
- Preliminary analyses
- Data collection

Develop conceptual model
- Processes/ budget
- Boundary conditions
- Hydro-geologic framework
- Data collection

Develop mathematical model
- Choose model code
- Choose how to represent processes and boundary conditions
- Construct the model

Parameter Estimation
- History matching
- Sensitivity & uncertainty analysis
- Data collection

Assessment of problem using model

Present results

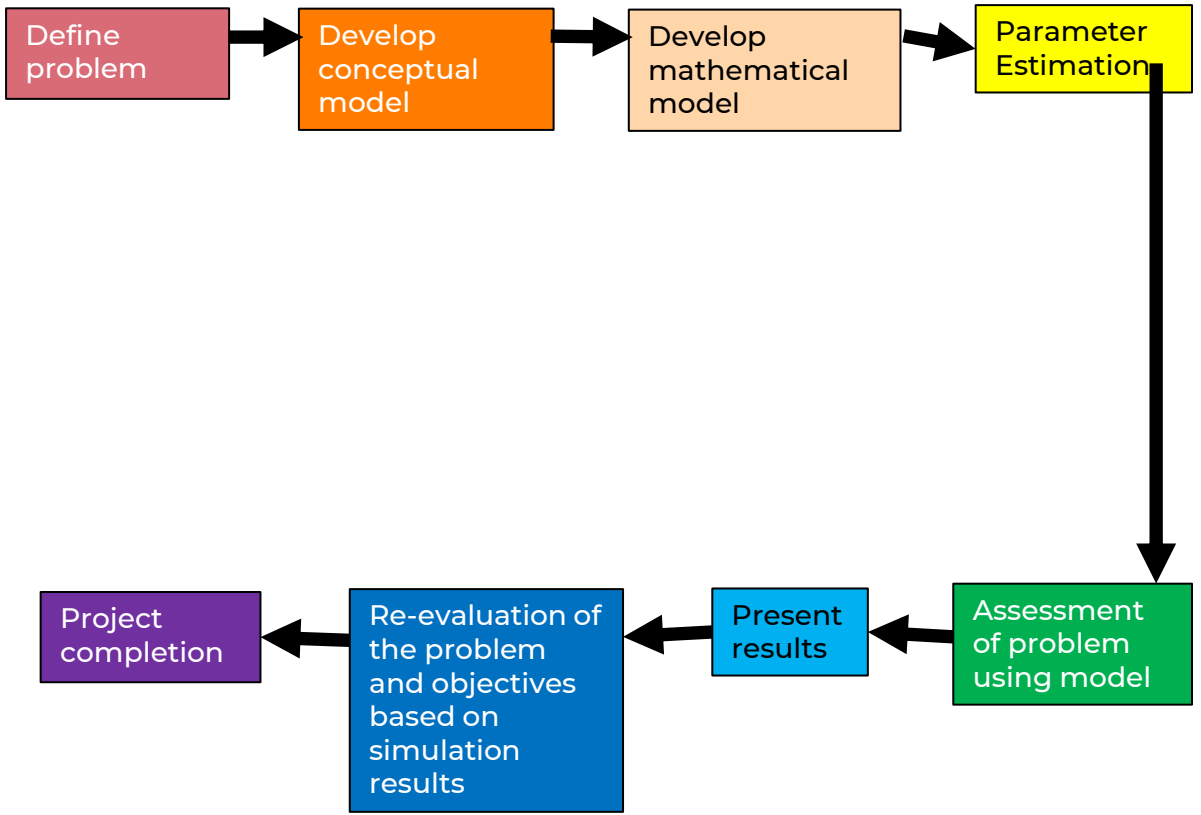
Re-evaluation of the problem and objectives based on simulation results

Project completion

After Reilly (2001) TWRI 3,B8

USGS science for a changing world

Sequential Development



After Reilly (2001) TWRI 3,B8

USGS science for a changing world

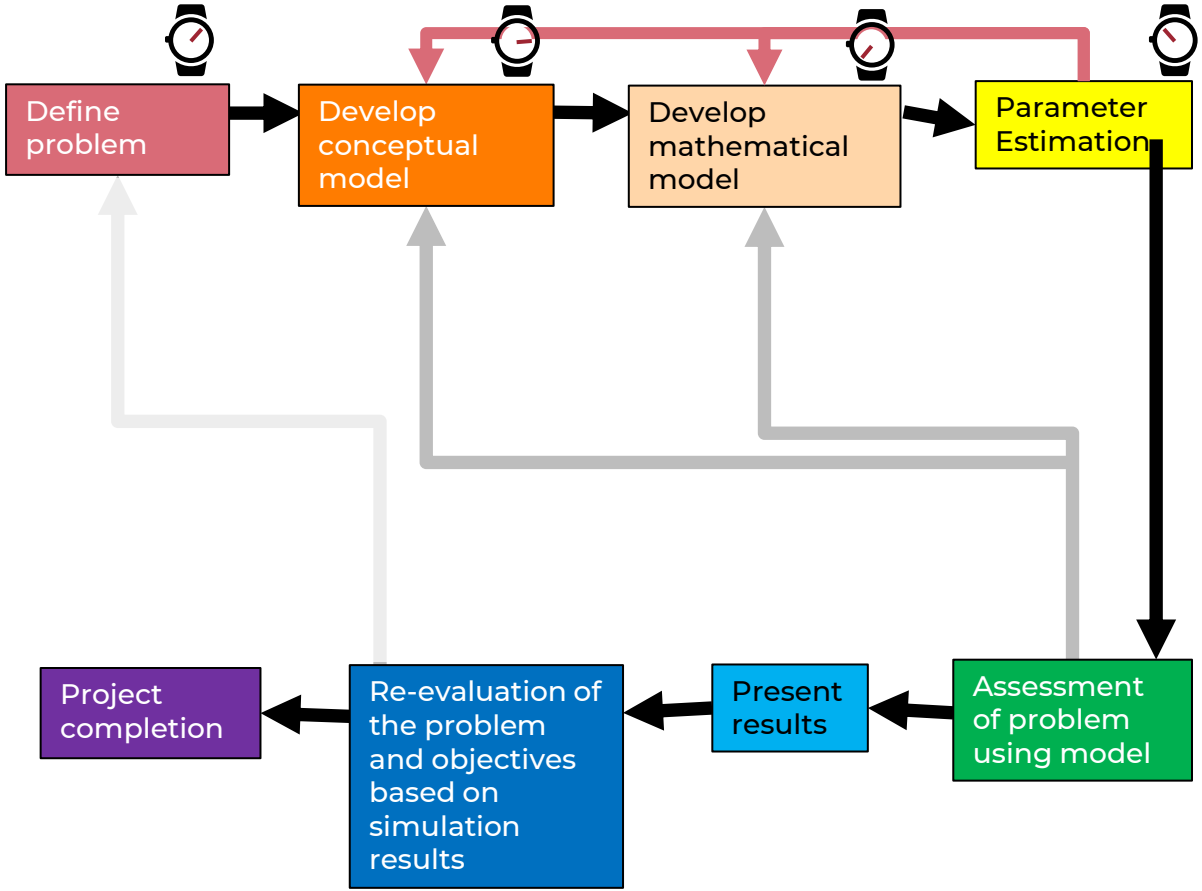
Sorta - Iterative Development

Run out of time to reevaluate problem and adjust conceptual and mathematical model

Difficult to notice and track down errors

Hard to assess the importance of different model features (especially if they don't include parameters for sensitivity analysis)

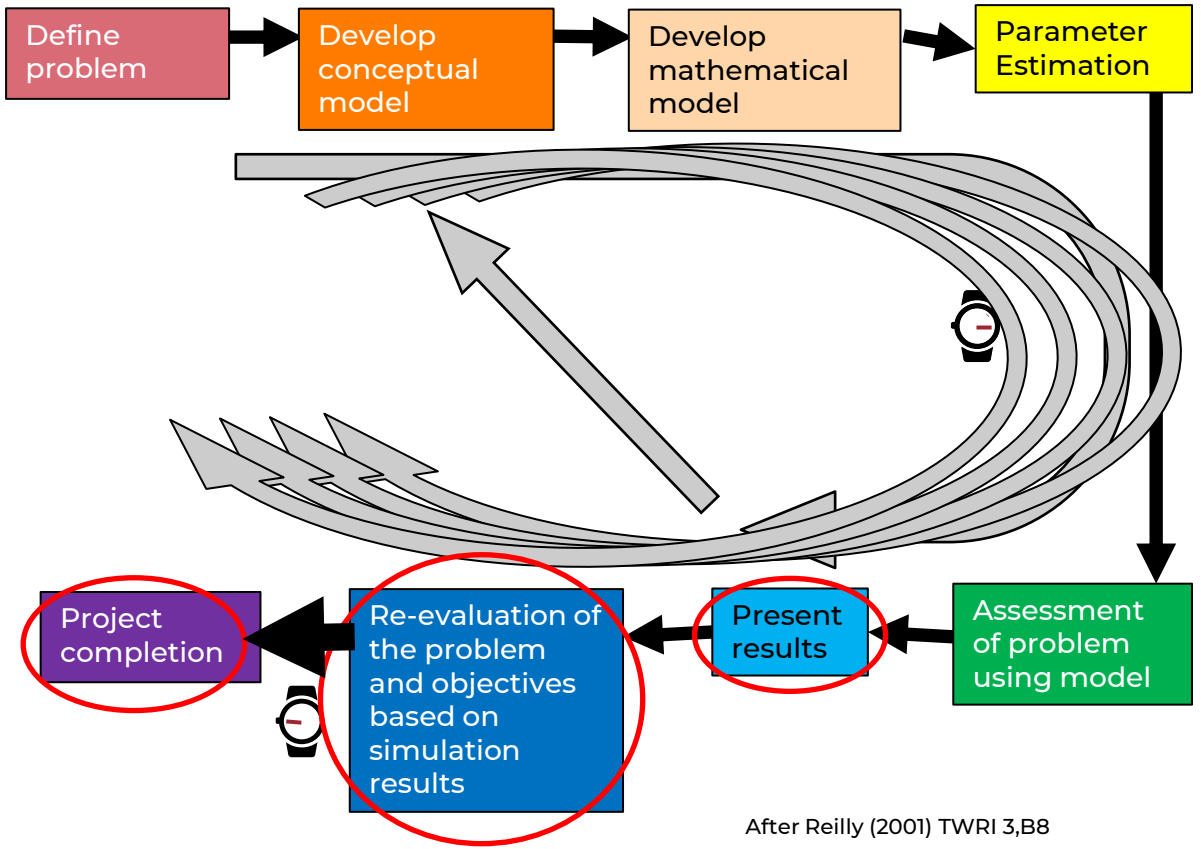
Less familiar with 'results' -> less insightful documentation



After Reilly (2001) TWRI 3,B8

USGS science for a changing world

How to get to truly iterative model development?



