Model Calibration

Stephen Hundt

science for a changing world

Where we are

The modeling process

Define problem
- Literature review
- Preliminary analyses
- Data collection

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

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

Calibration
- History matching
- Sensitivity analysis
- Data collection

Assessment of problem using model

or
le
and objective
based on simulation results

results

Project completion

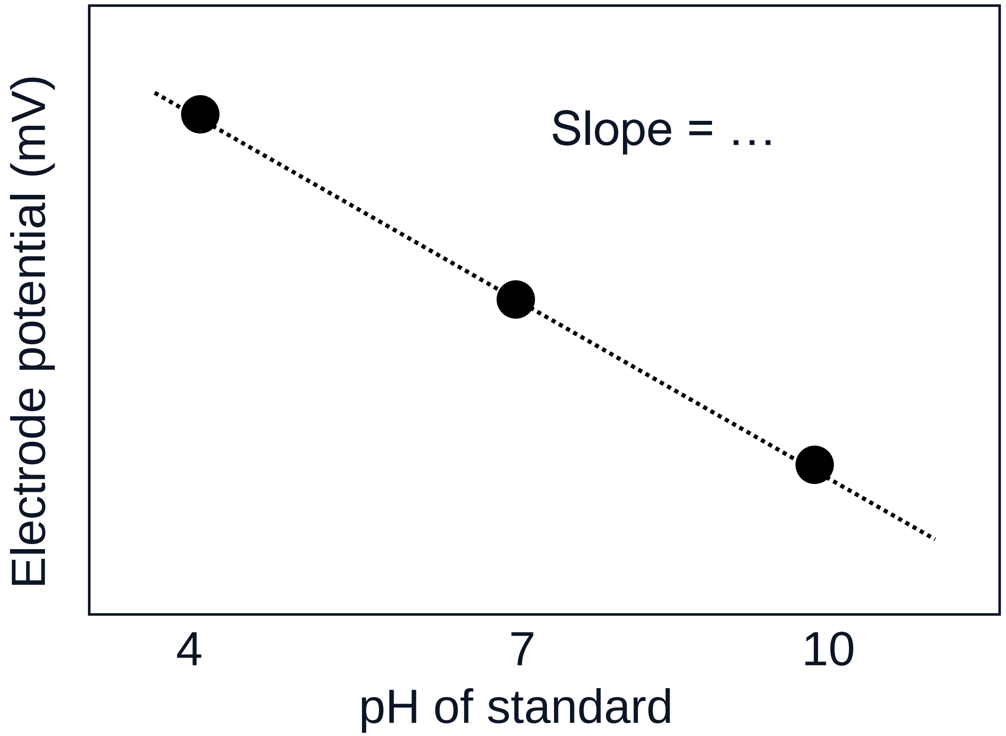
After Reilly (2001) TWRI 3,B8

USGS
science for a changing world

What is Model Calibration?

Other Disciplines

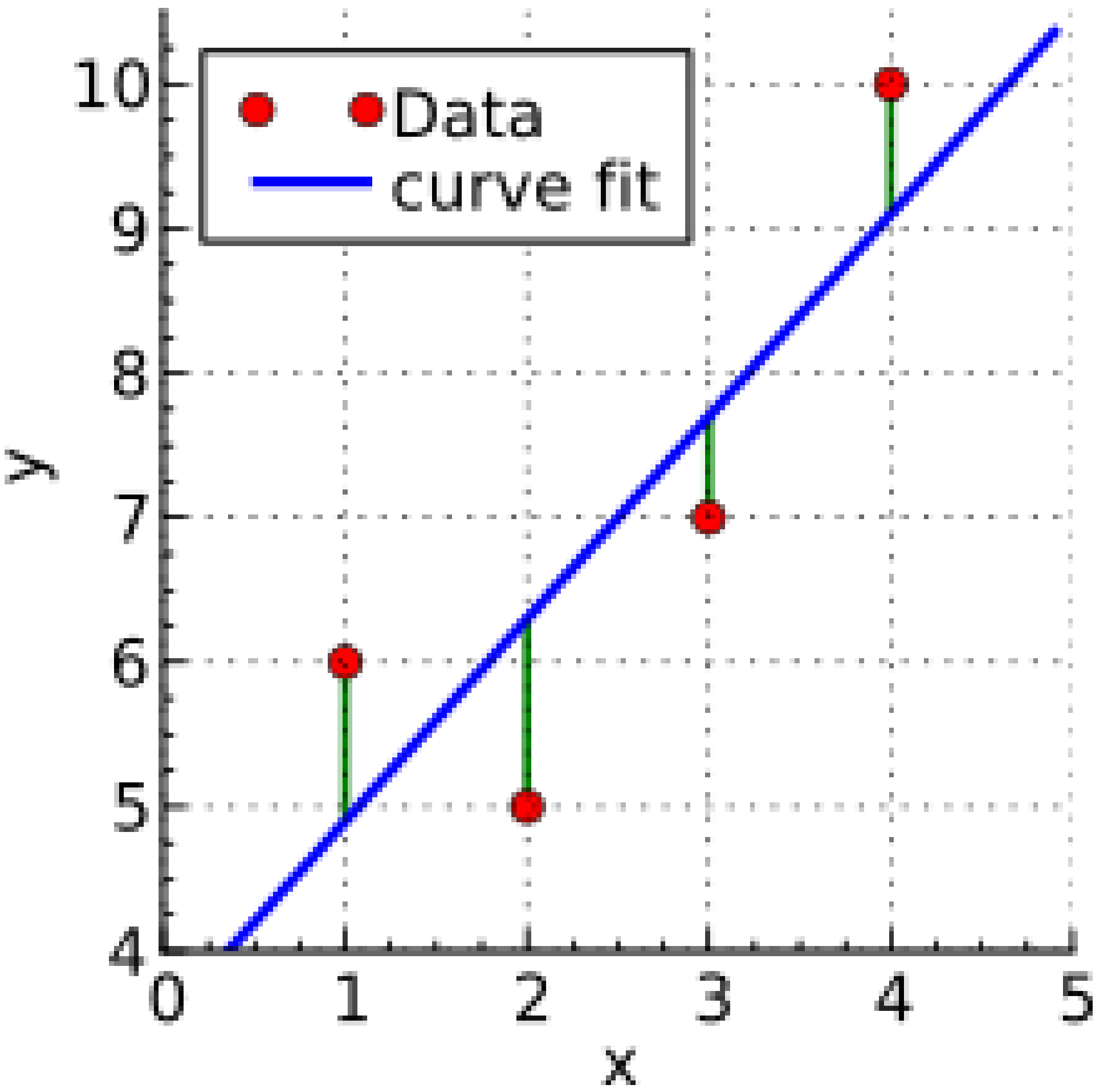
instrumentation 'calibration'



USGS
science for a changing world

Other Disciplines

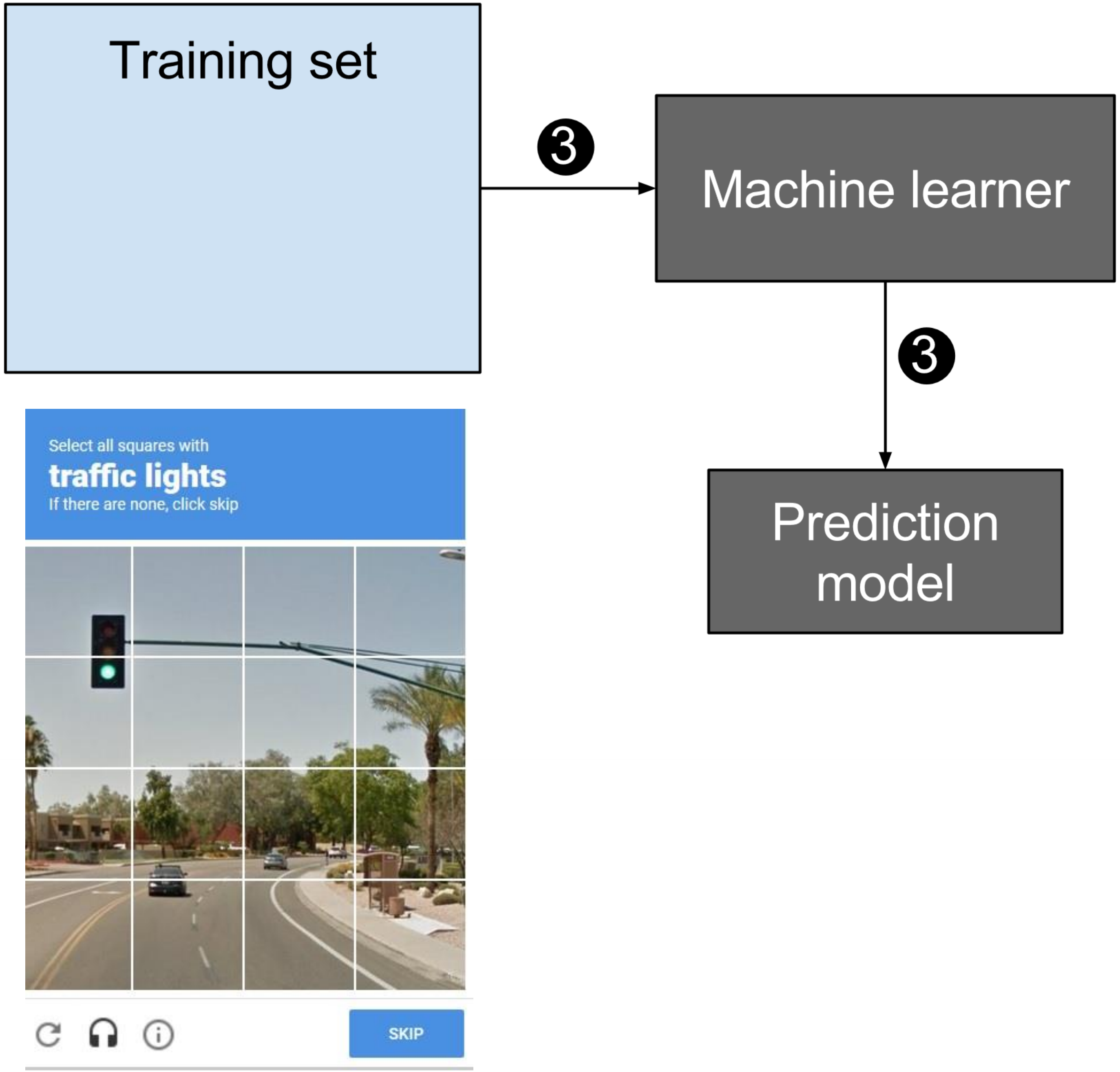
‘fitting’ curves



USGS
science for a changing world

Other Disciplines

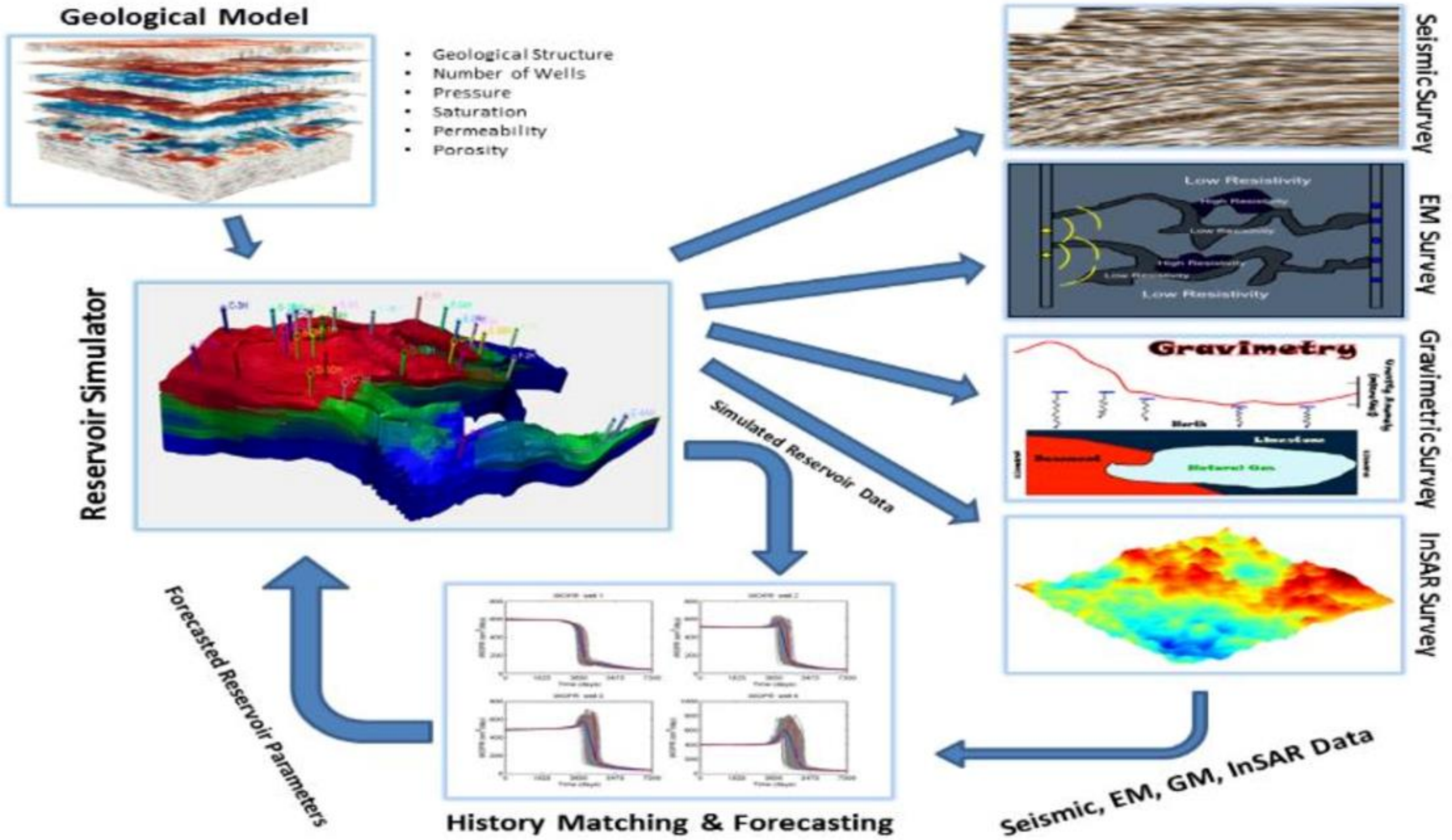
‘training’ machine learning models



USGS
science for a changing world

Other Disciplines

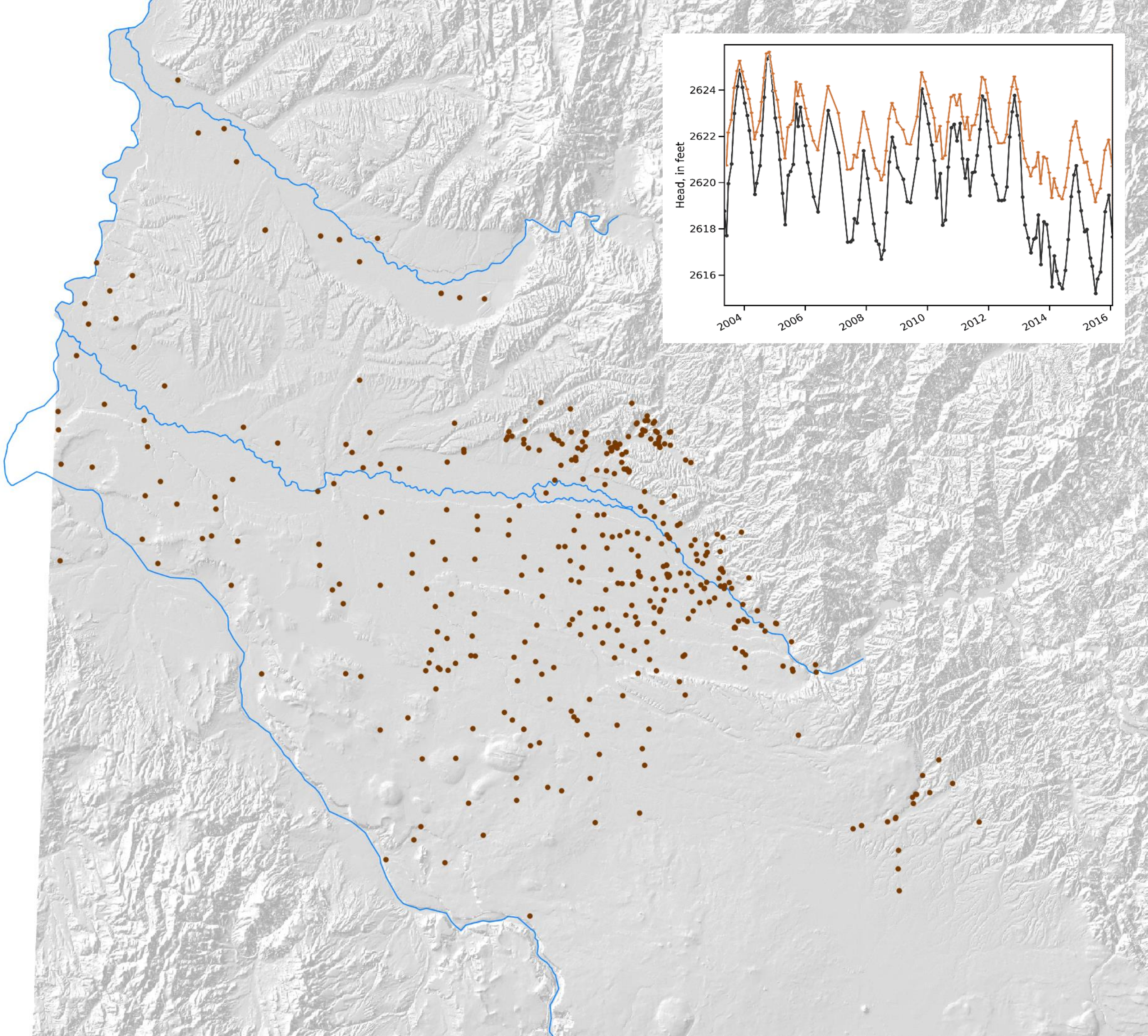
‘history-matching’ petroleum reservoir models