After Reilly (2001) TWRI 3,B8

USGS science for a changing world

Our Proposed Approach

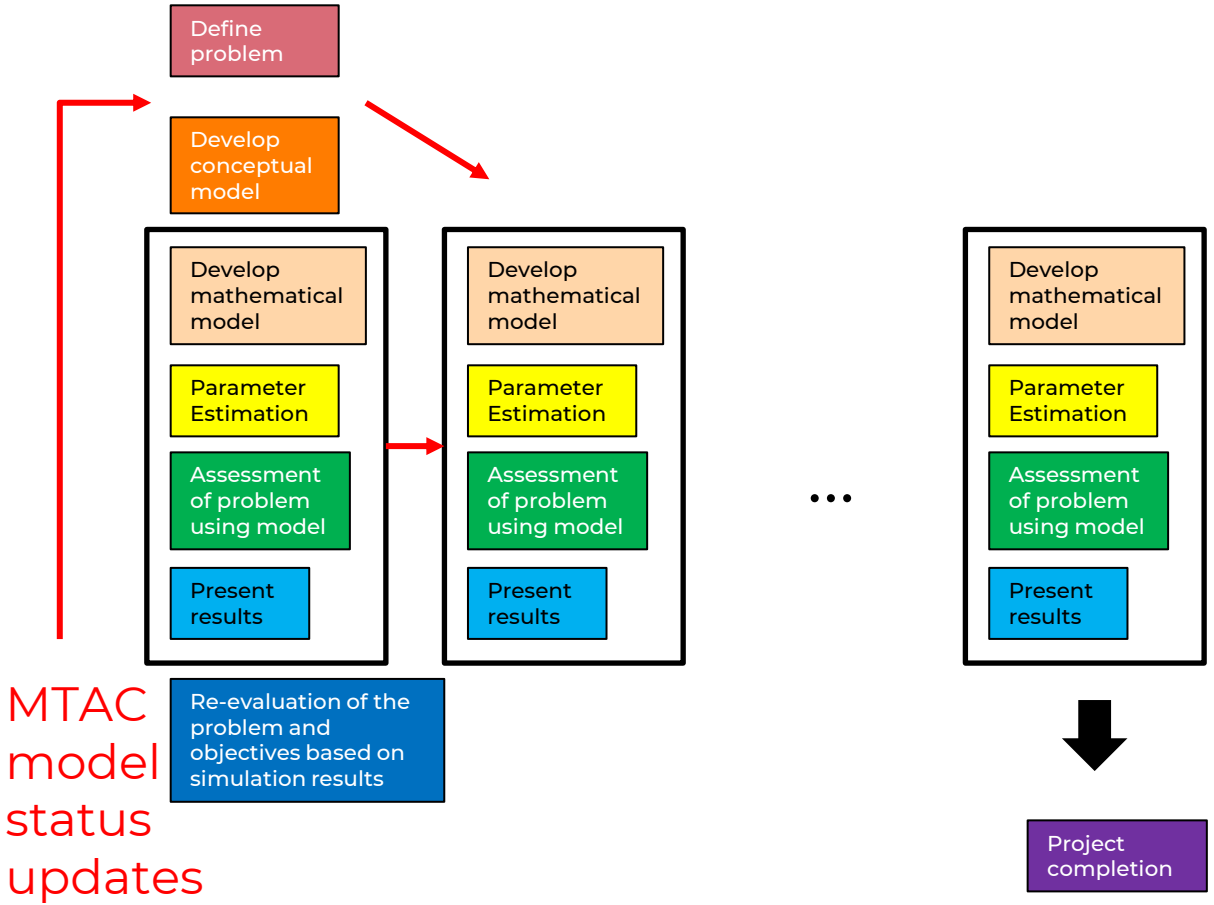
Including how we'll involve the MTAC

Approach in summary

Build scripts that automatically complete all steps from data retrieval through running and plotting scenario output

Get rudimentary version working and incrementally step through working versions from there

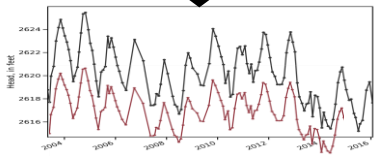
The “results” will include history-matching, parameter estimates, scenario output, and other model metrics **and will be presented at every MTAC**



We'll build fully working versions as we go

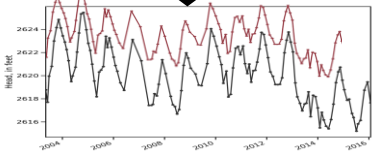
v0.01

- uniform recharge
- Big Lost as drain
- uniform K values



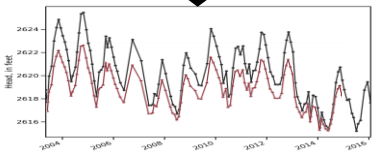
v0.02

- uniform recharge
- Big Lost as drain
- variable K values
- tributary underflow



v0.03

- ag recharge
- Big Lost as RIV
- variable K values
- tributary underflow



...

Communication with MTAC

Show newest working version

Specific results to facilitate feedback

Act on MTAC feedback

Proposed format

- Review objectives
- What you told us last time
- What we've done since last time
- What the model does
- RESULTS
- What the model doesn't do
- What the model should do next

USGS science for a changing world

Case Study and Example Model Status Update

Version 0.0.1:
A pure fantasy starting point

USGS
science for a changing world

Jake's Turn...

USGS
science for a changing world

Reproducible Model Development, Rapid Deployment, and the 'Workmanship of Certainty'

Jake Knight

Hydrologist

Arizona Water Science Center

USGS
science for a changing world

INTRODUCTION

The following presentation contains a worked example of RRR modeling, a sales pitch for scripted model development, and a first attempt at conducting a model review session

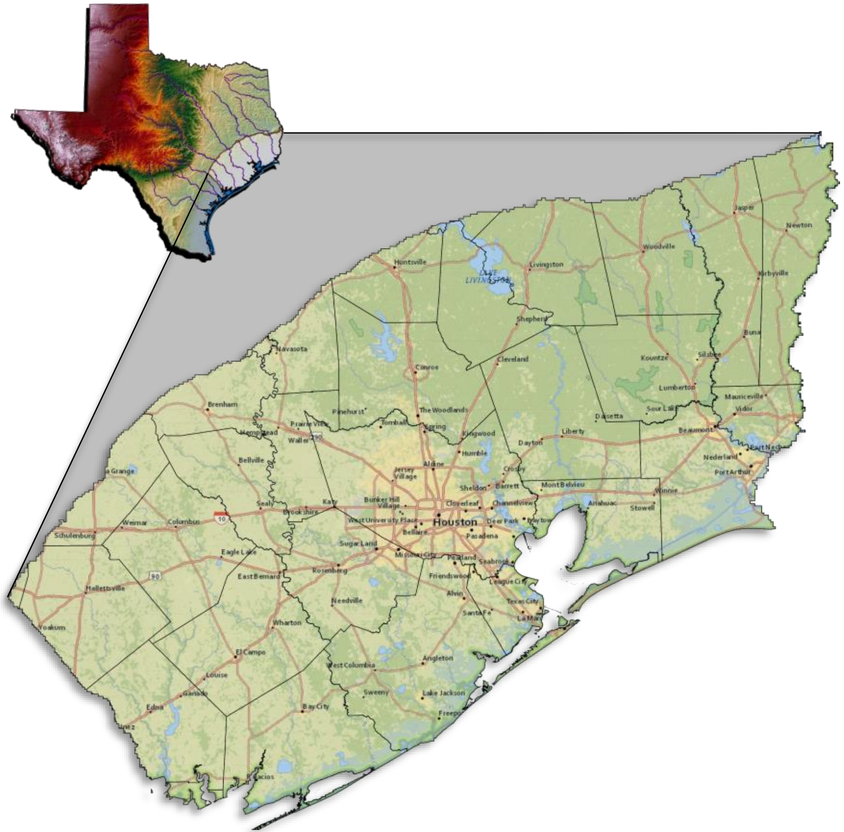
Presentation Outline

- **The GULF model experience**
- **Script-based model development**
- **Getting started with BLRM**