USGS
science for a changing world

Groundwater Modelling

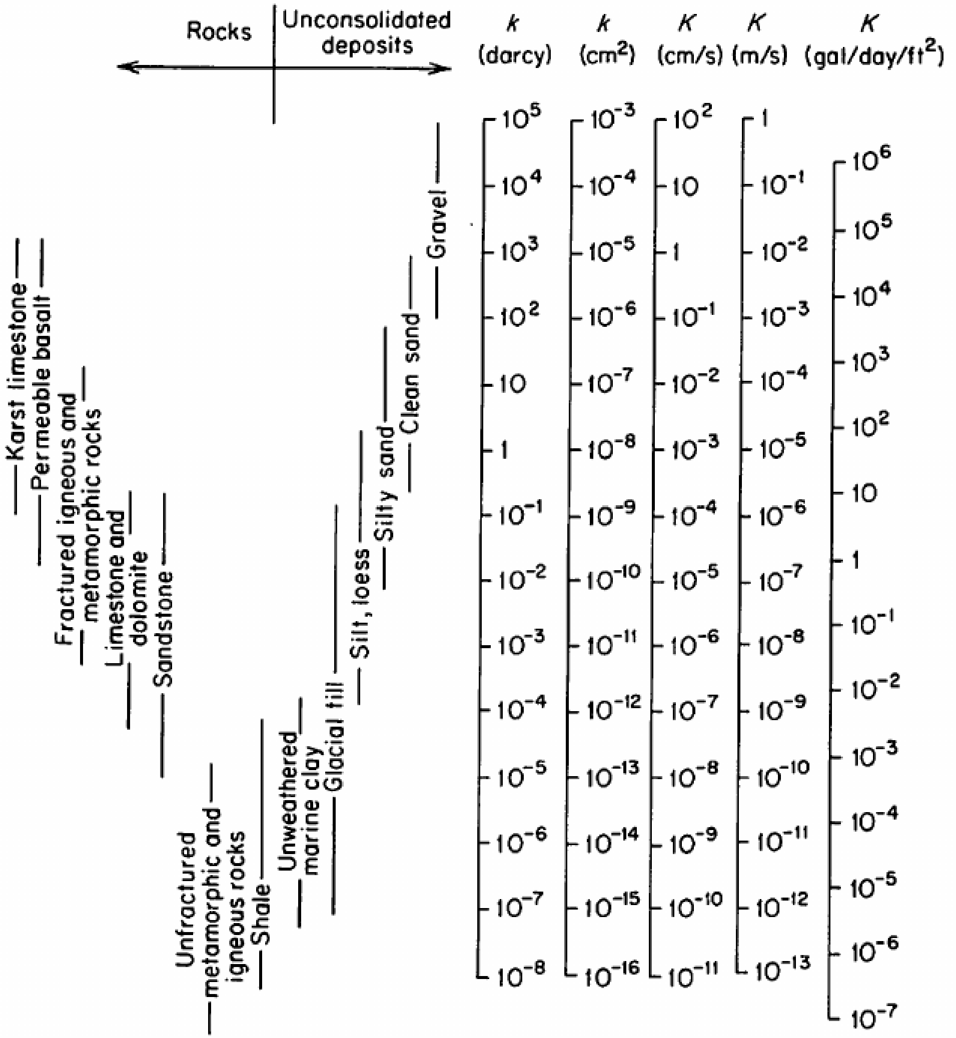
Calibration: History-matching, Parameter estimation, Inverse modelling



USGS
science for a changing world

Why do we calibrate?

Unknown Material Properties

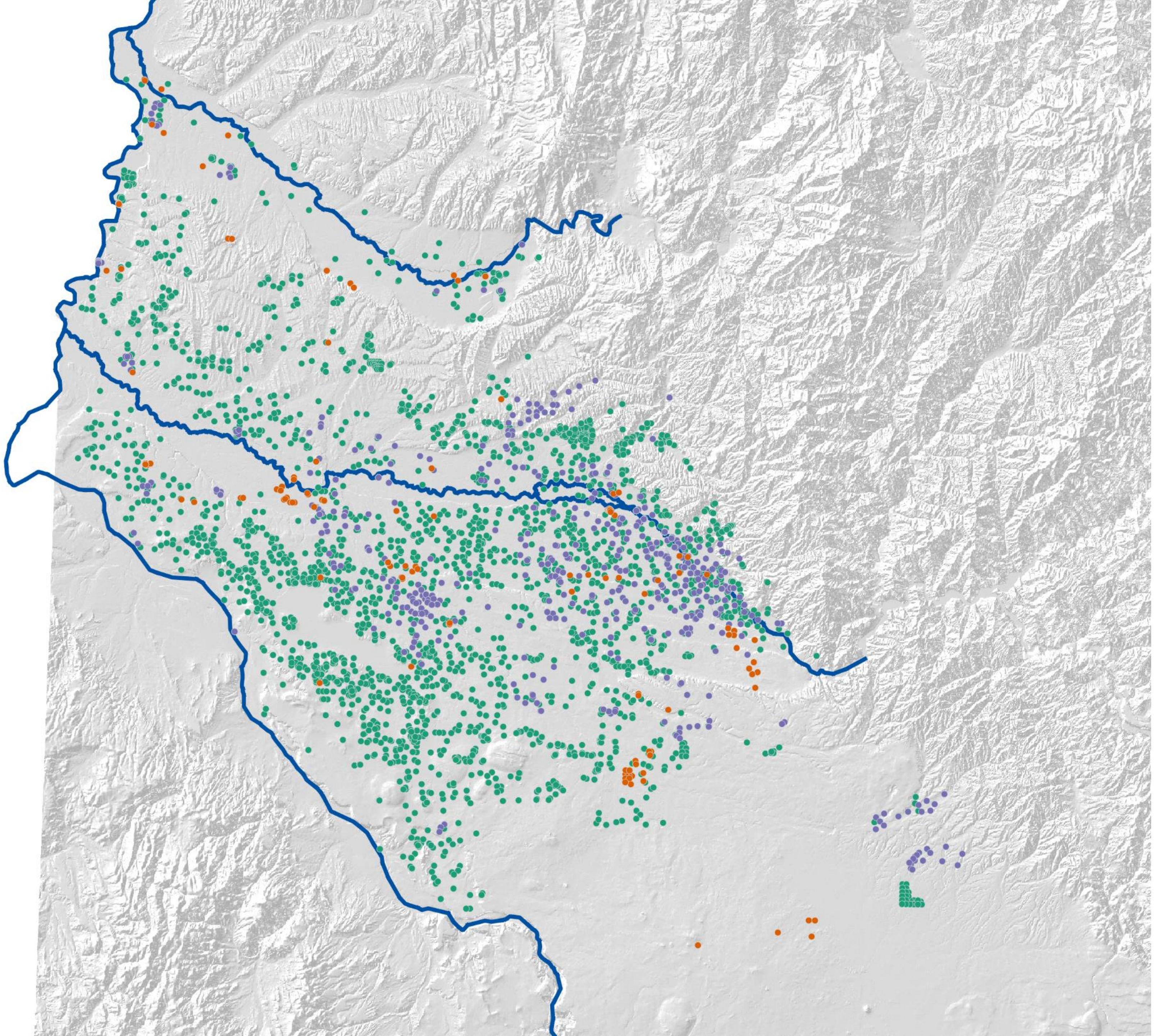


13. LITHOLOGIC LOG: (Describe repairs or abandonment)

Bore Dia.	From	To	**Remarks:** Lithology, Water Quality & Temperature	Water Y	Water N
18"	0'	2'	brown top soil		X
18"	2'	4'	brown clay		X
18"	4'	27'	brown gravel		X
18"	27'	46'	brown clay		X
18"	46'	54'	brown sand & gravel	X	
18"	54'	61'	brown clay		X
18"	61'	89'	brown sand & gravel	X	
18"	89'	91'	brown sandy clay		X
18"	91'	98'	brown sand & gravel	X	
18"	98'	108	brown sand fine	X	
18"	108	114	brown sand & clay strips	X	
18"	114	127	brown sand fine	X	
18"	127	130	brown clay		X
18"	130	143	brown sand & clay strips	X	
			RECEIVED		

science for a changing world

Unknown Boundary Fluxes



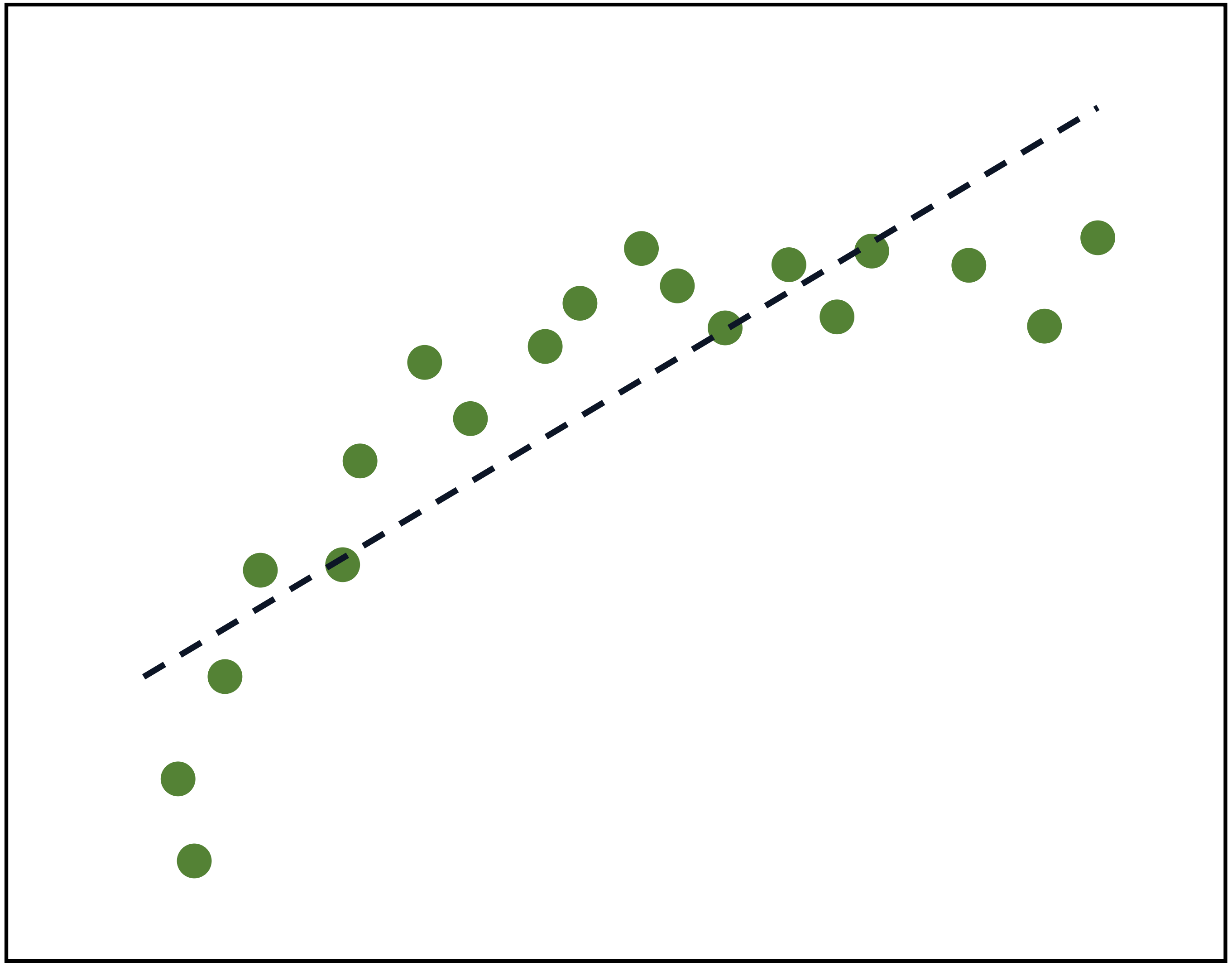
science for a changing world

Model Structural Error

Model represents a process ‘wrong’

Too simple?

Parameter compensation of model error



Unavoidable?

Simple

Complex

too much structural error

too time consuming and unstable to calibrate

more resolution

more parameters

more processes

more realistic process representation

less numerically stable

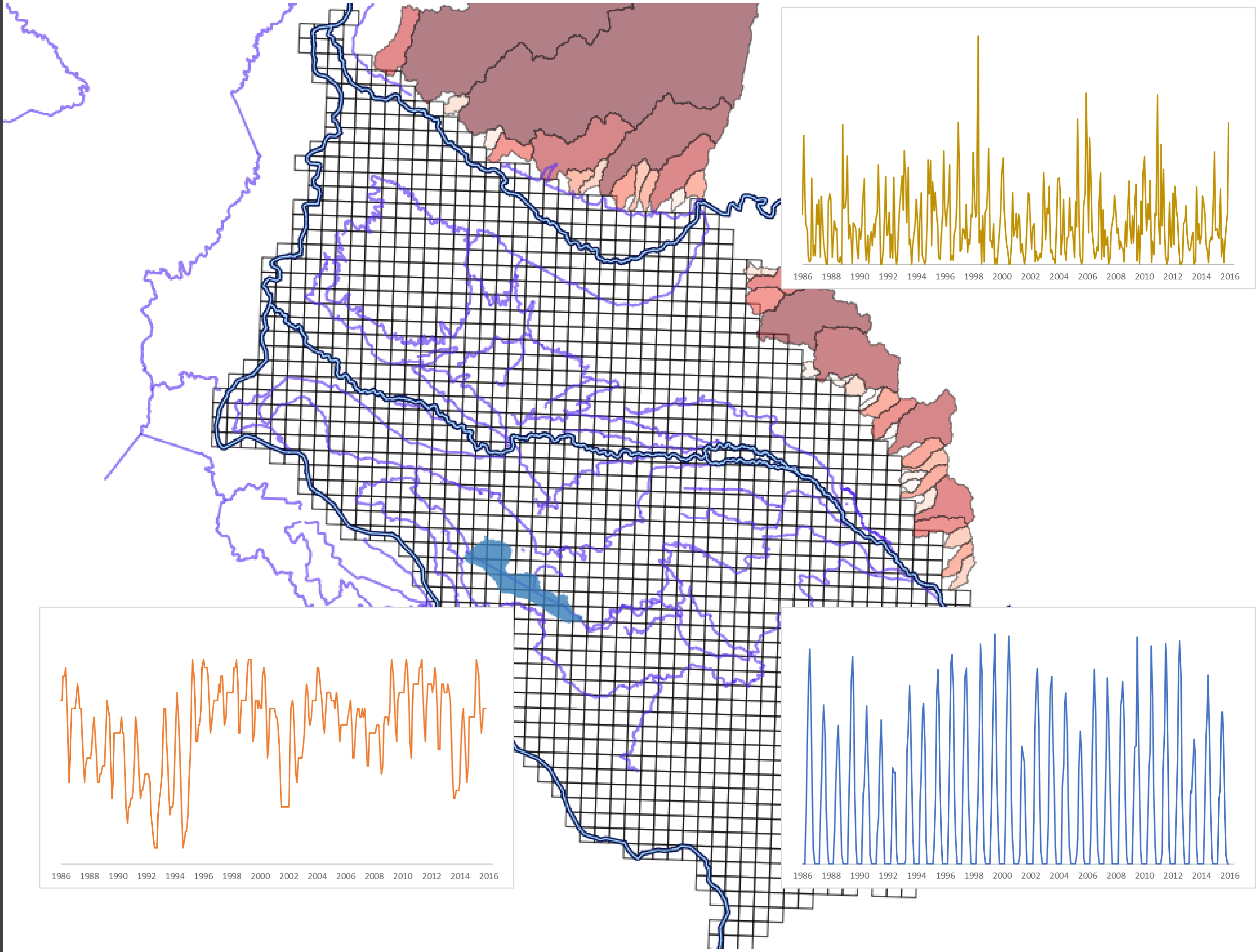
longer run times

…still structural error

USGS
science for a changing world

How do we calibrate?