The GULF Model

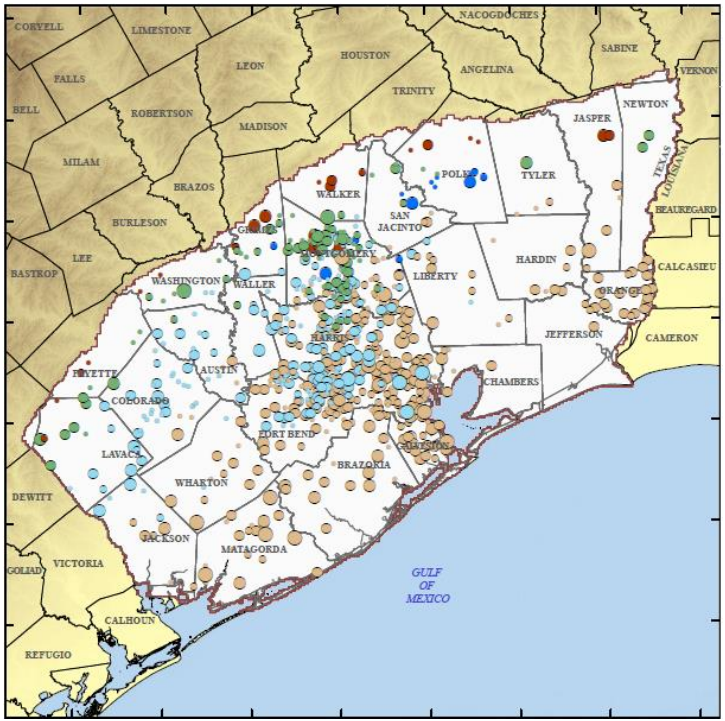
Contentious, Complex, Compressed

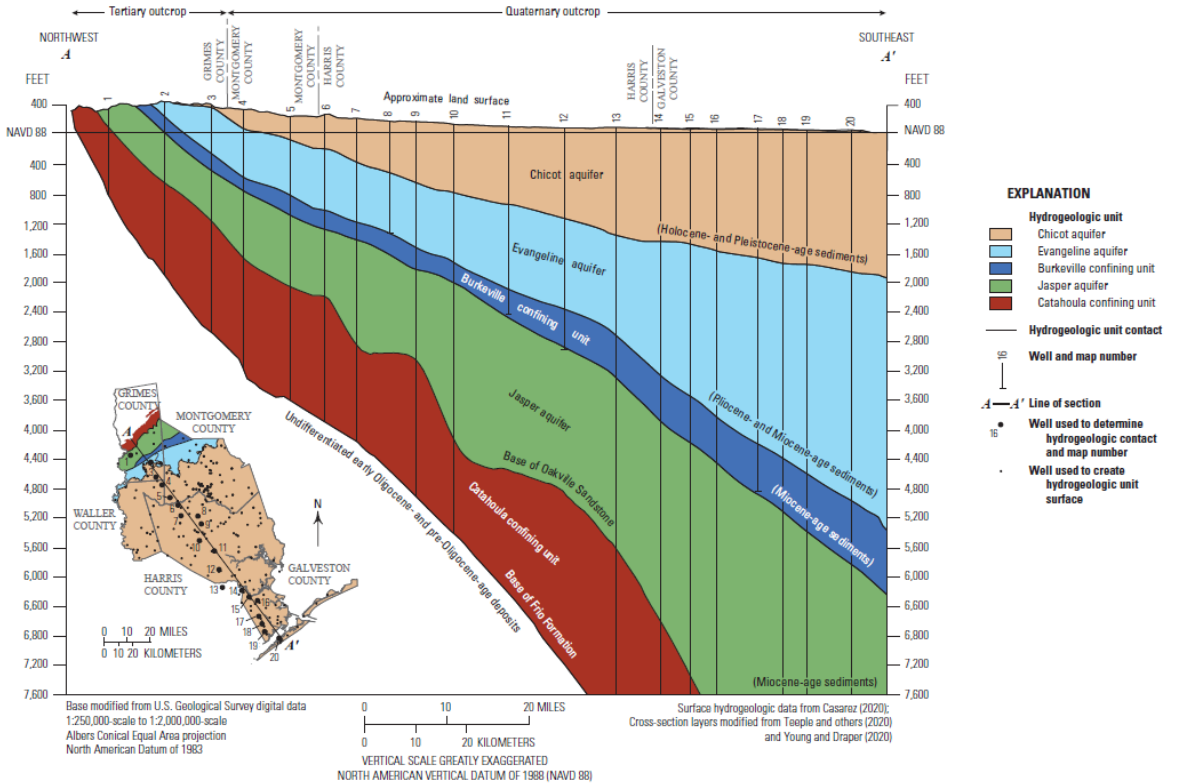
- **Next “Model of Record” for determining allowable groundwater extraction rates in Houston area.**
- **Multi-objective model that must adequately simulate widespread historical groundwater-level declines in multiple aquifer systems, and associated land-surface subsidence**
- **Substantial uncertainty of historical groundwater pumping rates**
- **1st large-scale implementation of CSUB package in MODFLOW 6**
- **Quick turnaround time**



The GULF Model

Observation Data – GW Levels

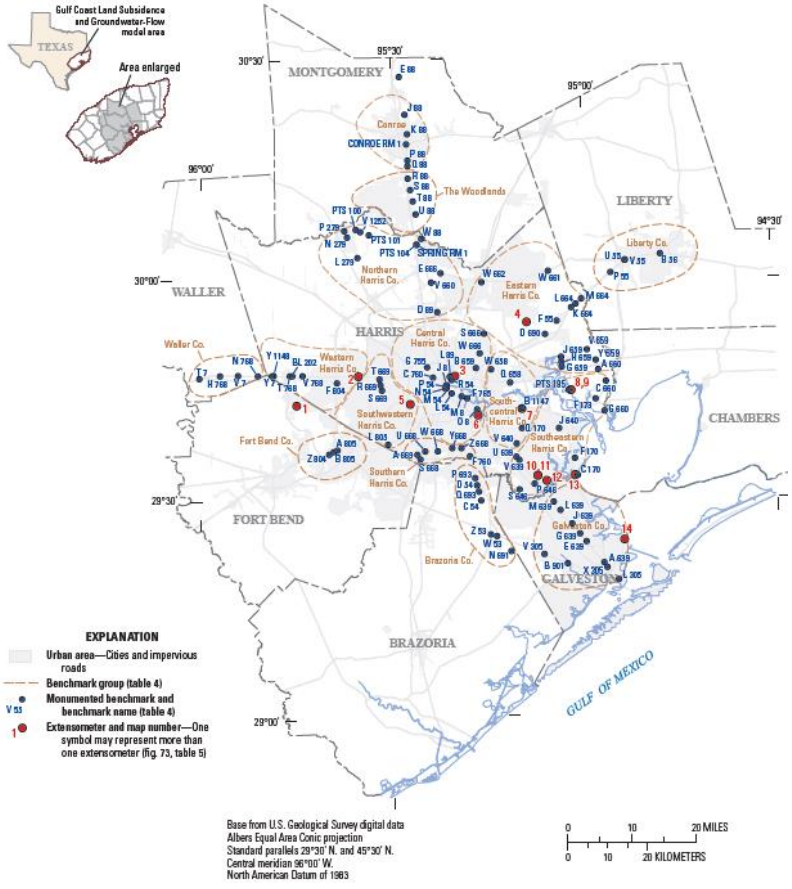


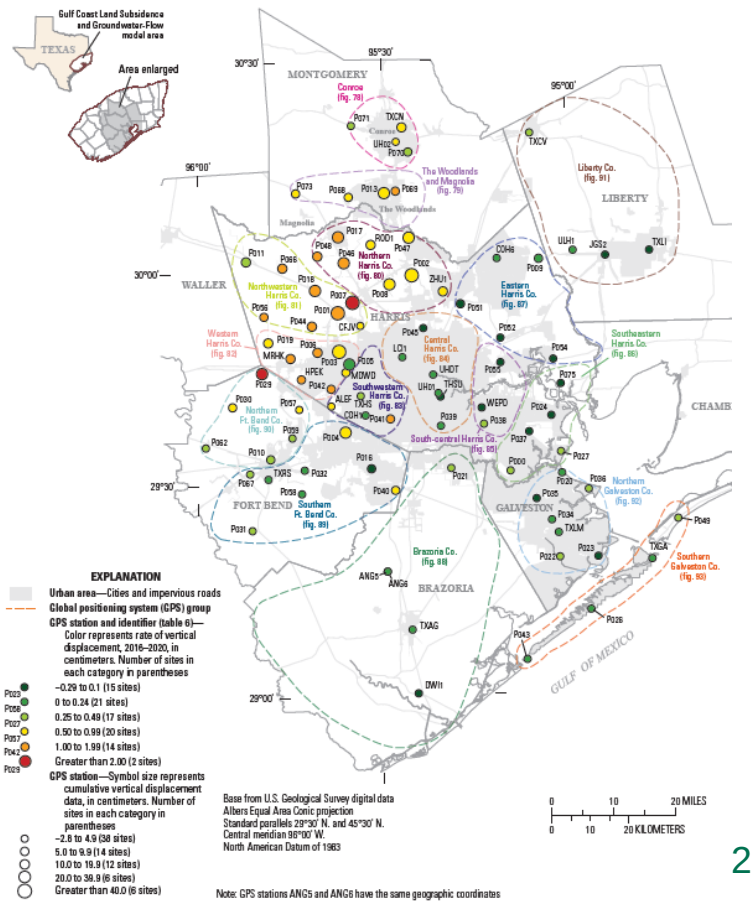


USGS
science for a changing world

The GULF Model

Observation Data - Subsidence





USGS
science for a changing world

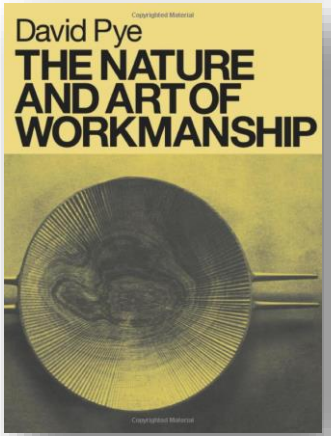
Script-based model development

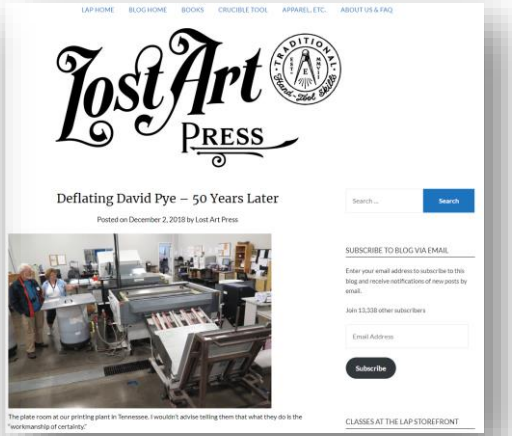
Risk vs. Certainty

From "The Nature and Art of Workmanship" - David Pye (1968)

The most typical and familiar example of the workmanship of risk is writing with a pen, and of the workmanship of certainty, modern printing…

…But all this judgment, dexterity and care has been concentrated and stored up before the actual printing starts. Once it does start, the stored up capital is drawn on and the newspapers come pouring out in an absolutely predetermined form with no possibility of variation between them…