Construct Historic Scenario

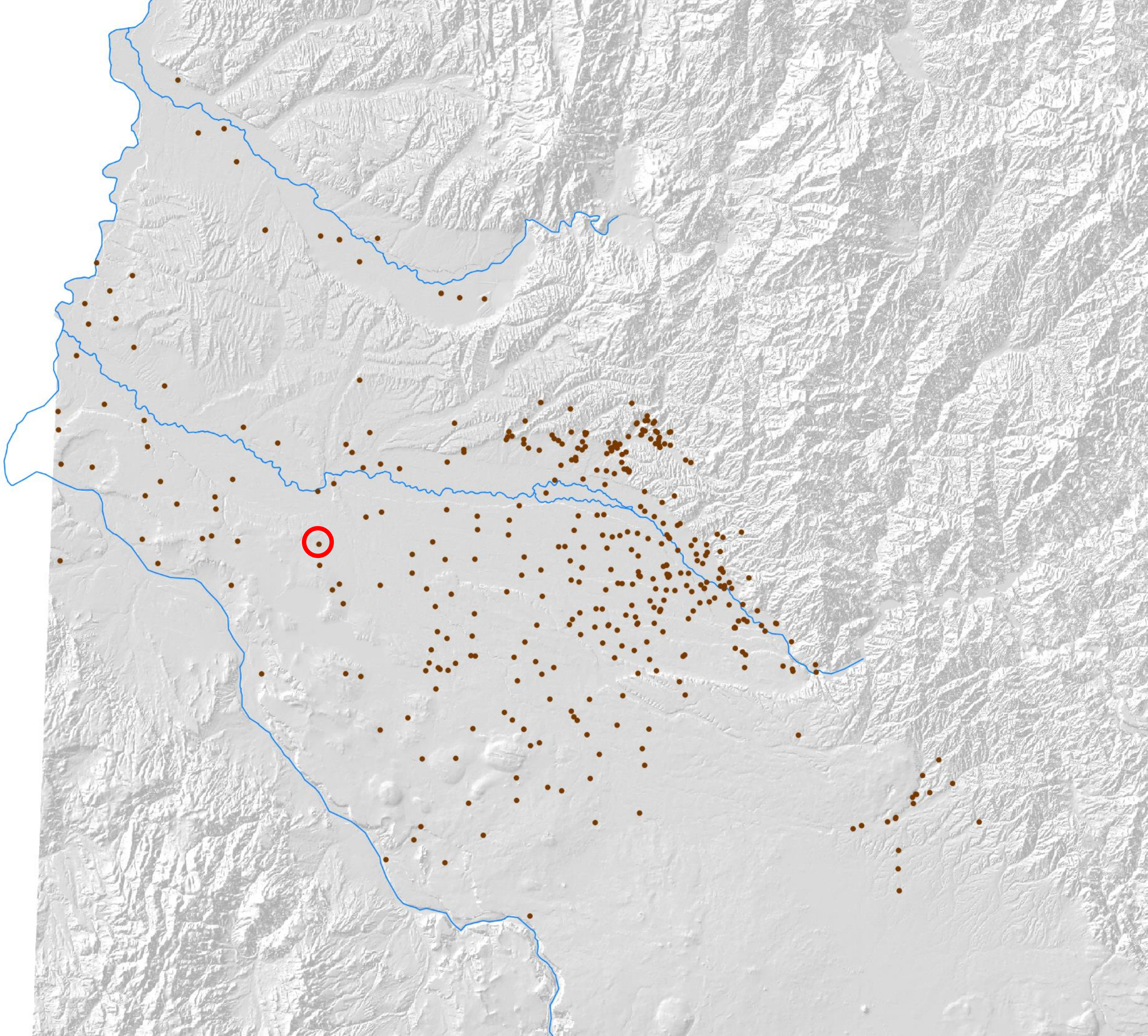


science for a changing world

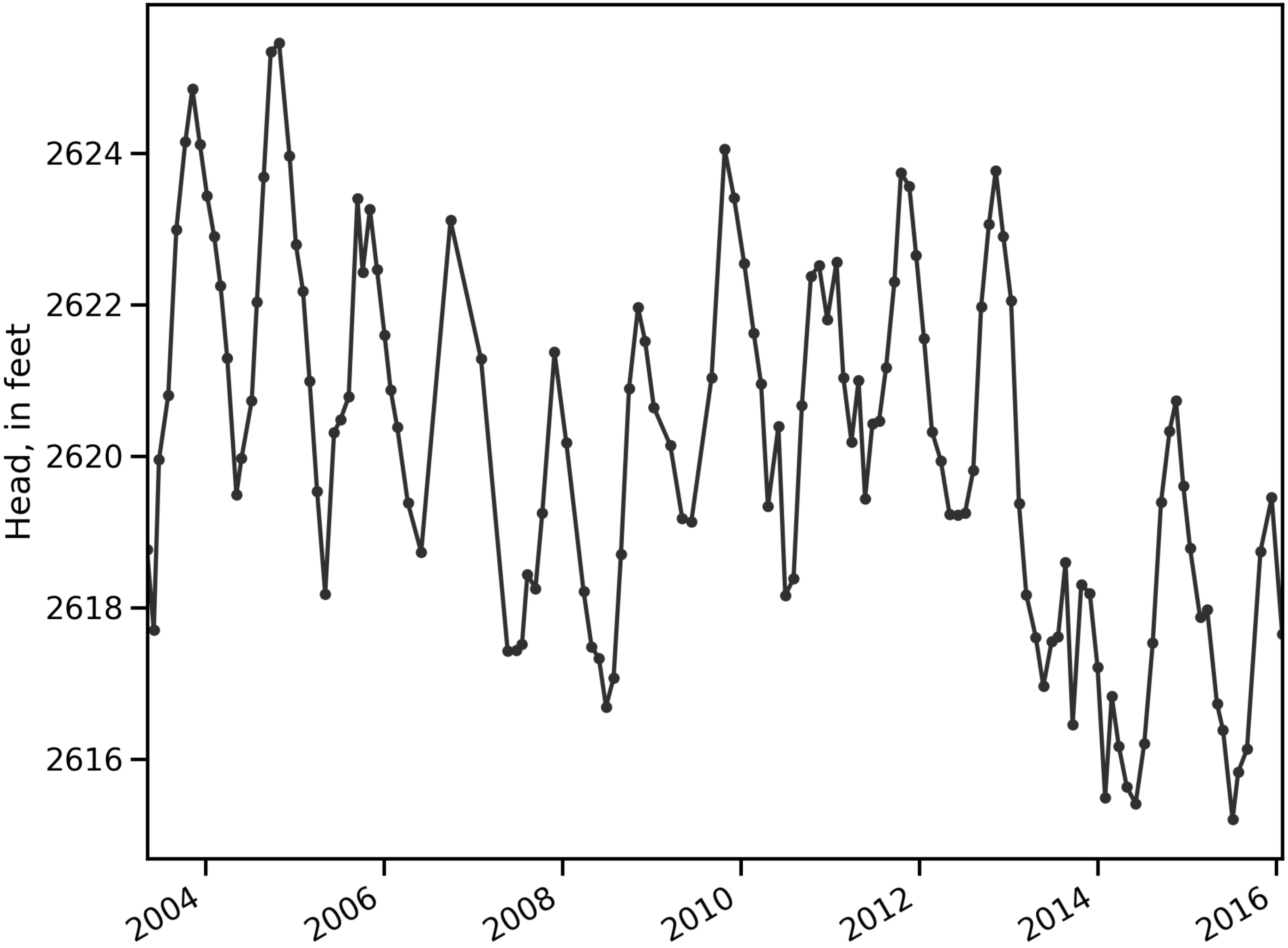
Field Observations

USGS
science for a changing world

Field Observations

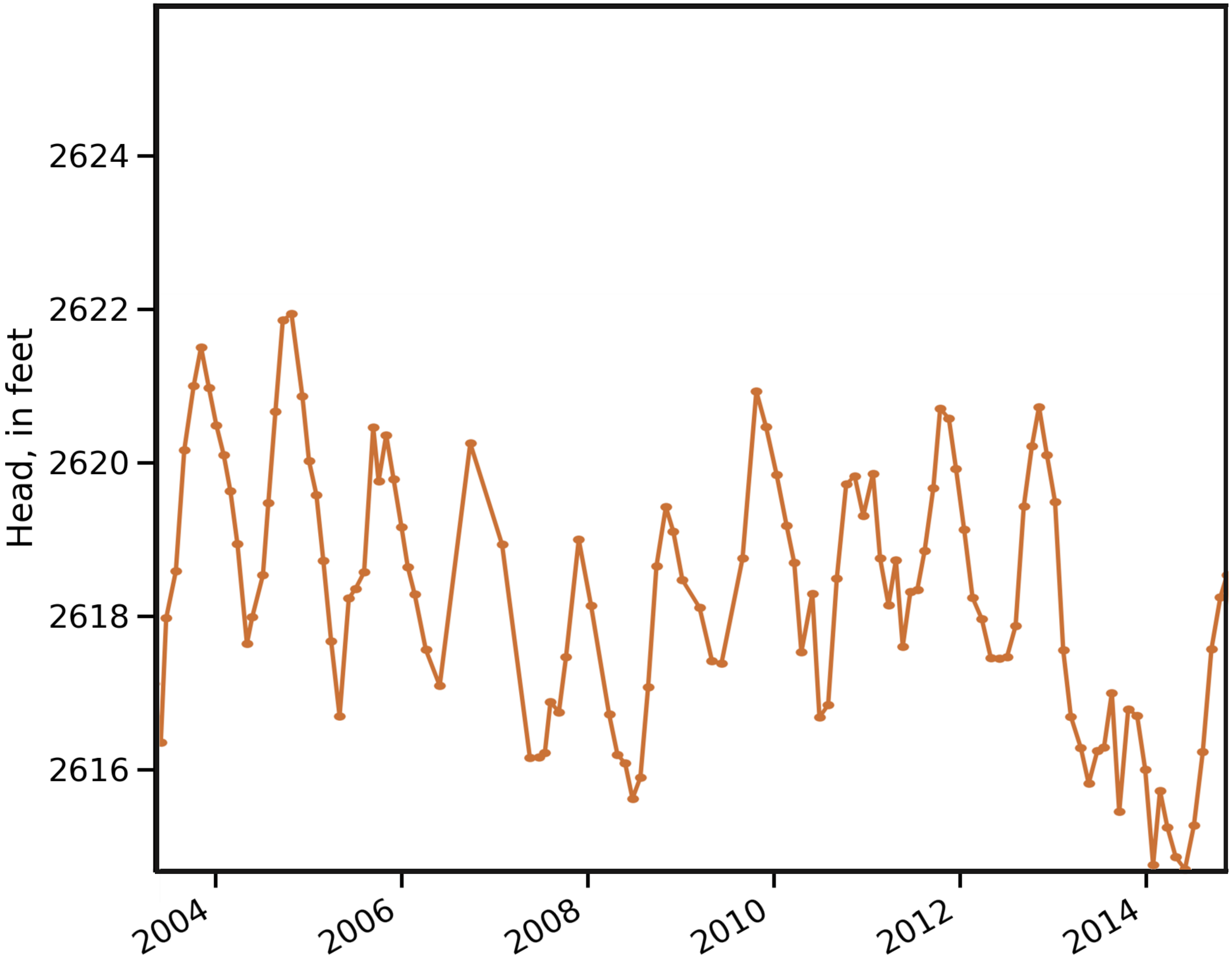


Field Observations



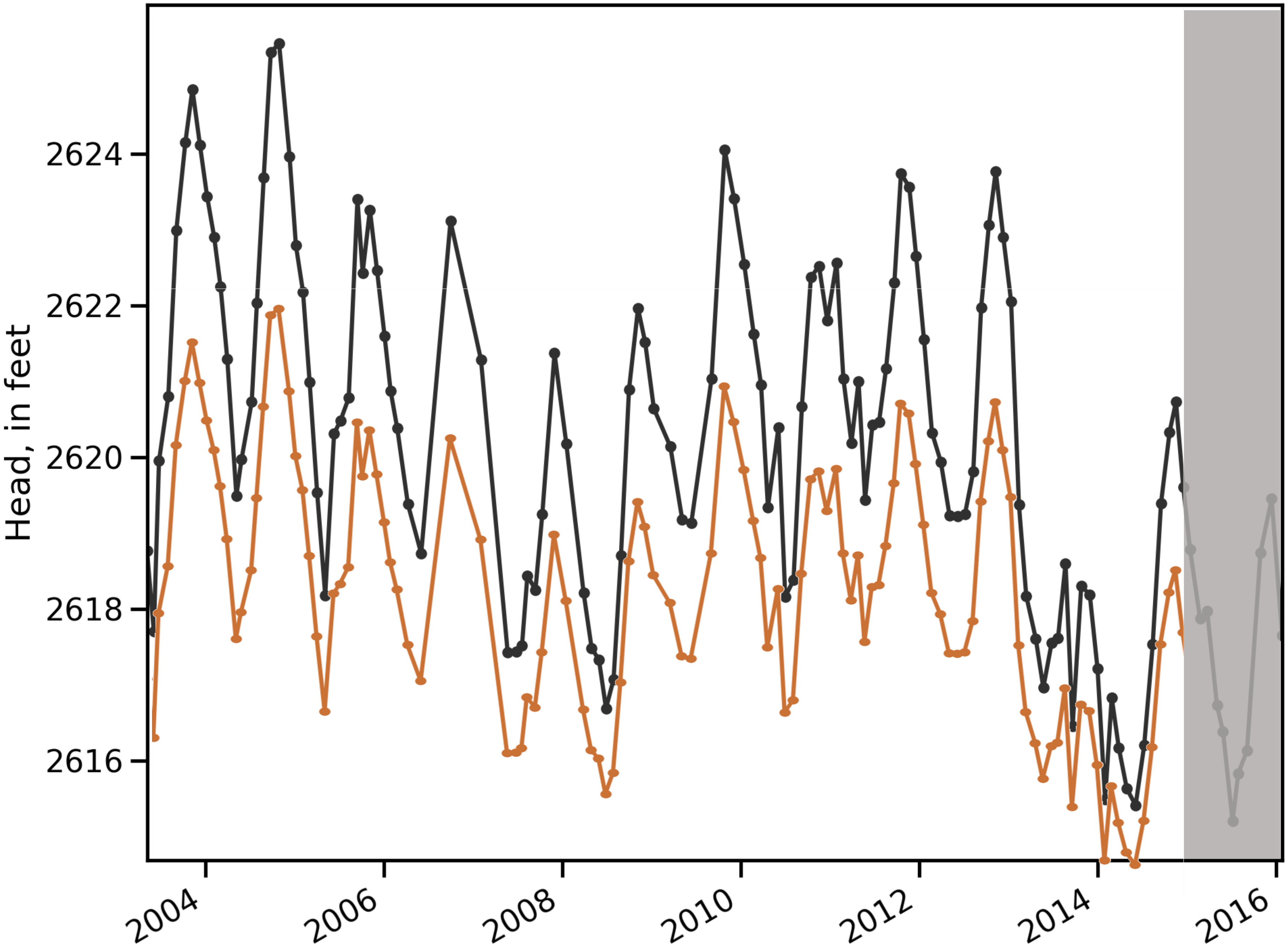
Simulated Equivalents

Simulated Equivalents

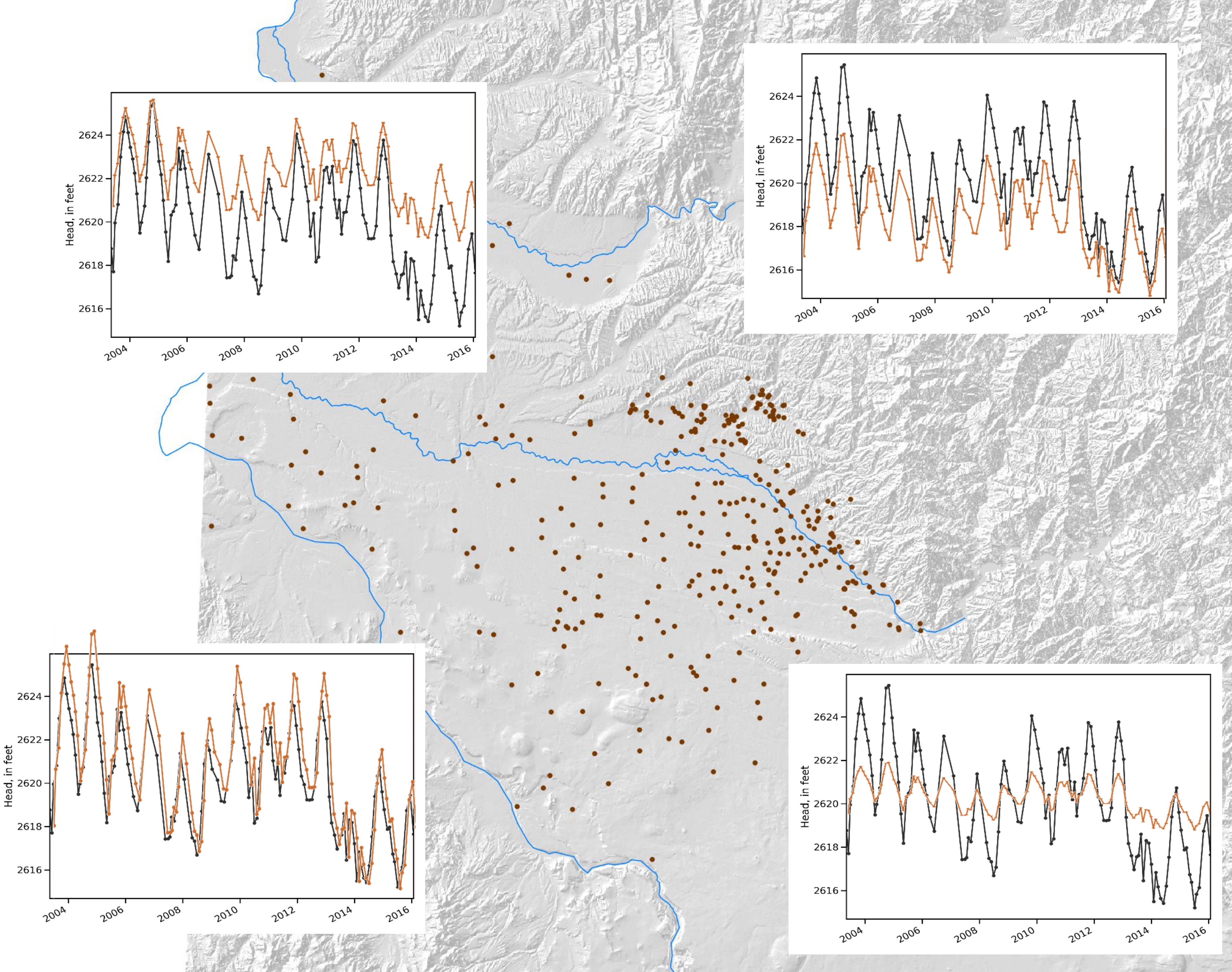


science for a changing world

Observed compared to simulated

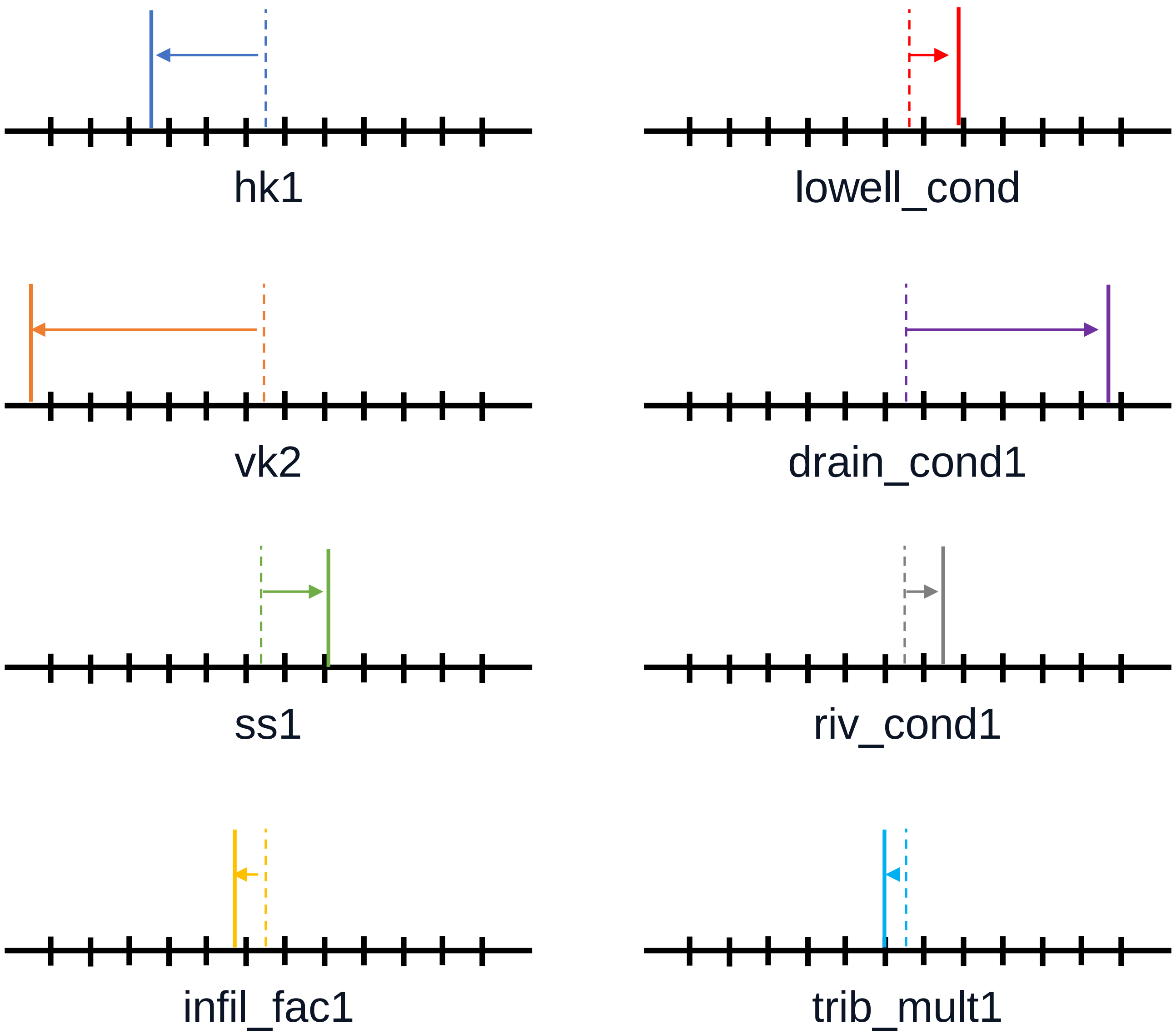


Observed compared to simulated



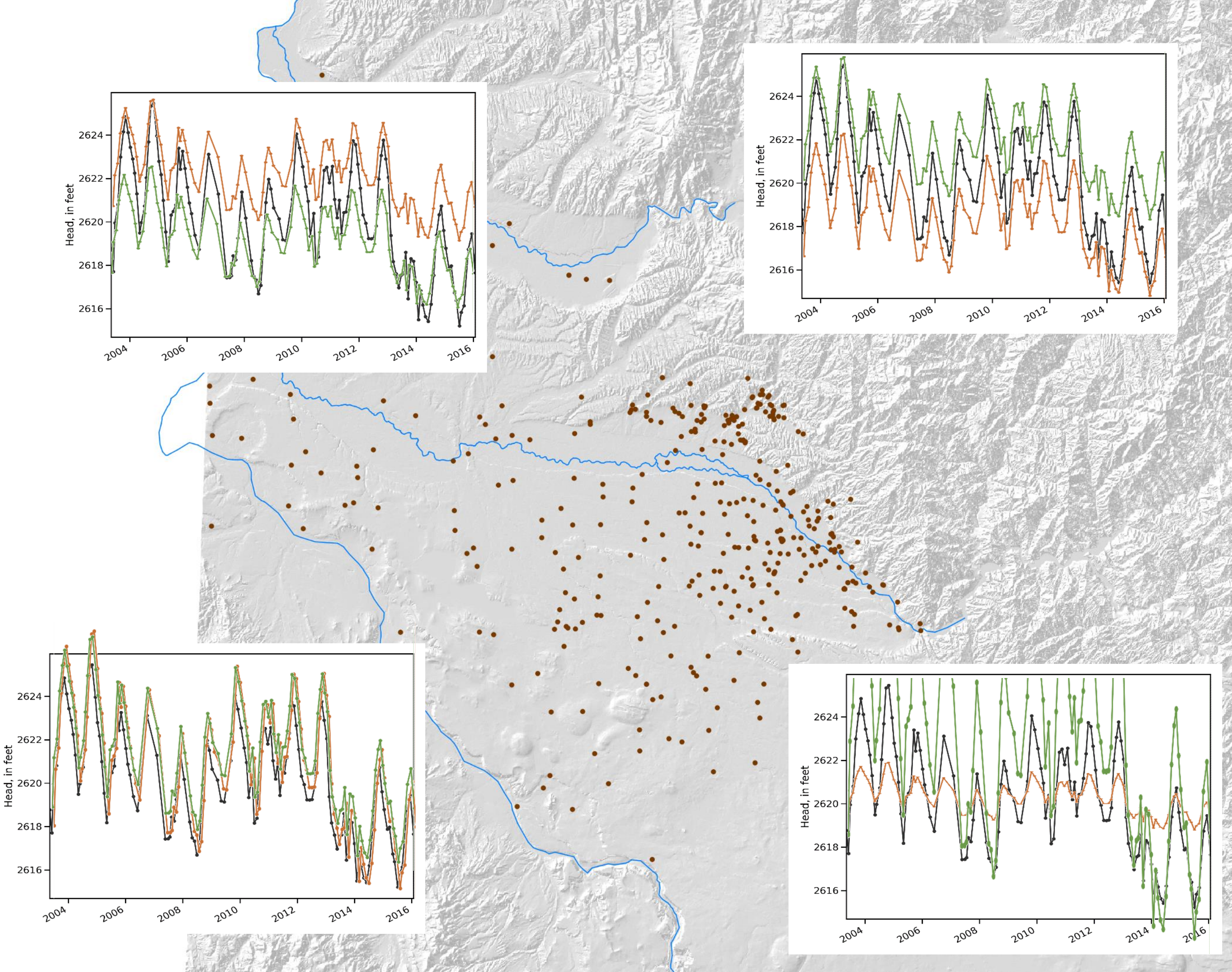
USGS
science for a changing world

Tweak Parameter Values

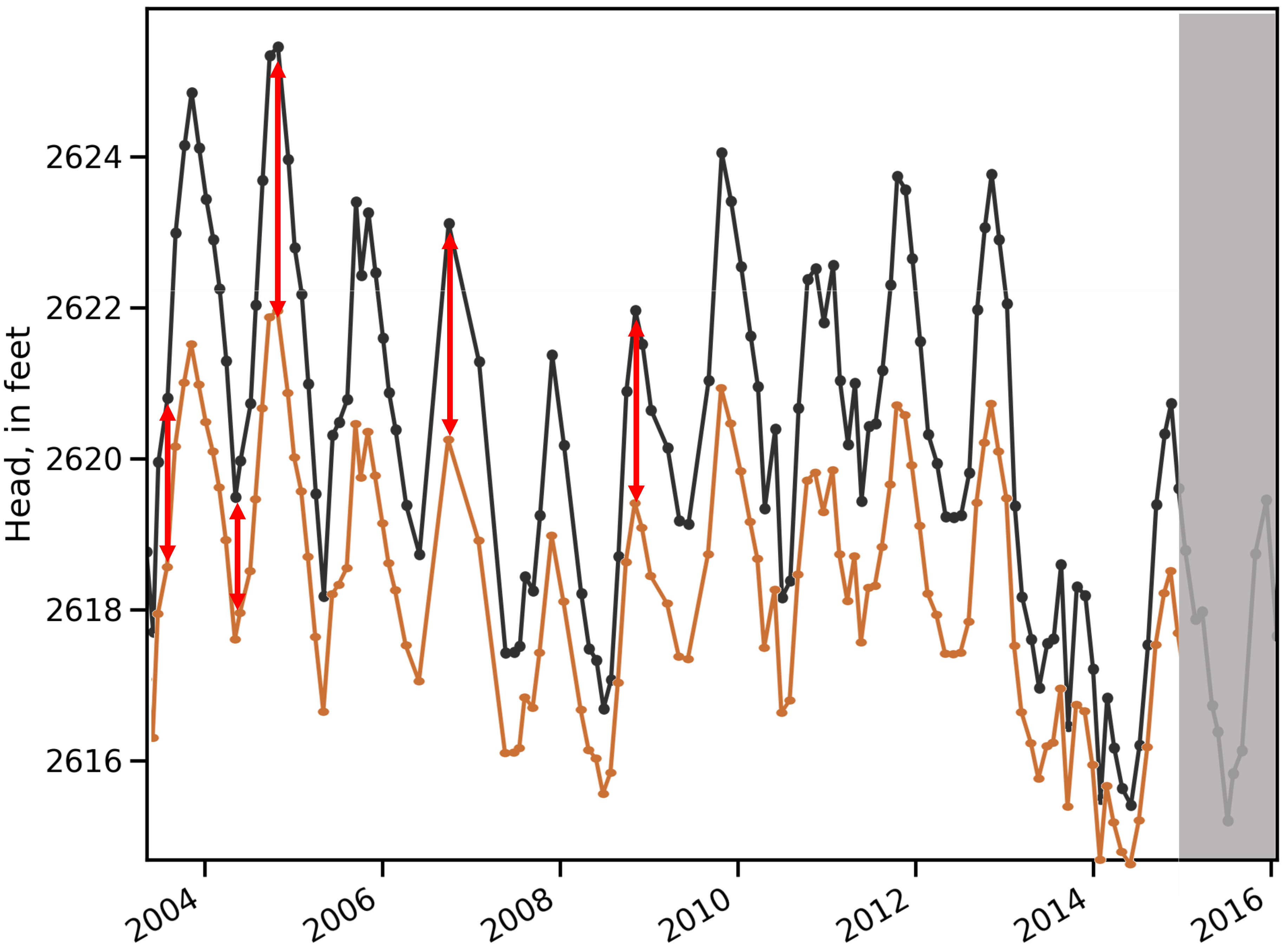


and so on…