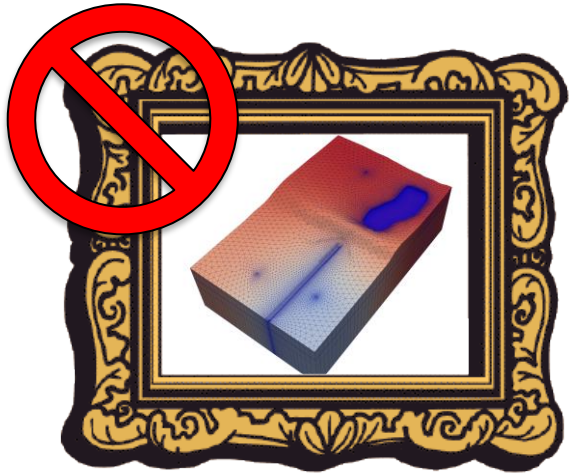




USGS science for a changing world

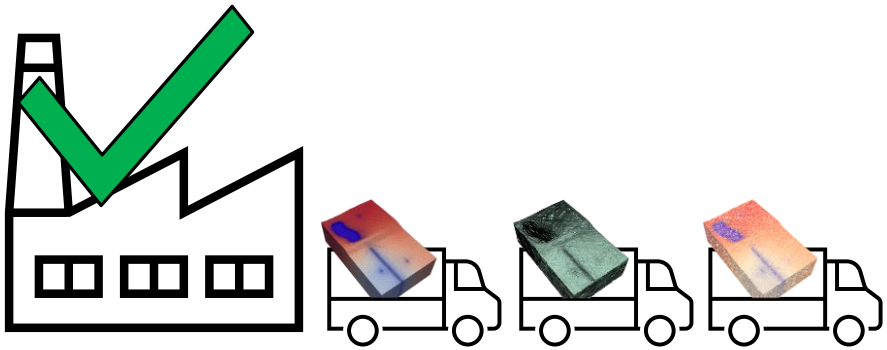
Script-based model development

Don't build a model, build a model factory.



Leverage scripts to:

- **Survive in the realm of "the ubiquity of error"**
- **"Concentrate and store up" your modeling "judgement, dexterity, and care."**
- **Automate input/output generation**
- **Execute *Plan-Do-Check-Act* cycles to move forward in short, quick steps**
- **Maintain flexibility to change decisions**



USGS science for a changing world

The GULF Model

Project Workflow

Plan-Do-Check-Act Cycle (Scrum-“ish”)

- **Build or deploy a model/ensemble variation by modifying and executing scripts build_model.py and/or build_pst.py**

./model_files/build_model.py

```python
if __name__ == "__main__":
    mod_dir = "model_test"

    #process_intera_wel()

    #shape_to_arrays(shp_dir = os.path.join("..","gulf_grid"), shp_f = "gulf_grid_20210602.shp")

    build_history_simulation(mod_name="gulf_history", mod_dir=mod_dir, sp_info="history_model_timestepping.csv",
                             ss_sim=False, sp_short=None, sub=True, build_predpar=True)

    run_history_simulation(mod_name="gulf_history", mod_dir=mod_dir, exe_file='mf6.exe', prev_strt=None)

    run_zonebudget(exe_file='zbud6.exe', mod_dir=mod_dir, mod_name='gulf_history')

    #plot_output(ifpdir=mod_dir, ofpdir='history_plots', mod_name='gulf_history',
    #            sp=19, sp_info="history_model_timestepping.csv", shp_f="gulf_grid_20210324_active_area.shp", sub=True)

    #grab_strt(mod_name="gulf_history", mod_dir="model_base", nlay=6, ofpdir="strt_heads")

    #shutil.copy2("build_model.py",os.path.join(mod_dir,"build_model_{0}.py".format(mod_dir)))
```

./model_files/build_pst.py

```python
if __name__ == "__main__":

    org_dir = "model_base_gm"
    pri_dir = "model_prior_gm"
    ies_dir = "model_ies_gm"
    noptmax = 3


    # CHOOSE LOCAL TEST OR SERVER RUN
    if multiprocessing.cpu_count() < 13:
        local = True
        nreal = 10
    else:
        local = False
        nreal = 600

    # # RUN MODEL
    # run_history_simulation(mod_name="gulf_history", mod_dir=org_dir, exe_file=exe_name)

    # # BUILD PEST INTERFACE

    # setup_control_file(mod_name="gulf_history", org_ws=org_dir, template_ws="model_template", quick=False, sub=True, ies_num_reals=nreal)

    # # ASSIGN STANDARD DEVIATION TO OBS
    # assign_stdev_obs(mod_name="gulf_history", m_d="model_template", pst_name="gulf_history.pst", draw_ensemble=False, num_reals=nreal)

    # # FREEZE SOME PARS
    # freeze_pars(mod_name="gulf_history", m_d="model_template", pst_name="gulf_history.pst")

    # # RUN BASE MODEL
    # run_pst_model(mod_name="gulf_history", template_ws="model_template")

    # # PRIOR MONTE CARLO
    # if os.path.exists(pri_dir):
    #     shutil.rmtree(pri_dir)

    # prep_for_parallel(b_d="model_template", pst_name="gulf_history.pst",
    #                   m_d=pri_dir, mod_name="gulf_history", noptmax=-1, ies_num_reals=nreal, overdue_giveup_fac=1.5, overdue_giveup_time=480.0, make_zip=True, sen=False)

    # # RUN PRIOR
    # if local:
    #     run_local(b_d="model_template_parallel", m_d=pri_dir, pst_name="gulf_history.pst", num_workers=5)
    # else:
    #     run_condor(pst_name="gulf_history.pst", m_d=pri_dir, port=4044, sub_file="gulf_workers.sub")

    # # REWEIGHT BASED ON MEAN RESIDUALS IN PRIOR
    # reweight_obs(mod_name="gulf_history", t_d="model_template", m_d=pri_dir, pst_name="gulf_history.pst", run=True, sub=True)

    # # # APPLY SOME COMMON SENSE LOCALIZATION
    # # build_localizer(t_dir="model_template_reweight", pst_name="gulf_history.pst", mod_name="gulf_history", save=True)

    # # # # APPLY SOME HEAVY HANDED LOCALIZATION
    # # # build_localizer_no_forcing(t_dir="model_template_reweight", pst_name="gulf_history.pst", mod_name="gulf_history", save=True)

    # # # IES RUN
    # # prep_for_parallel(b_d="model_template_reweight", pst_name="gulf_history.pst",
    # #                   m_d=ies_dir, mod_name="gulf_history", noptmax=noptmax, ies_num_reals=nreal, overdue_giveup_fac=1.5, overdue_giveup_time=480.0, make_zip=True,
    # #                   sen=False, restart=pri_dir, with_loc=True, max_ies_reals=999)

    # prep_for_parallel(b_d="model_template_reweight", pst_name="gulf_history.pst",
    #                   m_d=ies_dir, mod_name="gulf_history", noptmax=noptmax, ies_num_reals=nreal, overdue_giveup_fac=1.5, overdue_giveup_time=600.0, make_zip=True,
    #                   sen=False, restart=pri_dir, with_loc=False, max_ies_reals=999)

    # # RUN IES
    # if local:
    #     run_local(b_d="model_template_reweight_parallel", m_d=ies_dir, pst_name="gulf_history.pst", num_workers=5)
    # else:
    #     run_condor(pst_name="gulf_history.pst", m_d=ies_dir, port=4044, sub_file="gulf_workers.sub")

    # CONDITION POSTERIOR
    condition_posterior(pst_name="gulf_history.pst", m_d=ies_dir, noptmax=1, threshold=0.95)
    # # MISCELLANY
    # #initialize_bm_obs(template_ws="model_template")
```

USGS
science for a changing world

The GULF Model

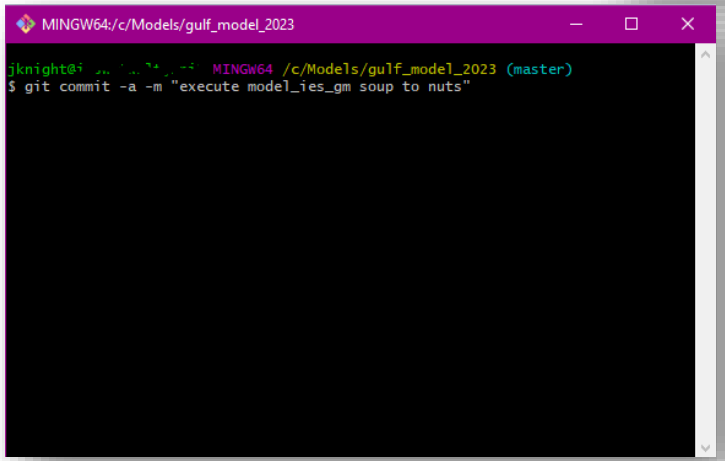
Project Workflow

Plan-Do-Check-Act Cycle (Scrum-“ish”)

- **Build or deploy a model/ensemble variation by modifying and executing scripts build_model.py and/or build_pst.py**
- **Document changes and expectations.**
- **Commit changes to GIT repository.**

./model_files/gulf_ies_run_notes.txt

```
model_ies_gm:
estimate coarse grain properties (ske and theta) independently by layer, no more cascading down from top layer
fix bug in build pest that quintuples number of clay_frac mfinput obs because ib in par name
last call for parameterhol
SY seems to be maxing out at 0.20?
decrease ssv ult lower bound from 5e-5 to 1e-5

model_base_bm5:
stop tracking sk at submap obs locs



model_ies_bm4 (bm4a after correcting weight scheme bug:
new pest exe 5.1.0
weighted cmp obs > 0 for 1943 and 1978 maps
removed bm obs minimum value filter for weighting
theta mult range back to larger distributions of fc2 era, sacrificial par
clay_frac parameterized, used to change thick_frac values in csub pkgdata
rch pars fixed
twel pars fixed
wel and irr grid pars adjustable

RESULTS: pretty good. considering de-weighting burkeville obs, might be messing with updip evg and jsp heads


model_base_bm4:
```