Tweak Parameter Values



Residuals

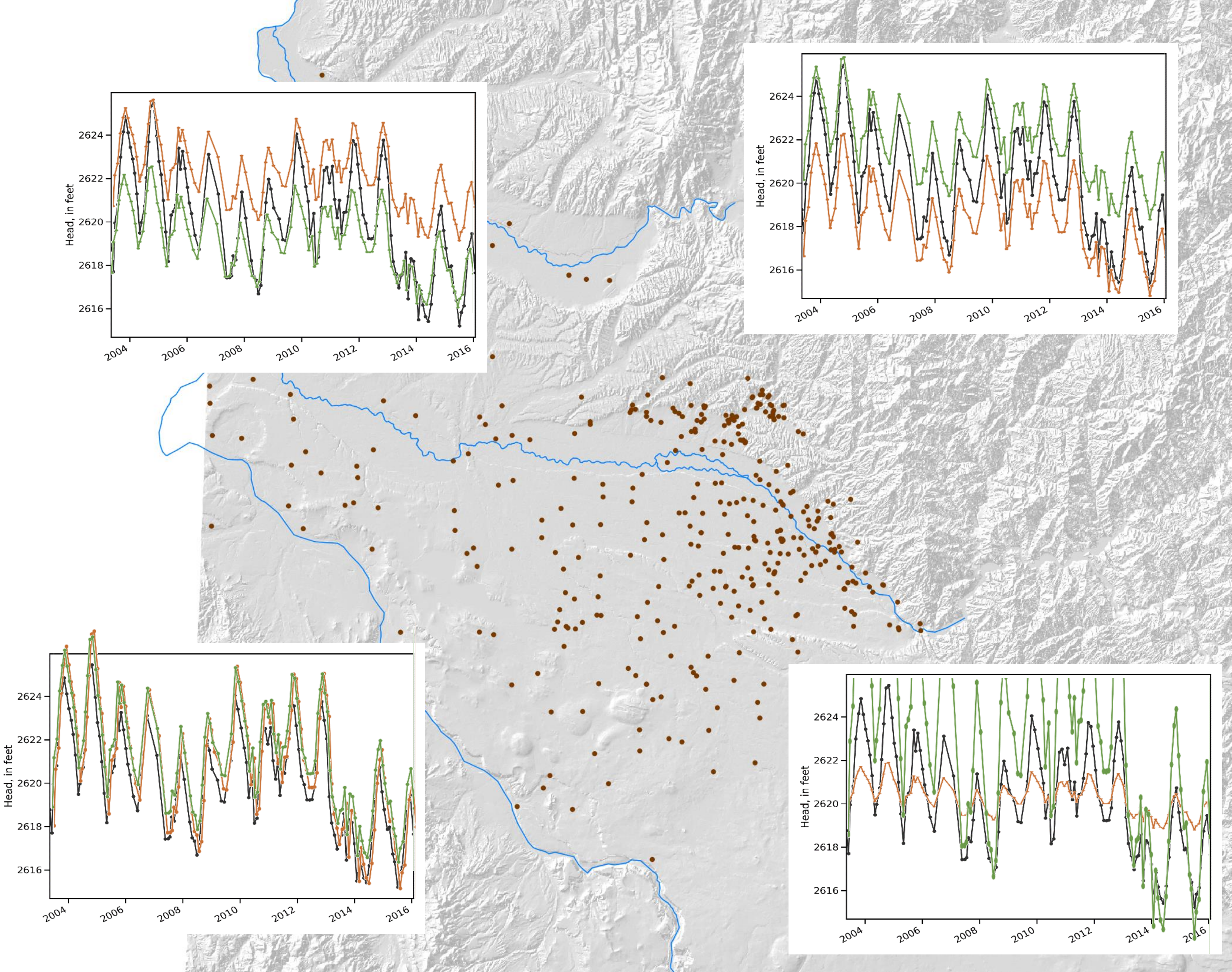


USGS
science for a changing world

Residuals

Residuals = 2, -3, 4, …

Residuals = 1, -5, 2, …



USGS
science for a changing world

Objective Function

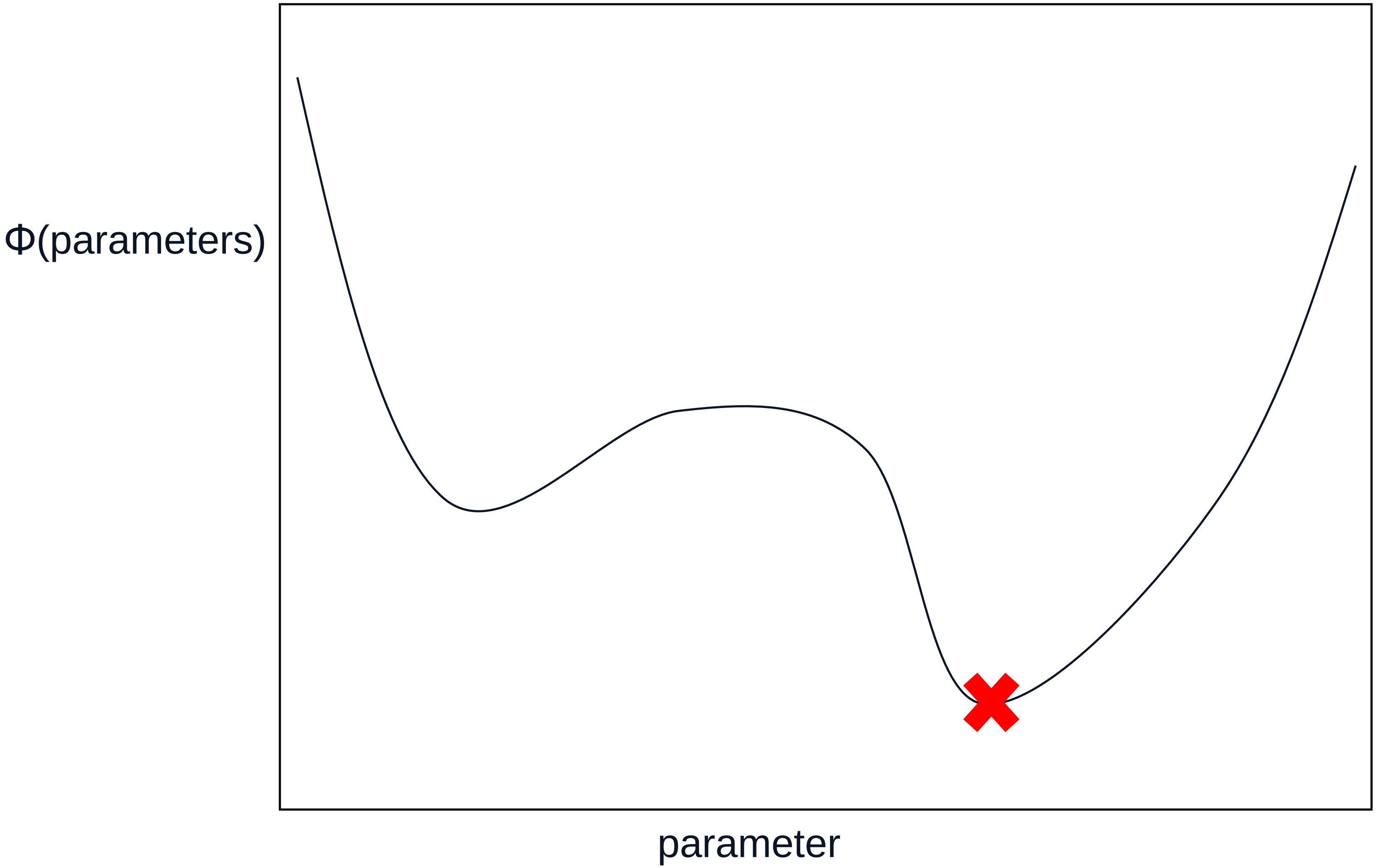
$$r_i = observed_i - simulated_i$$

$$w_i = weight$$

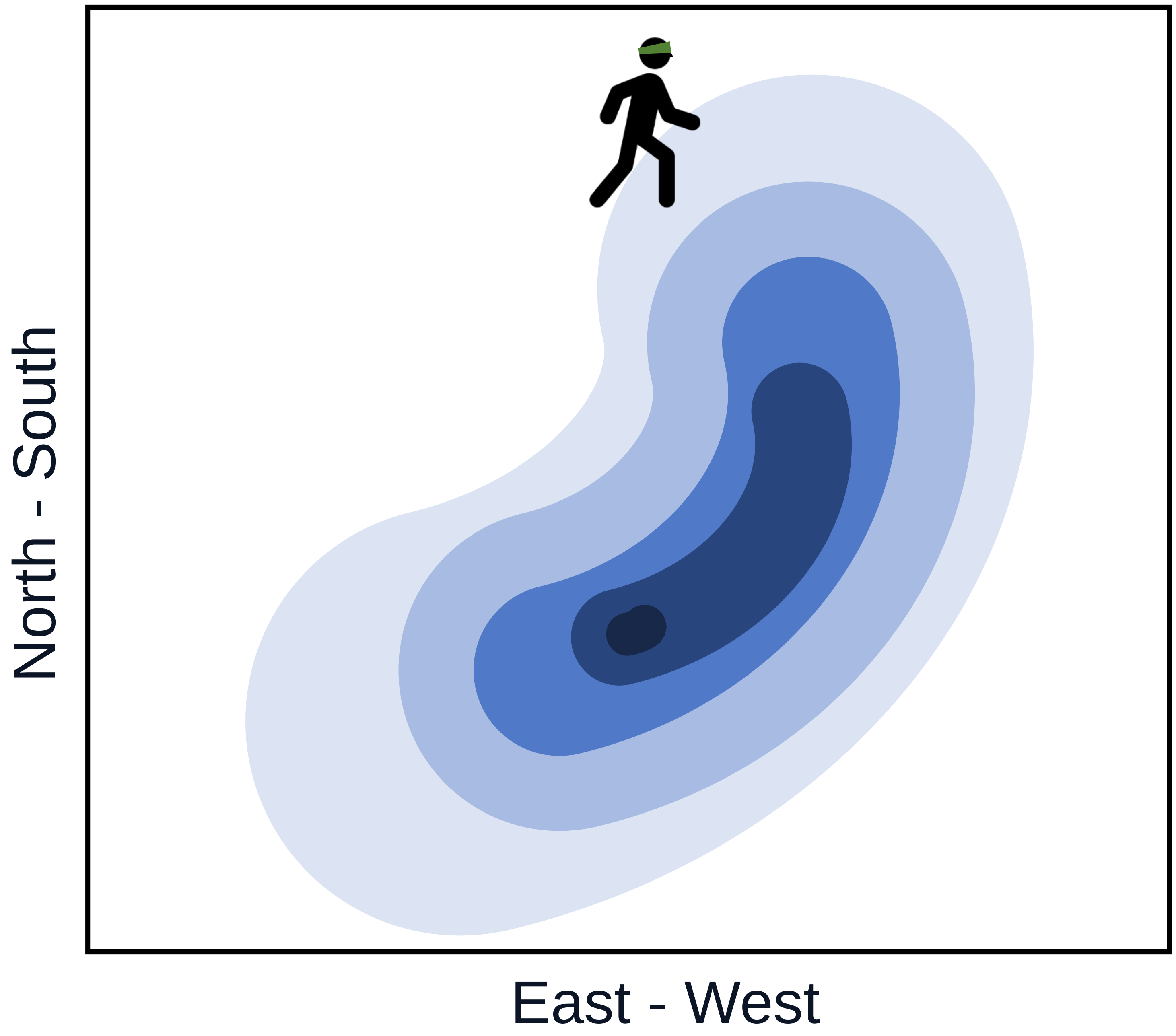
$$objective\ function\colon \Phi = \sum (w_i r_i)^2$$

$$\begin{aligned}\Phi \;=\; & \sum (w_{heads} r_{heads})^2 \\ & + \sum (w_{vert\_head\_diff} r_{vert\_head\_diff})^2 \\ & + \sum (w_{temp\_head\_diff} r_{temp\_head\_diff})^2 \\ & + \sum (w_{drain\_flux} r_{drain\_flux})^2 \\ & + \sum (w_{lowell\_seepage} r_{lowell\_seepage})^2\end{aligned}$$