USGS science for a changing world

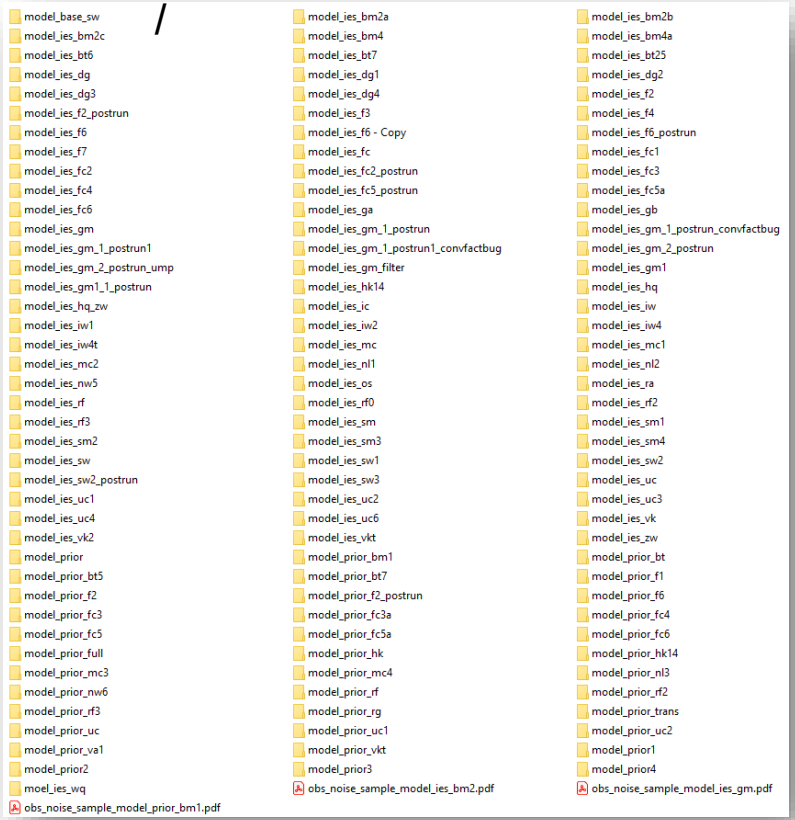
The GULF Model

Project Workflow

Plan-Do-Check-Act Cycle (Scrum-“ish”)

- **Build or deploy a model/ensemble variation by modifying and executing scripts build_model.py and/or build_pst.py**
- **Document changes and expectations.**
- **Commit changes to GIT repository.**
- **Review results plot_results.py**
- **Decide next course of action.**
- **Repeat, Repeat, Repeat, until…?**

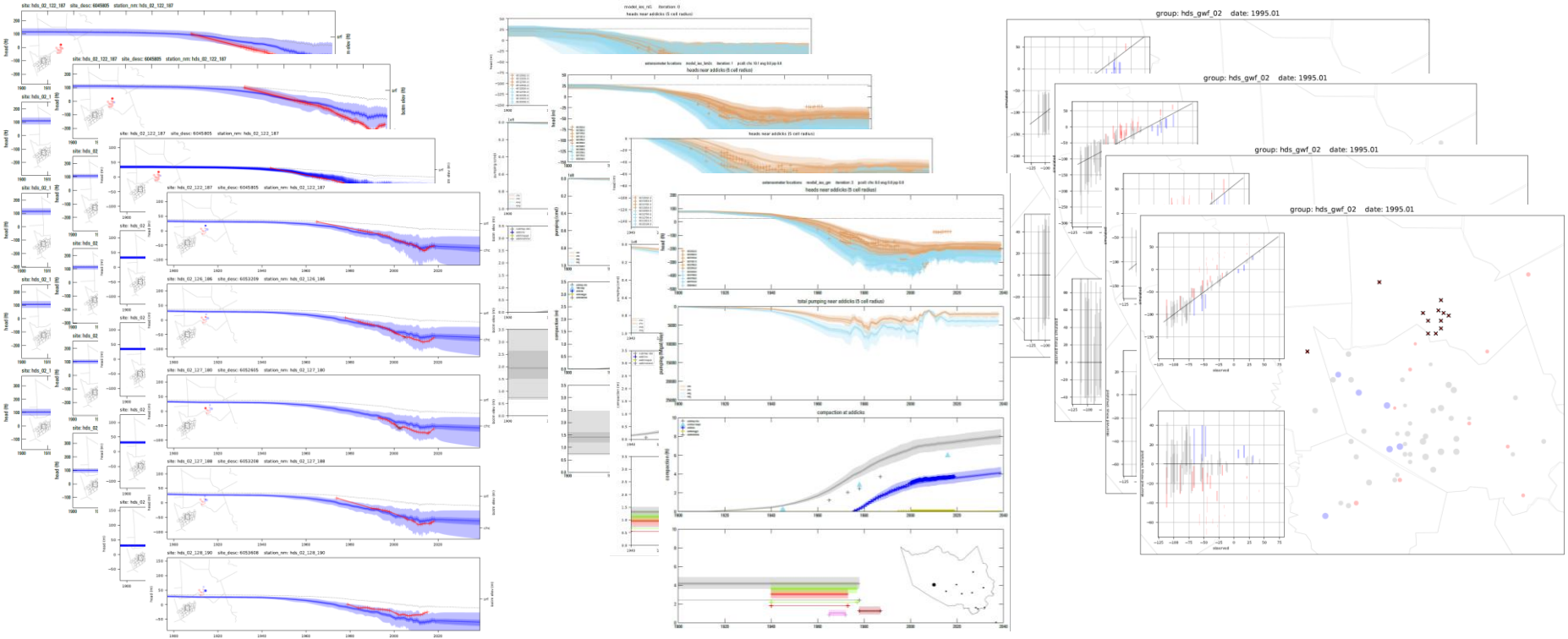
./model_files/model_plots/



USGS science for a changing world

The GULF Model

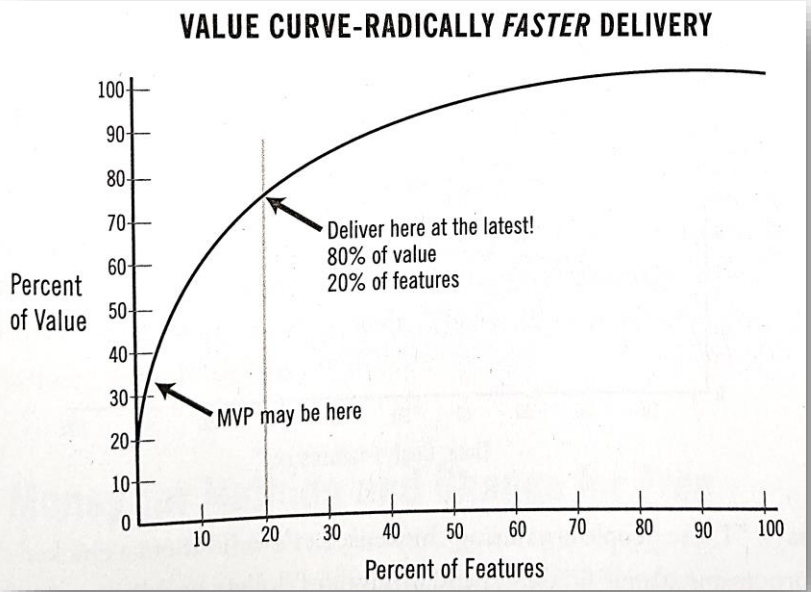
Project Workflow



The BLRM Model

Getting Started

- **“Minimum Viable Product” model running**
  - **Embarrassingly simple, far from complete, but it “works.”**
- **Many useful scripts inherited and adapted from GULF project**
- **Executing PDCA cycles with each added feature**
- **Repeat, Repeat, Repeat, until…?**

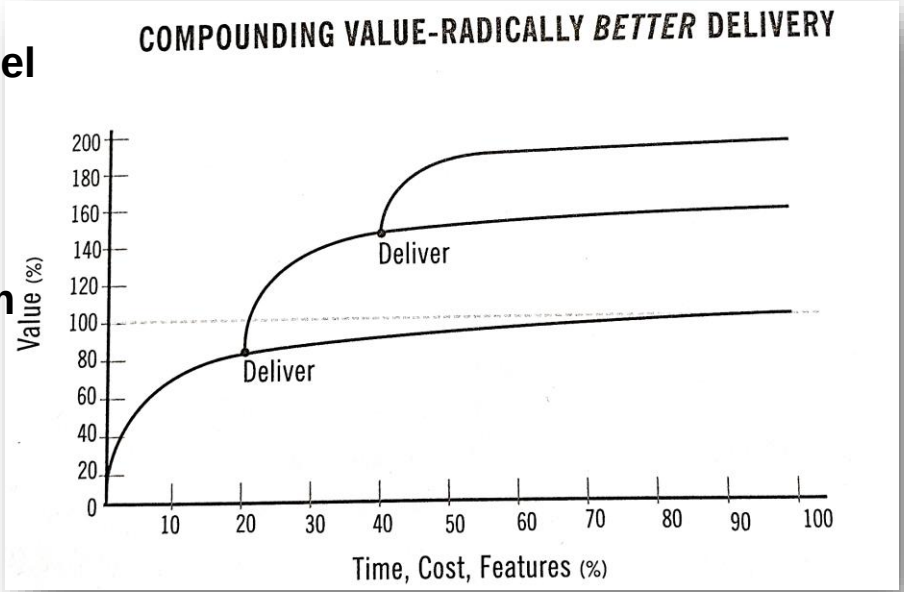


USGS science for a changing world

The BLRM Model

Getting Started

- **Each MTAC meeting is an opportunity for a "product delivery"**
  - **Review added features, improved model performance w.r.t. objectives**
  - **Discuss model shortcomings, challenges**
- **Feedback at each MTAC meeting will inform a refreshed list of target features and abilities to aim for before the next MTAC meeting**
- **Repeat, Repeat, Repeat, until…?**