Minimizing Objective Function

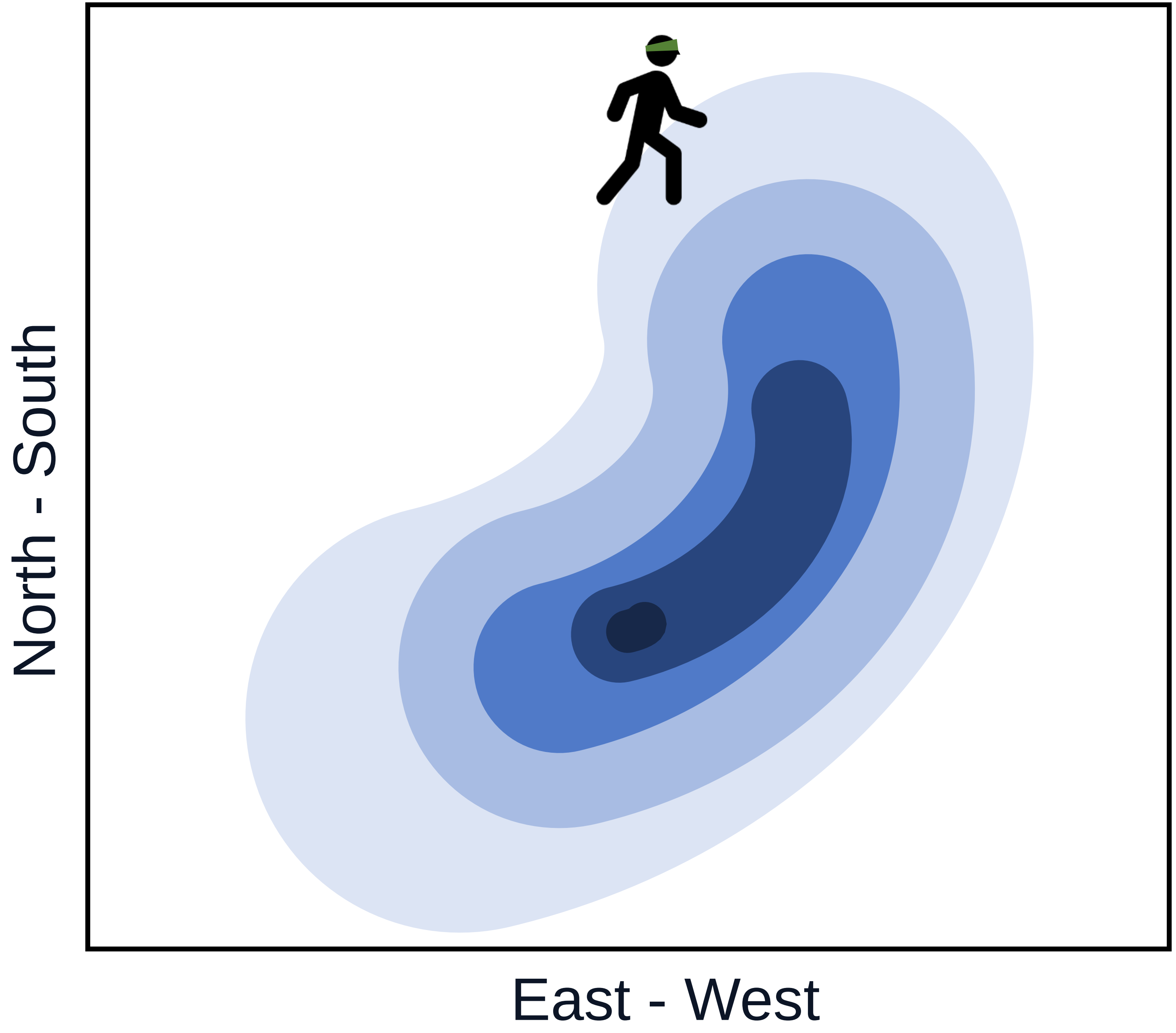


Minimizing Objective Function



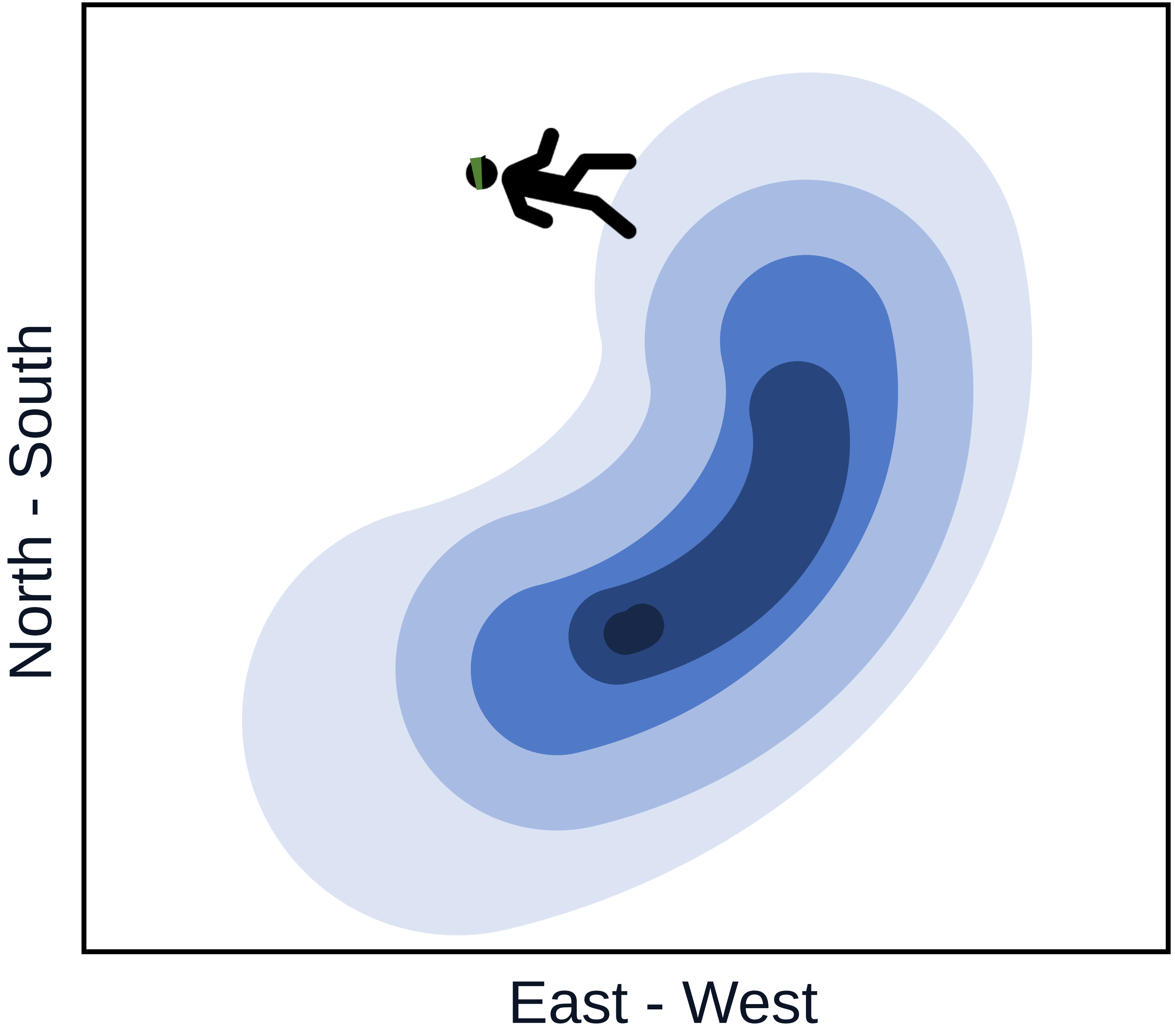
USGS
science for a changing world

Minimizing Objective Function



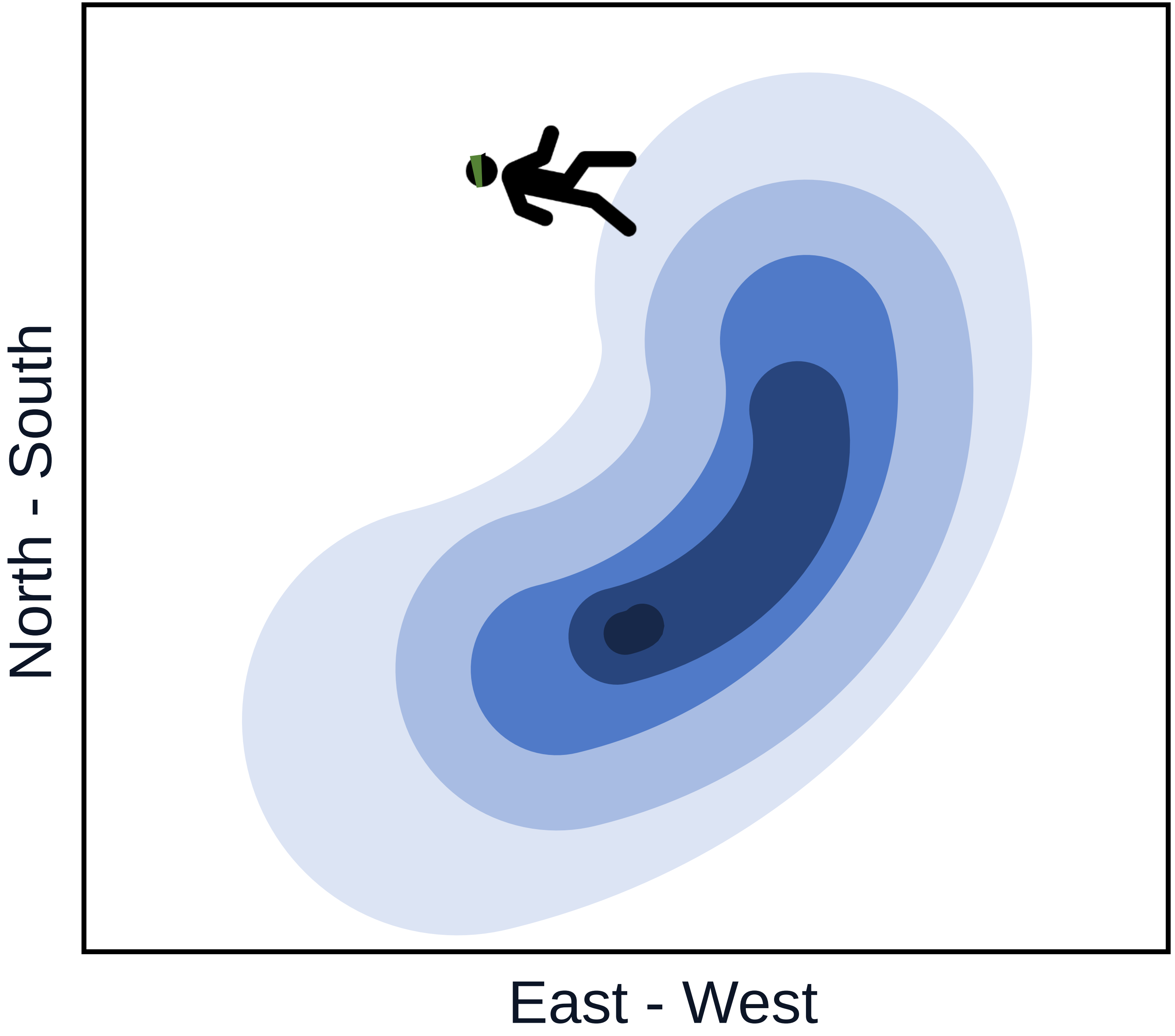
science for a changing world

Minimizing Objective Function



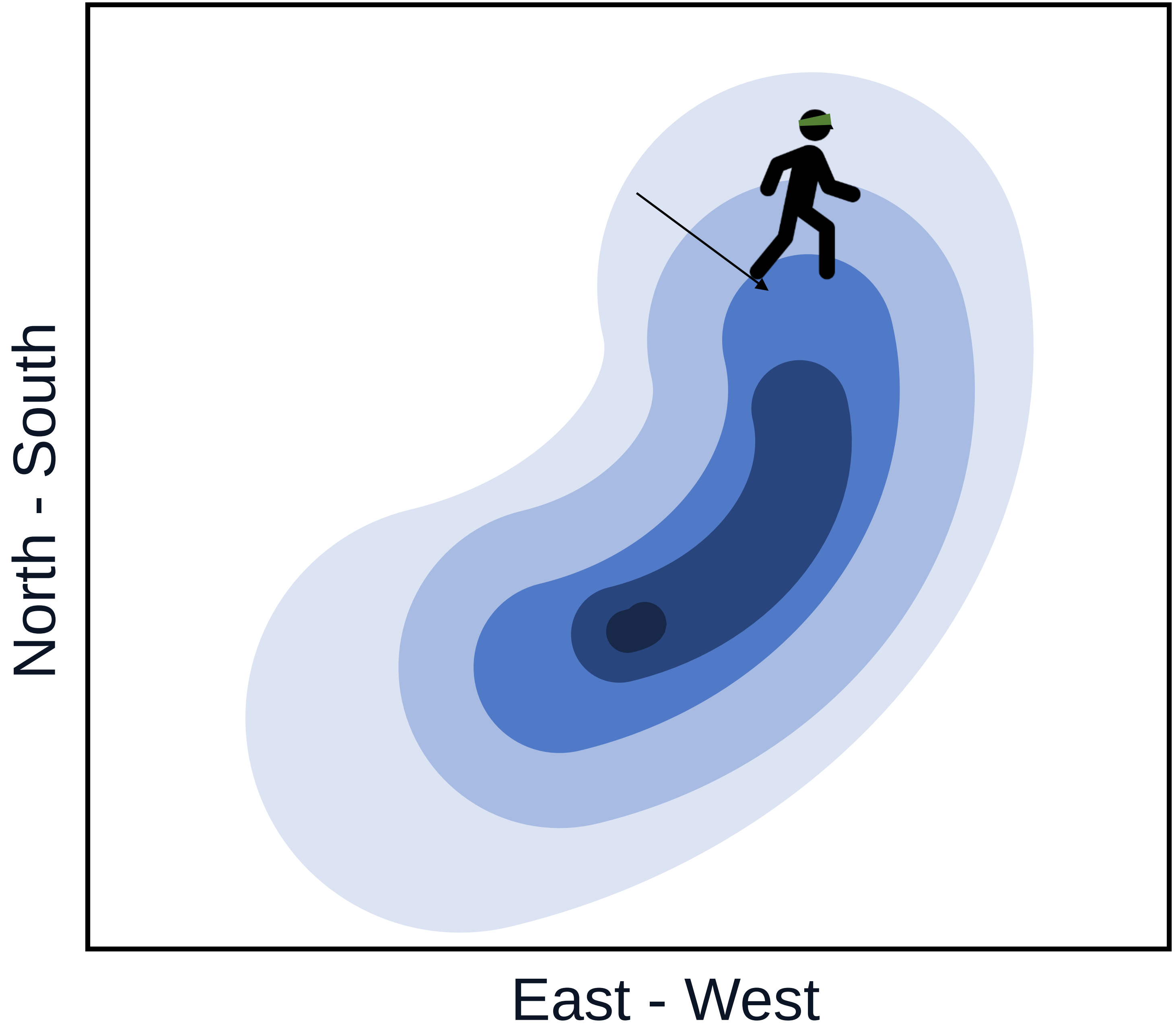
science for a changing world

Minimizing Objective Function



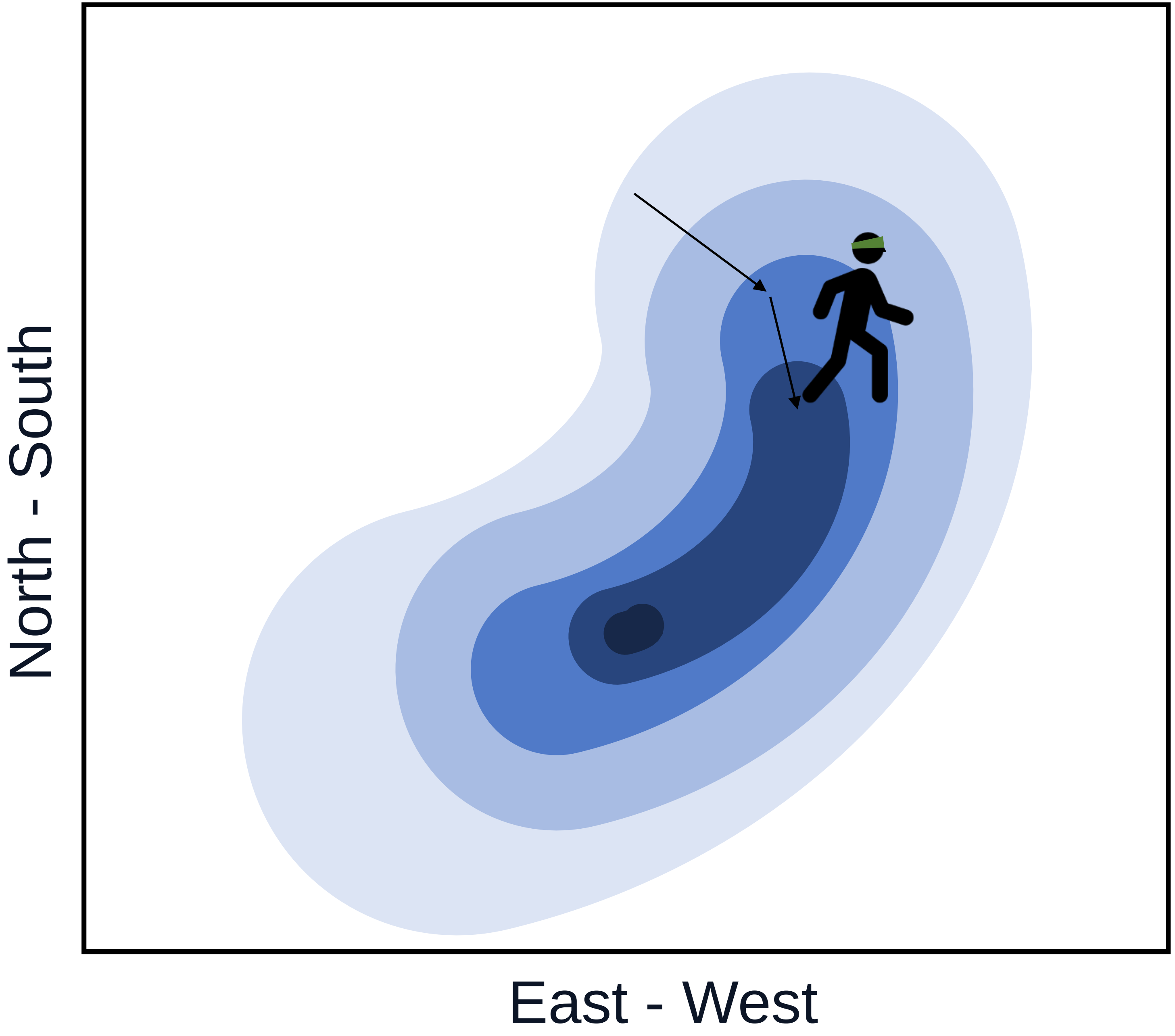
USGS
science for a changing world

Minimizing Objective Function

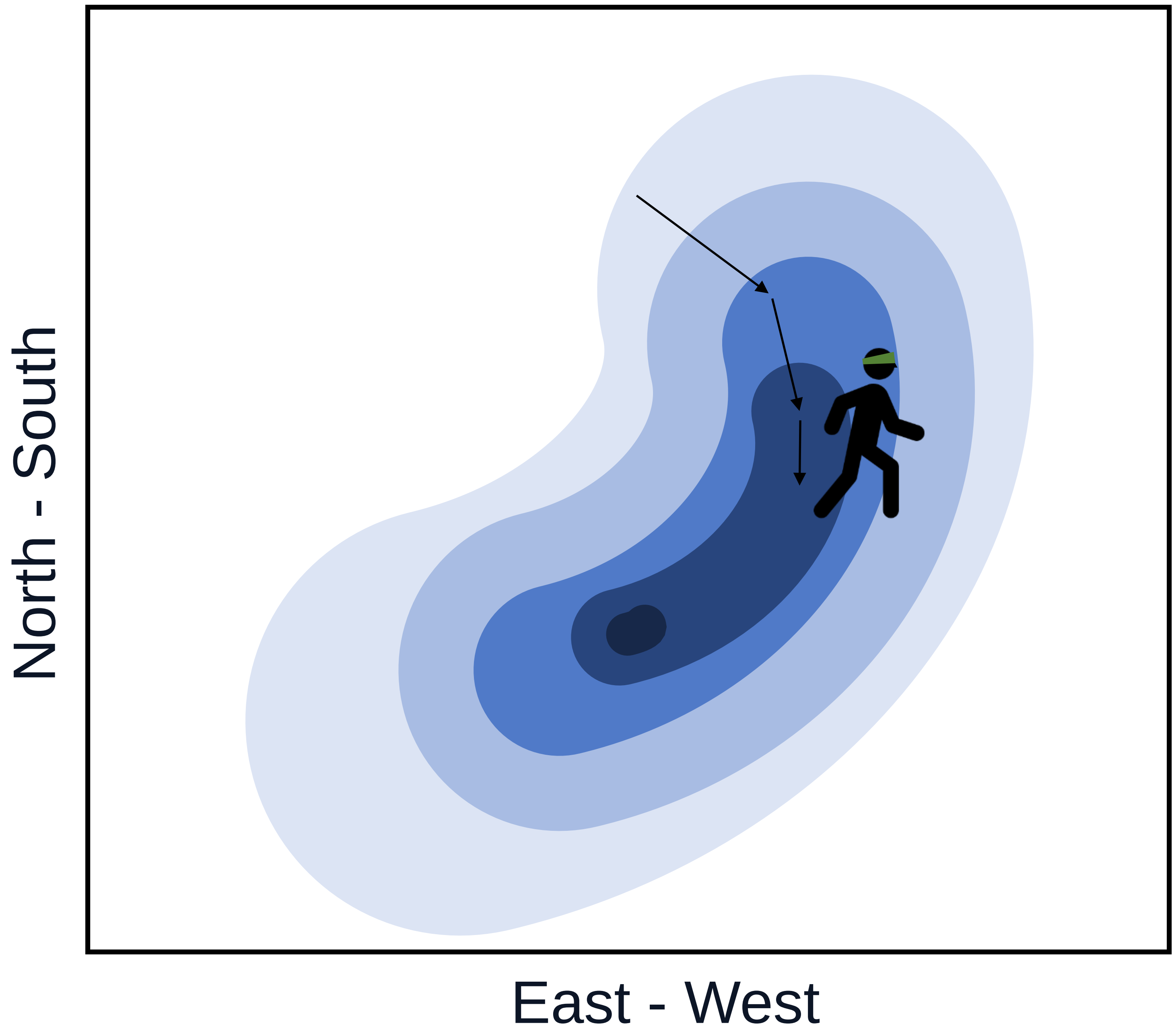


USGS
science for a changing world

Minimizing Objective Function

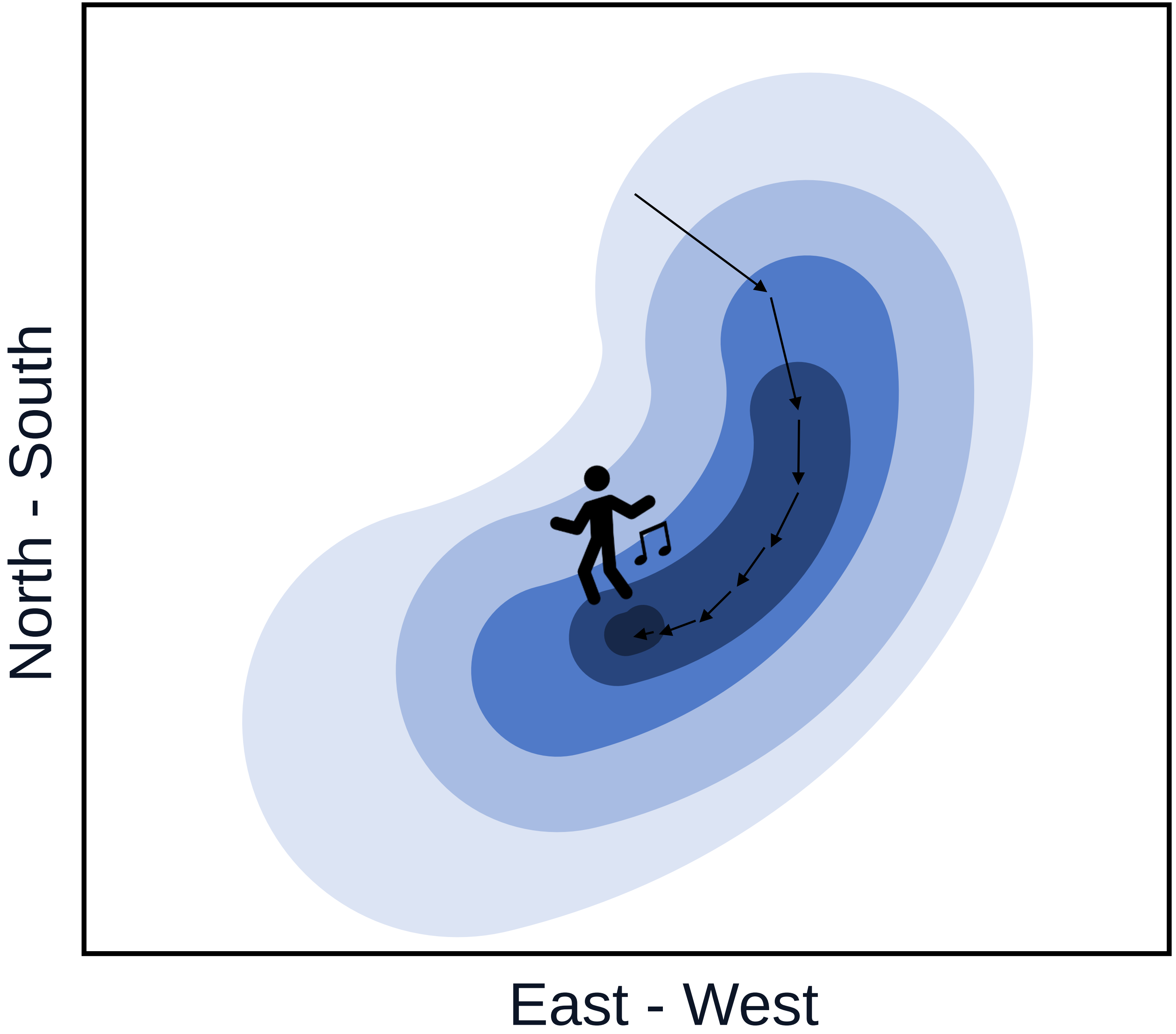


Minimizing Objective Function