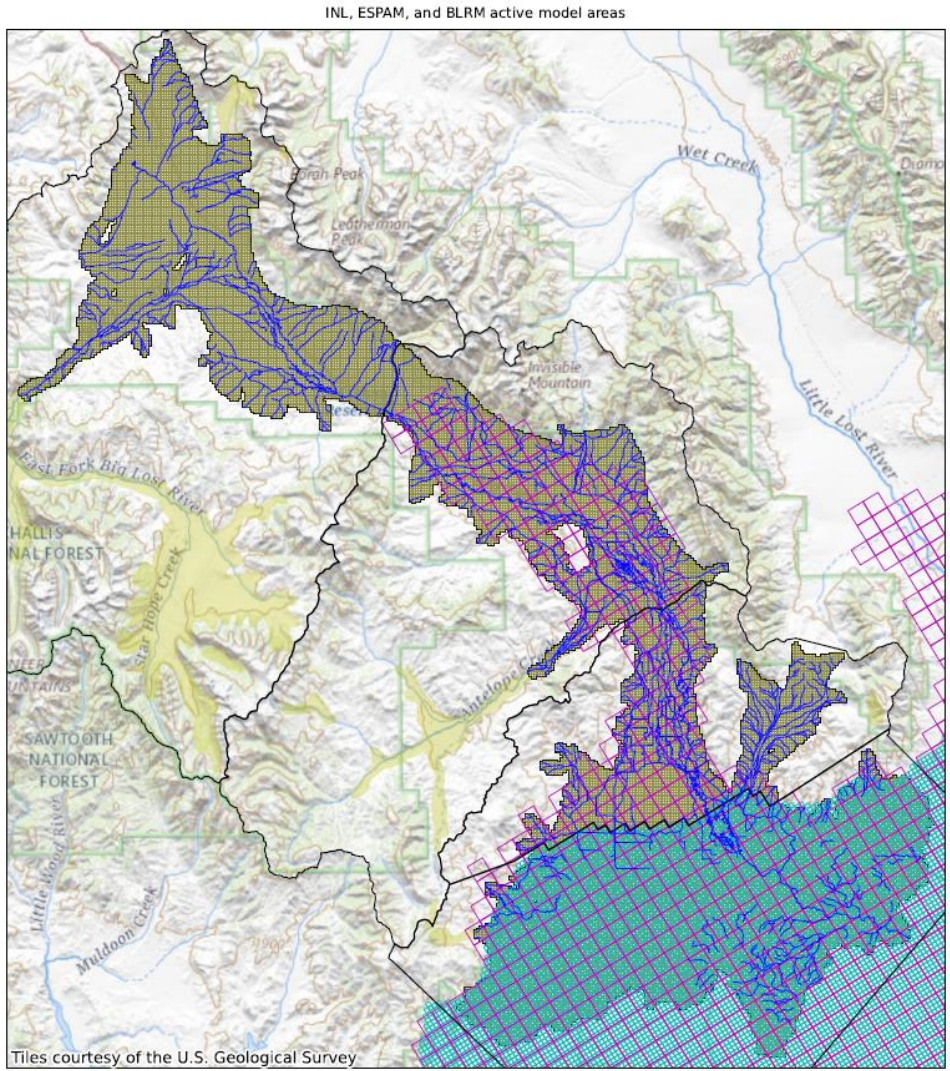


USGS science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Model build from raw data through parameterization and prior MC evaluation**
- **Script-based and fully automated (so far)**
  - **Python + FloPy, PyEmu, mfsetup, sfrmaker**
- **Arbitrary cell size (100m, 200m, 250m tested)**
- **Drain network from NHD+ streamlines**
- **General Head Boundary condition at southern extent set to ESPAM-simulated steady-state heads**
- **Constant RCH rates and HK**
- **Crude basement surface elevation**

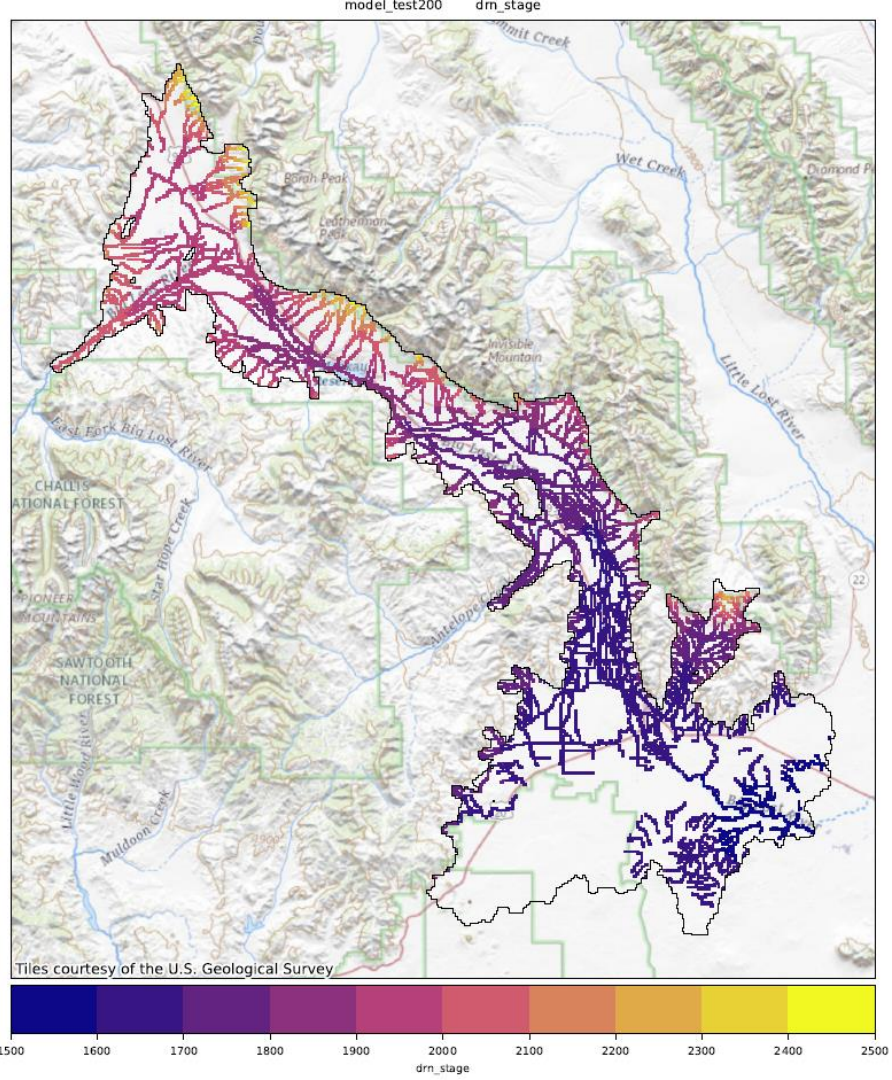


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Model build from raw data through parameterization and prior MC evaluation**
- **Script-based and fully automated (so far)**
  - **Python + FloPy, PyEmu, mfsetup, sfrmaker**
- **Arbitrary cell size (100m, 200m, 250m tested)**
- **Drain network from NHD+ streamlines**
- **General Head Boundary condition at southern extent set to ESPAM-simulated steady-state heads**
- **Constant RCH rates and HK**
- **Crude basement surface elevation**

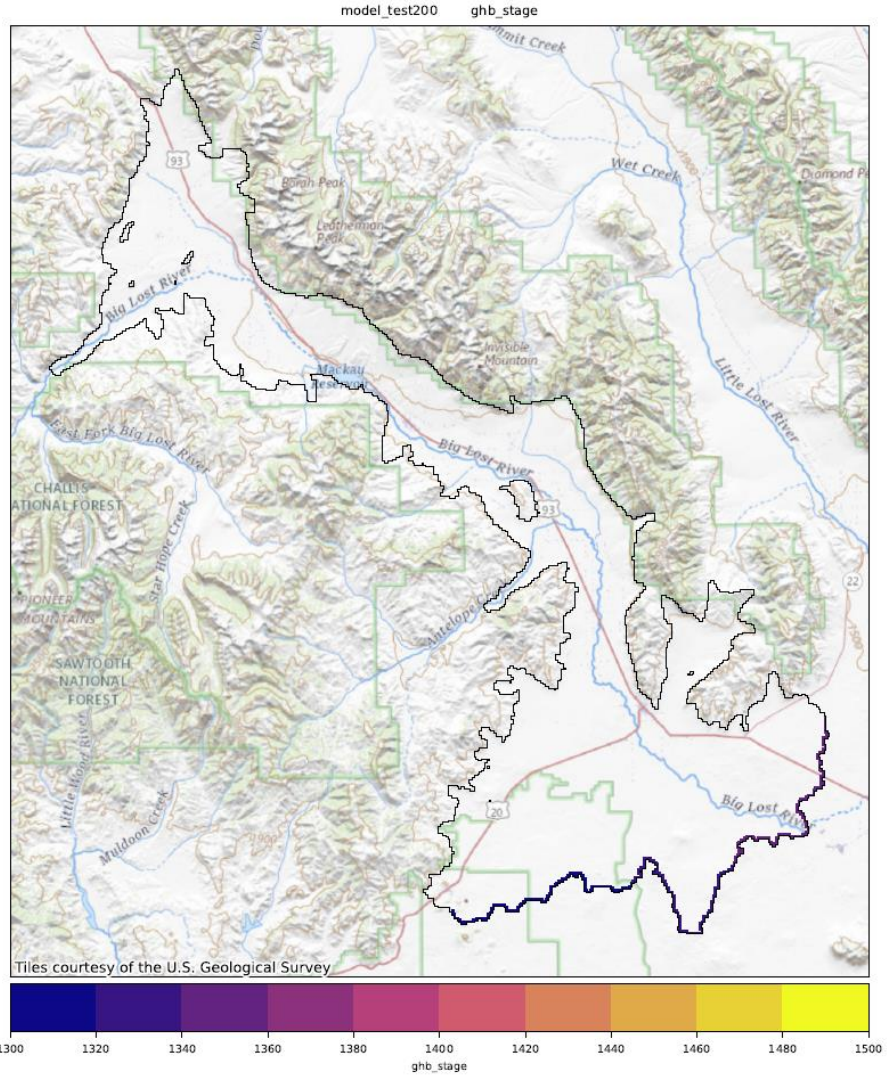


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Model build from raw data through parameterization and prior MC evaluation**
- **Script-based and fully automated (so far)**
  - **Python + FloPy, PyEmu, mfsetup, sfrmaker**
- **Arbitrary cell size (100m, 200m, 250m tested)**
- **Drain network from NHD+ streamlines**
- **General Head Boundary condition at southern extent set to ESPAM-simulated steady-state heads**
- **Constant RCH rates and HK**
- **Crude basement surface elevation**

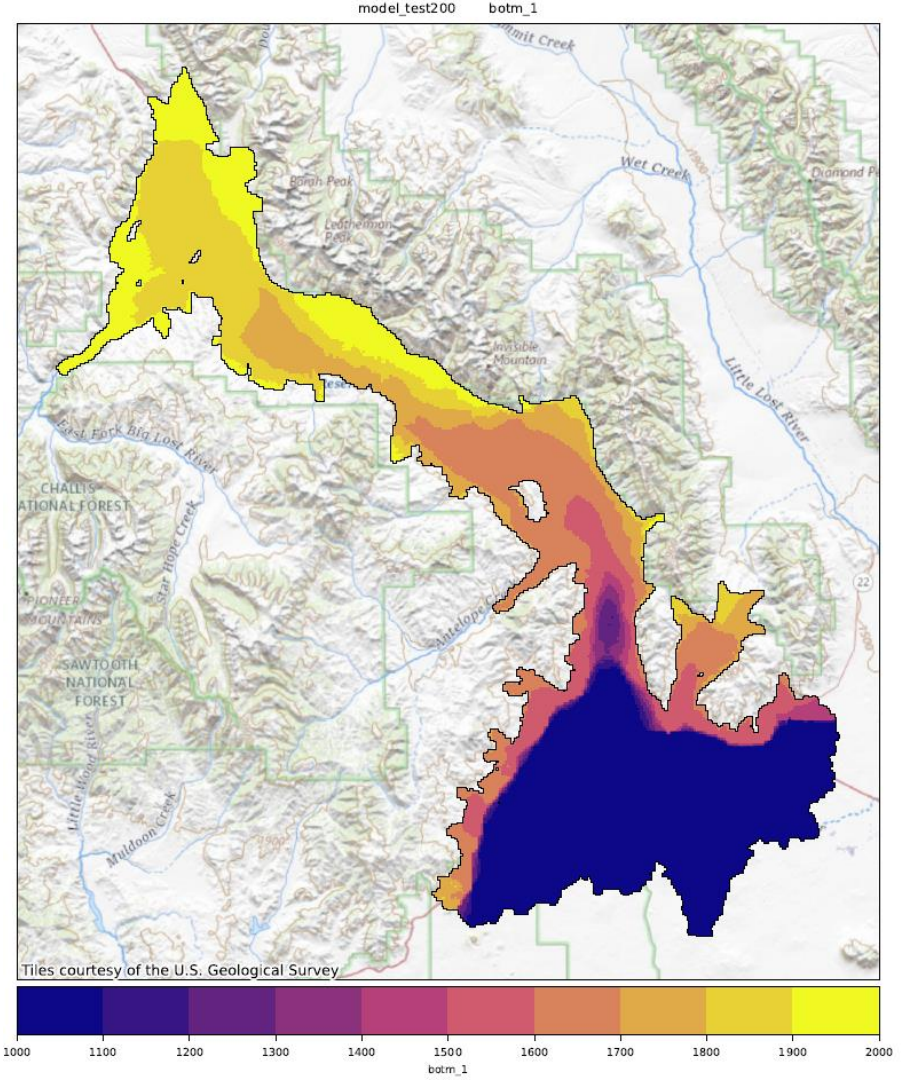


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Model build from raw data through parameterization and prior MC evaluation**
- **Script-based and fully automated (so far)**
  - **Python + FloPy, PyEmu, mfsetup, sfrmaker**
- **Arbitrary cell size (100m, 200m, 250m tested)**
- **Drain network from NHD+ streamlines**
- **General Head Boundary condition at southern extent set to ESPAM-simulated steady-state heads**
- **Constant RCH rates and HK**
- **Crude basement surface elevation**

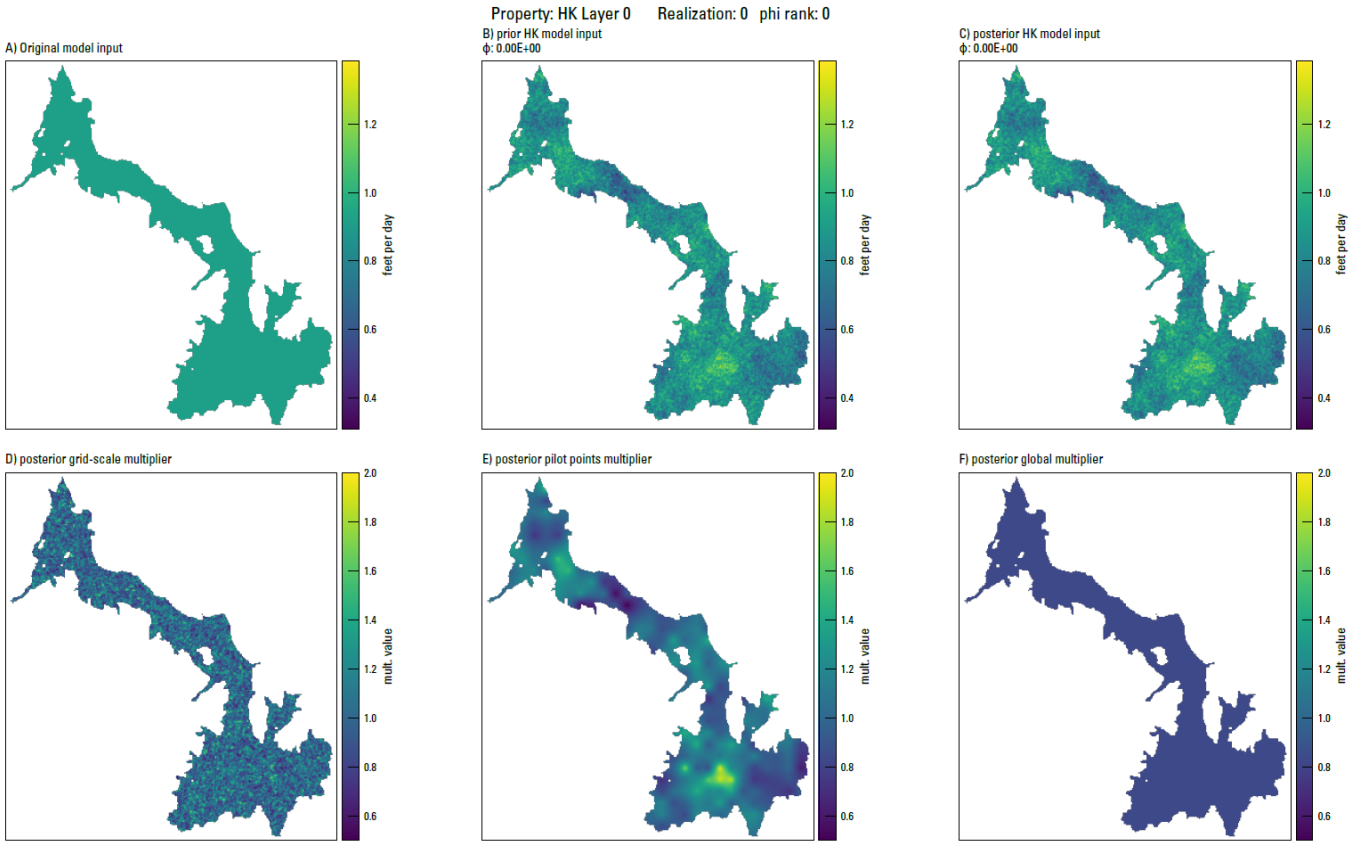


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **HK and RCH parameterization via multipliers deployed at 3 spatial scales:**
  - **Global**
  - **Pilot Point**
  - **Indv. Cell**