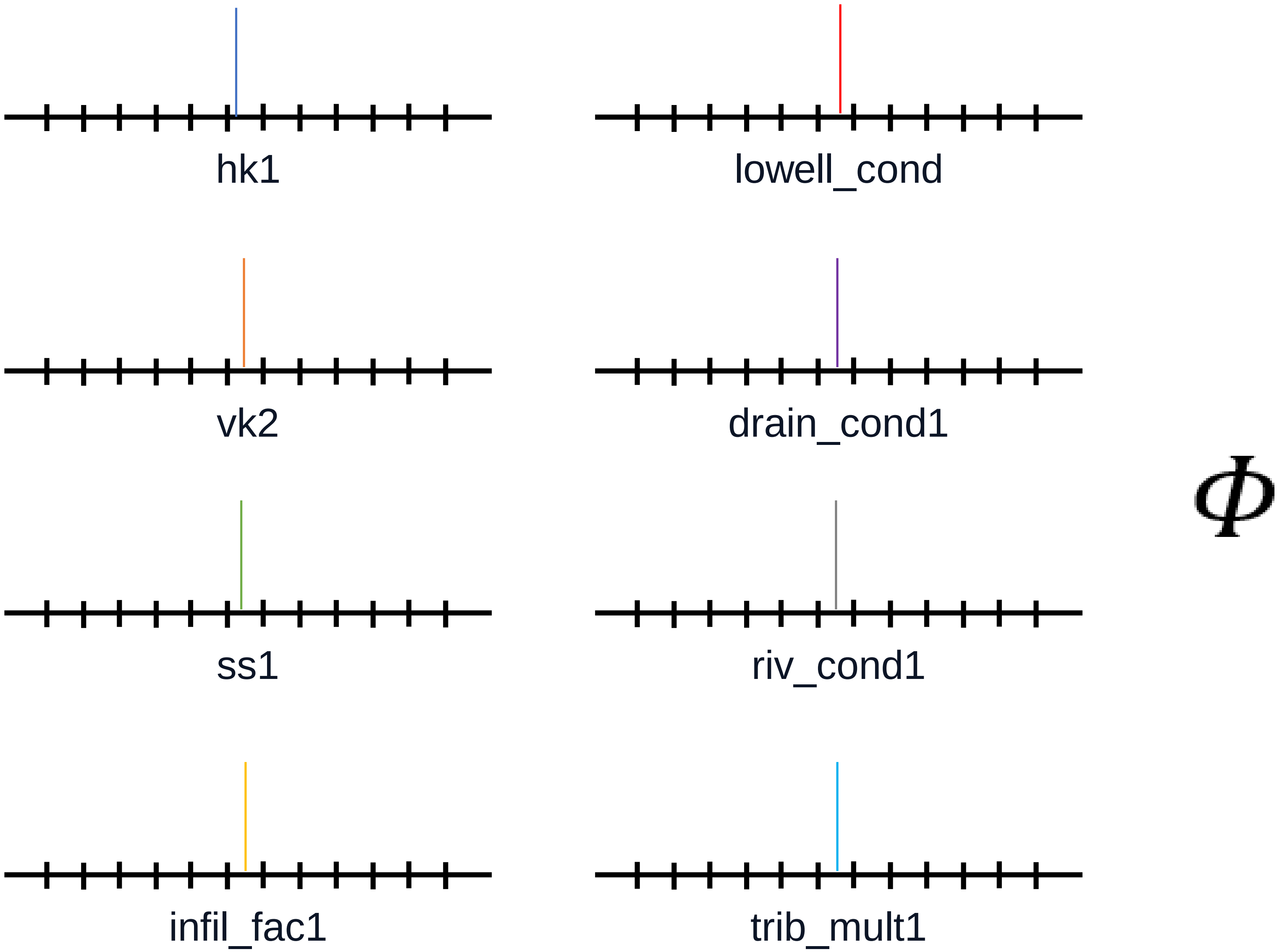


science for a changing world

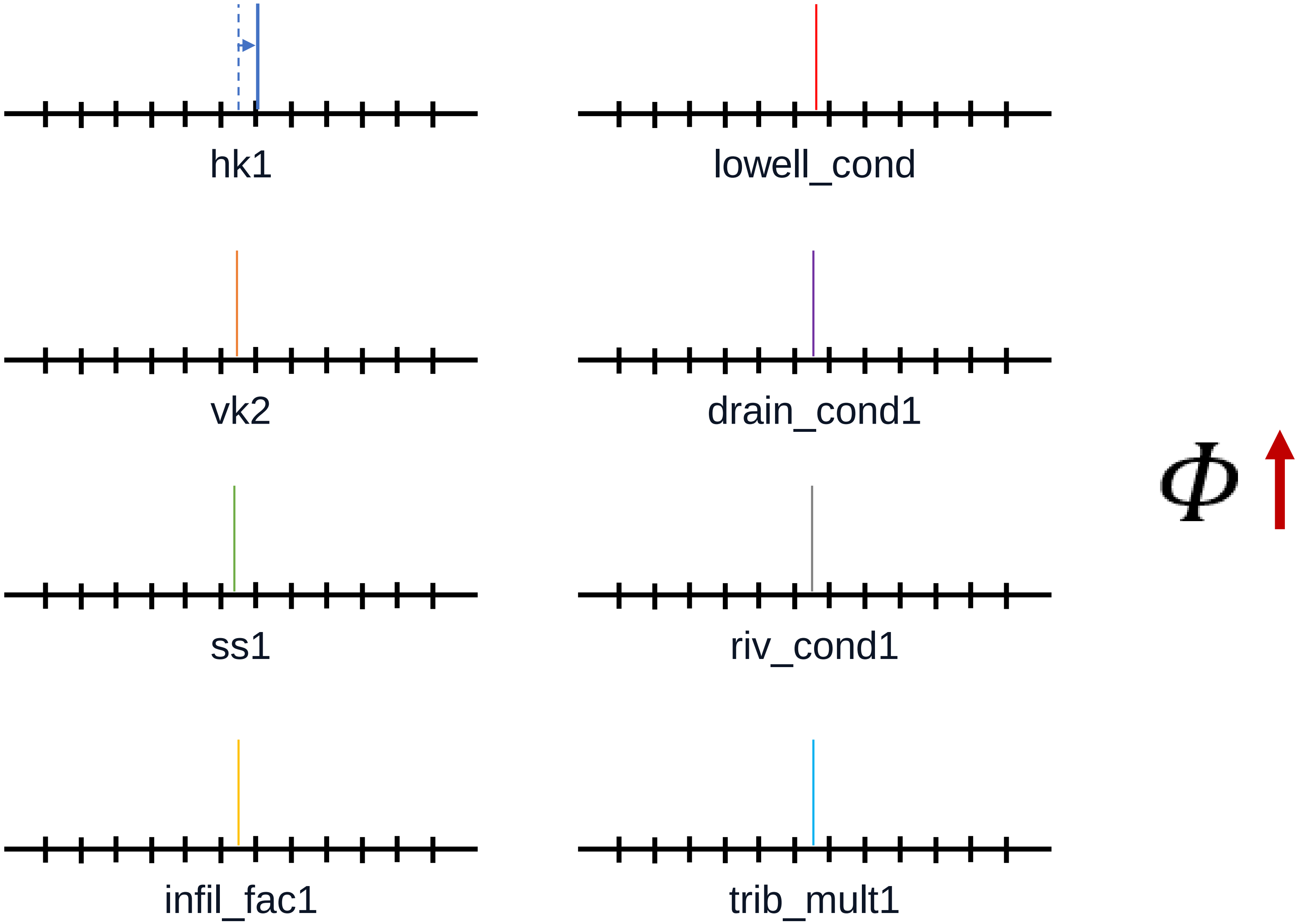
Minimizing Objective Function



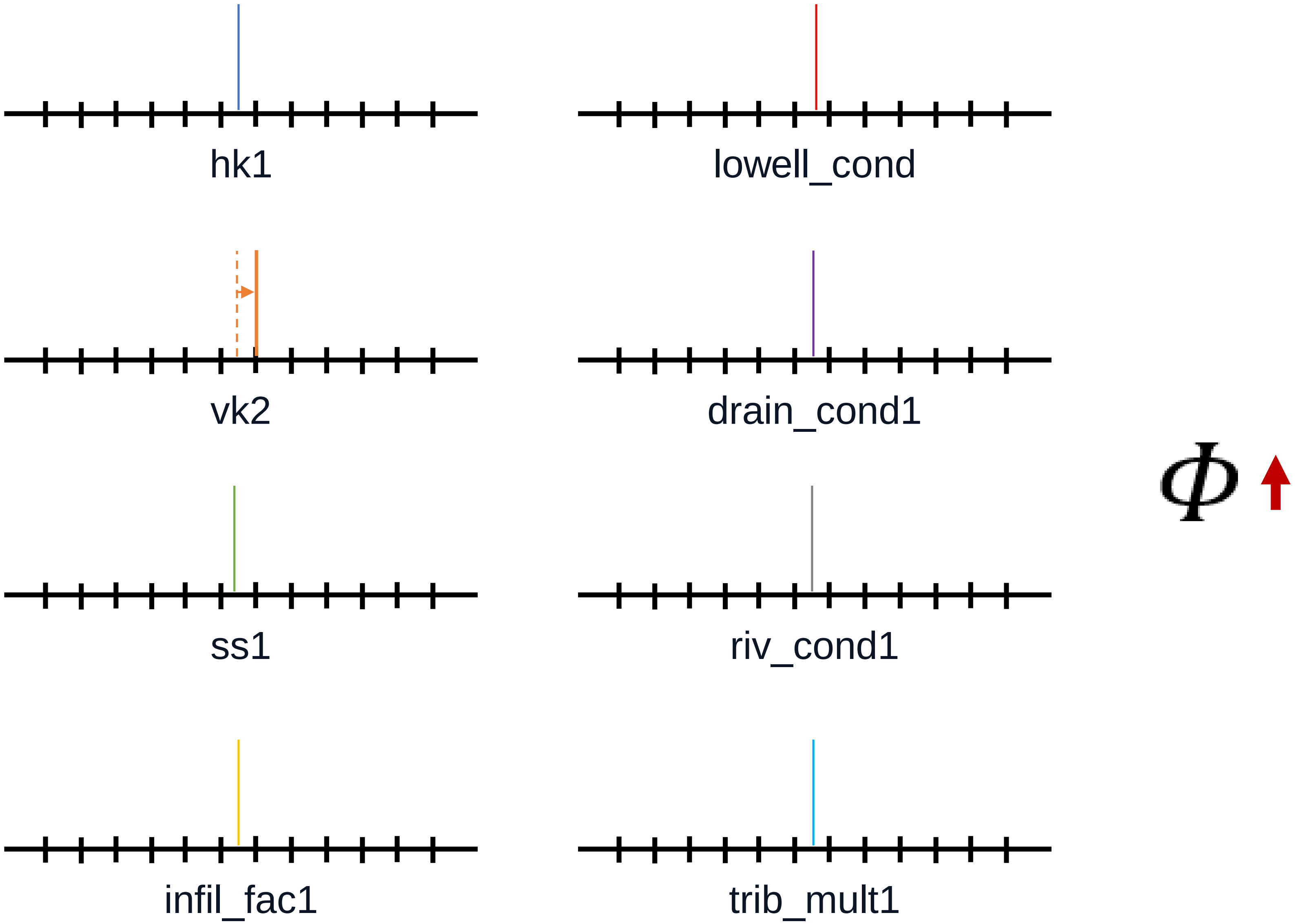
Minimizing Objective Function



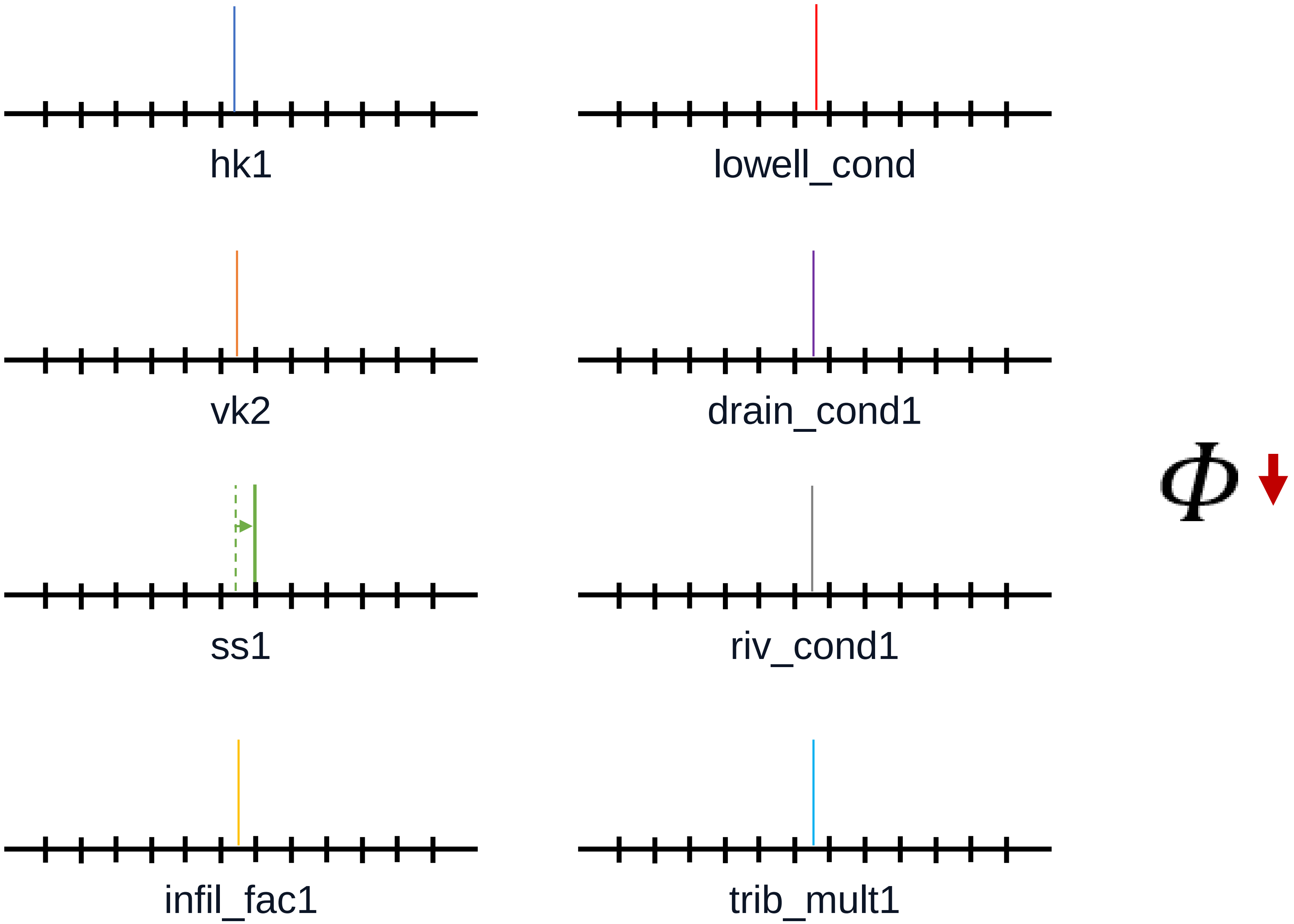
Minimizing Objective Function



Minimizing Objective Function



Minimizing Objective Function



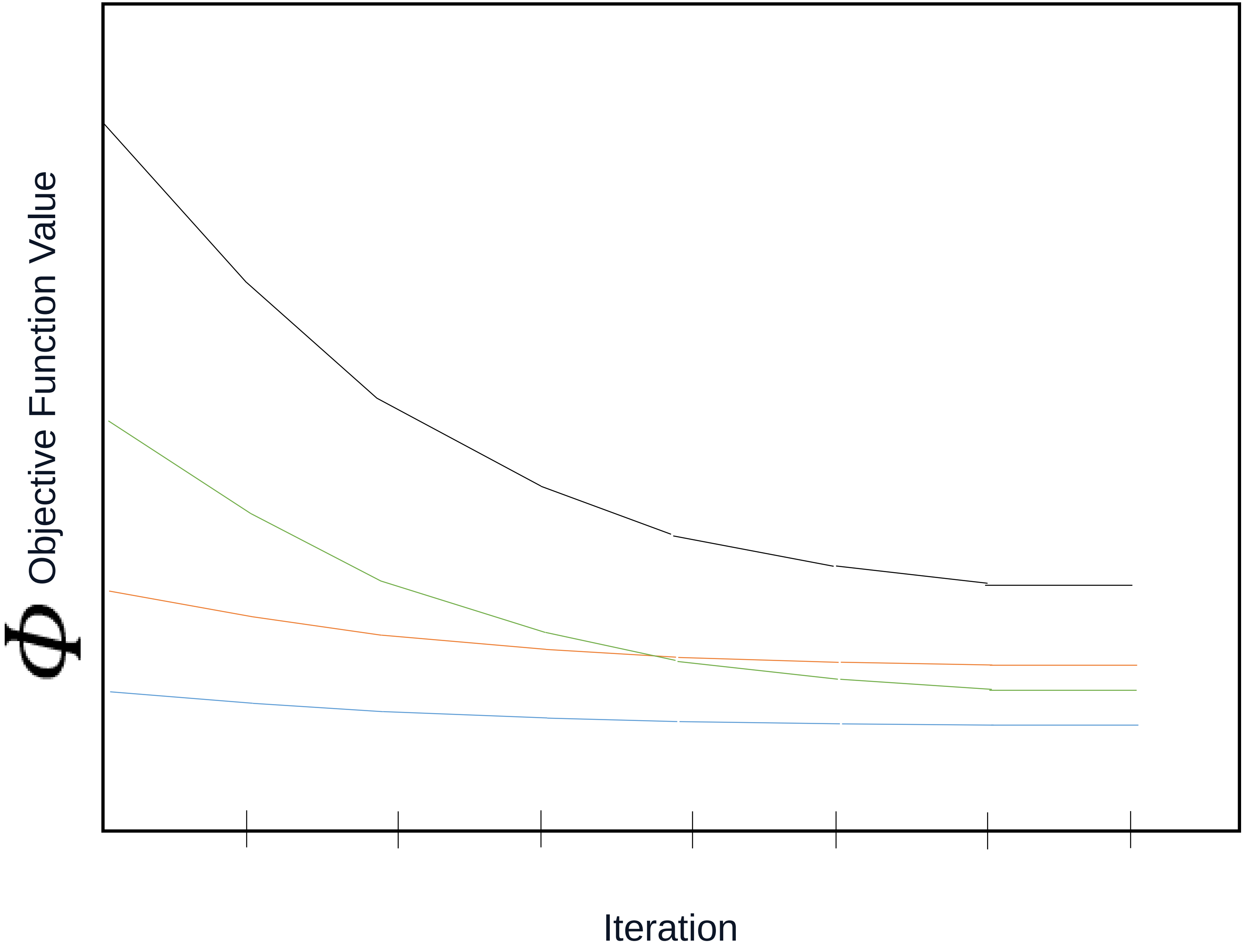
science for a changing world

Minimizing Objective Function

	obs1	obs2	obs3	. . .	
par1	$\frac{\Delta\phi_1}{\Delta par_1}$	$\frac{\Delta\phi_2}{\Delta par_1}$	$\frac{\Delta\phi_3}{\Delta par_1}$		$\frac{\Delta\phi}{\Delta par_1}$