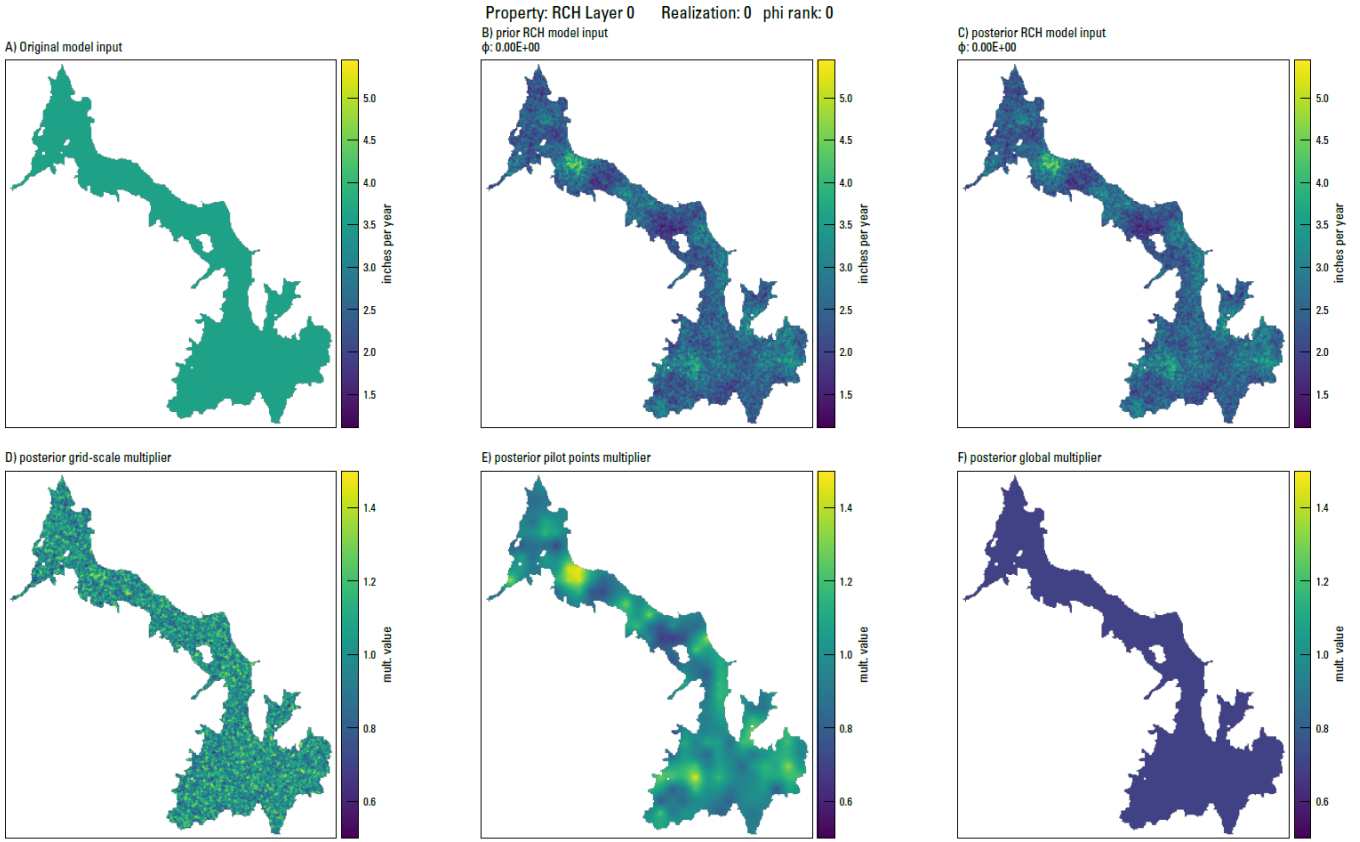


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **HK and RCH parameterization via multipliers deployed at 3 spatial scales:**
  - **Global**
  - **Pilot Point**
  - **Indv. Cell**

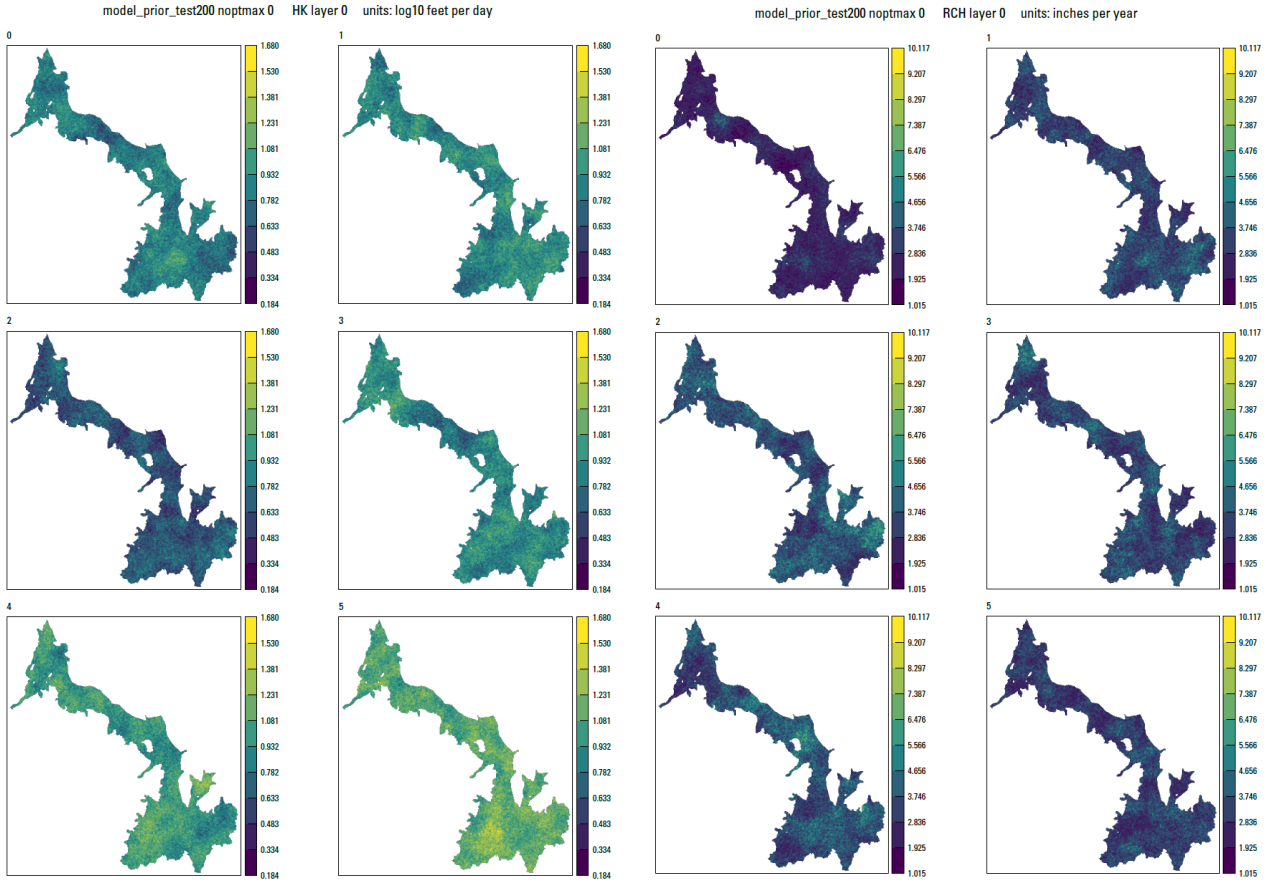


USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Randomly draw values from Prior Parameter Distribution to build Prior ensemble of model realizations**

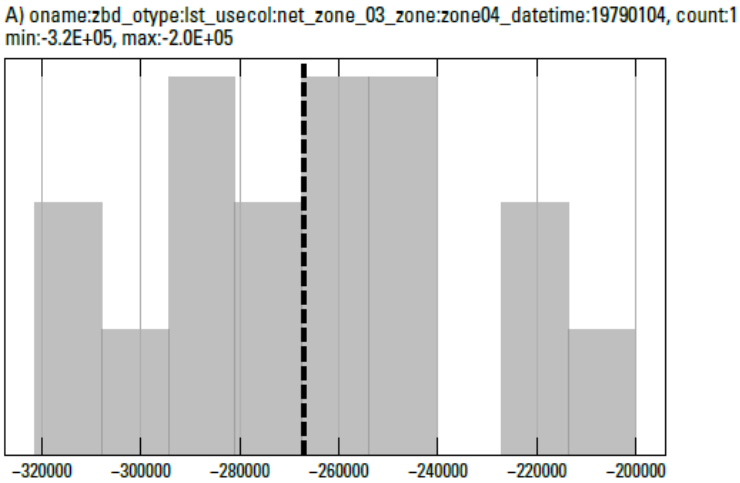


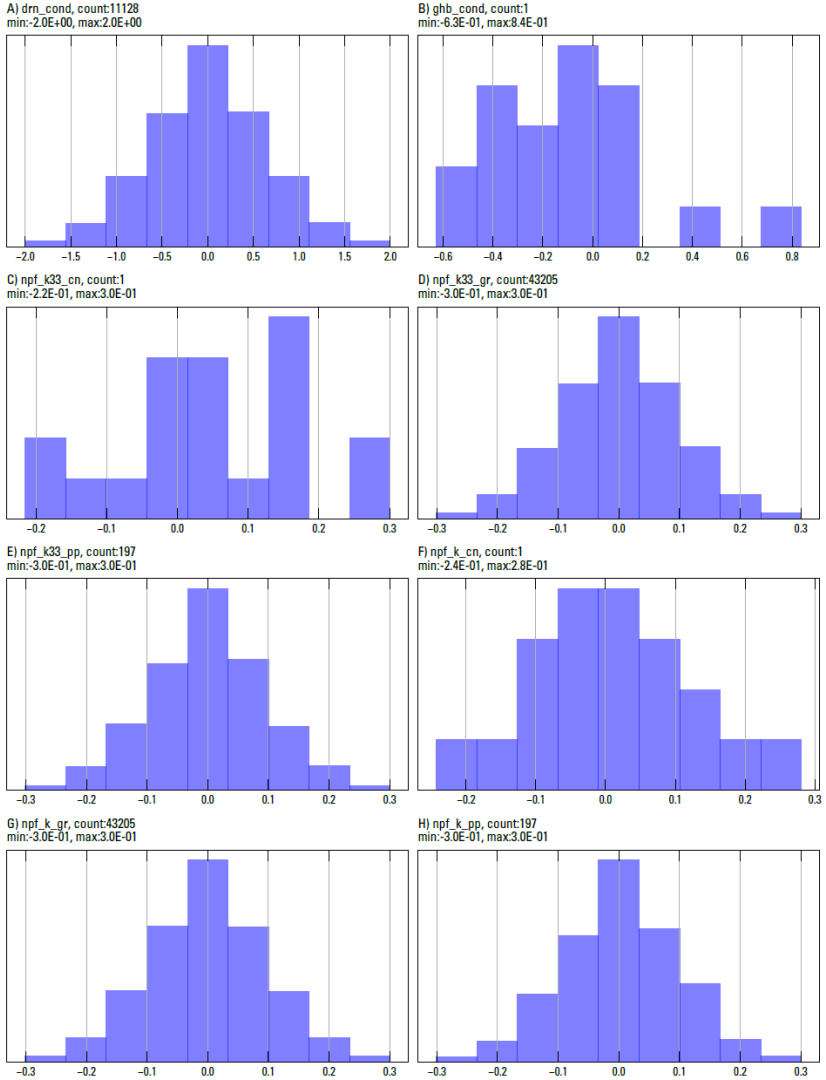
USGS
science for a changing world

The BLRM Model

minimum viable product (MVP) model

- **Prior Monte Carlo evaluates range of outputs resulting from range of inputs**
  - **Currently tracking underflow into ESPAM domain (est. 267k afyr)**





USGS
science for a changing world

Stephen's Turn...

Discussion

Summary and Questions

- How does this communication format sound?
- What should we include in model status update presentations?
  - Figures
  - Topics
- What should the model do next?
- What else do you want to tell us?
- Questions?

Thanks!

Stephen Hundt
shundt@usgs.gov
208-387-1390

Jacob Knight
jknight@usgs.gov
520-670-3336

USGS
science for a changing world