par2	$\frac{\Delta\phi_1}{\Delta par_2}$	$\frac{\Delta\phi_2}{\Delta par_2}$	$\frac{\Delta\phi_3}{\Delta par_2}$		$\frac{\Delta\phi}{\Delta par_2}$
par3	$\frac{\Delta\phi_1}{\Delta par_3}$	$\frac{\Delta\phi_2}{\Delta par_3}$	$\frac{\Delta\phi_3}{\Delta par_3}$	. . .	$\frac{\Delta\phi}{\Delta par_3}$
par4	$\frac{\Delta\phi_1}{\Delta par_4}$	$\frac{\Delta\phi_2}{\Delta par_4}$	$\frac{\Delta\phi_3}{\Delta par_4}$		$\frac{\Delta\phi}{\Delta par_4}$
⋮			⋮		

USGS science for a changing world

Minimizing Objective Function



Common Challenges

long run times

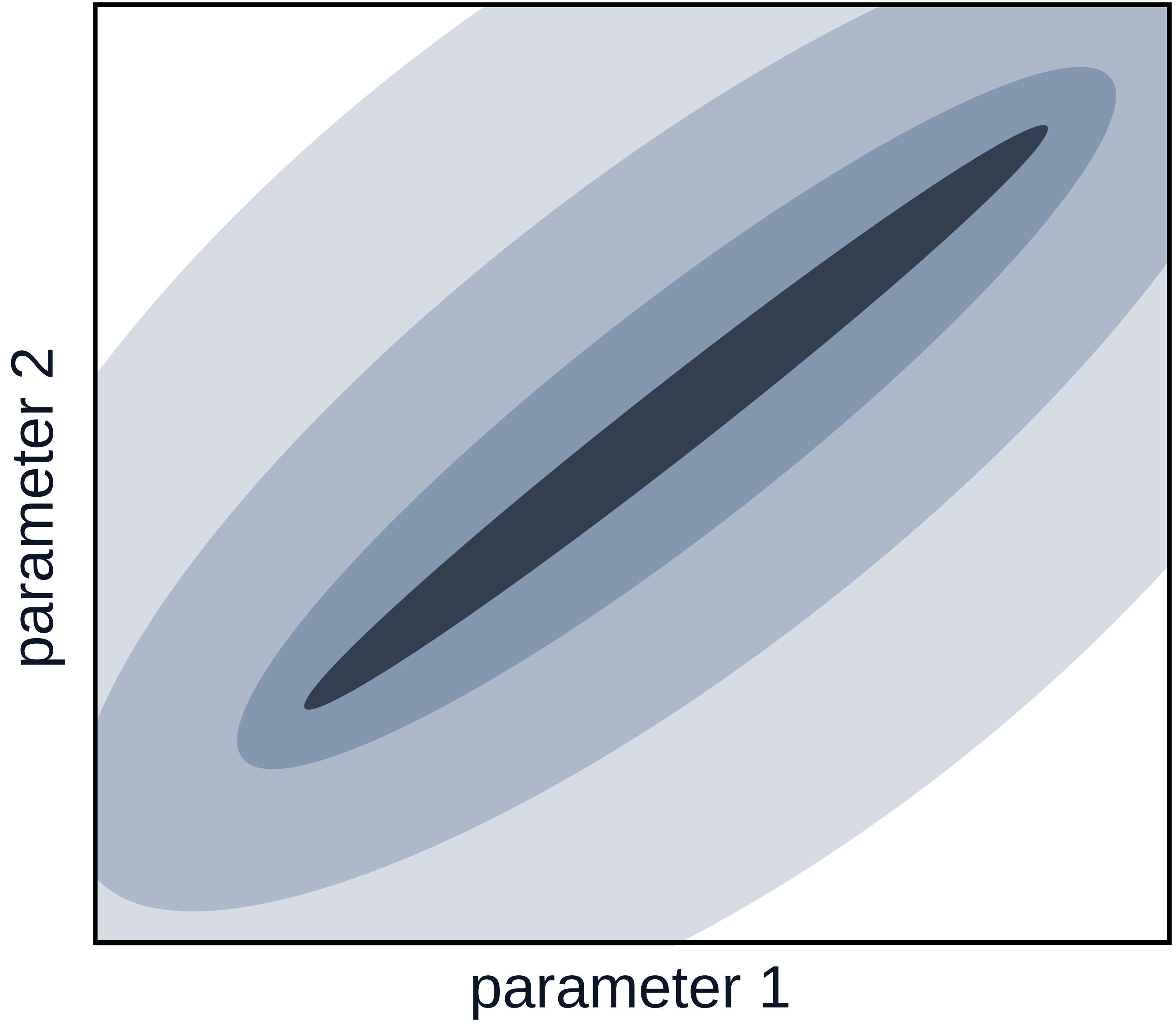
```
COMPUTING JACOBIAN:

  Iteration type: base parameter solution
  calculating jacobian...     running model 1488 times
```

science for a changing world

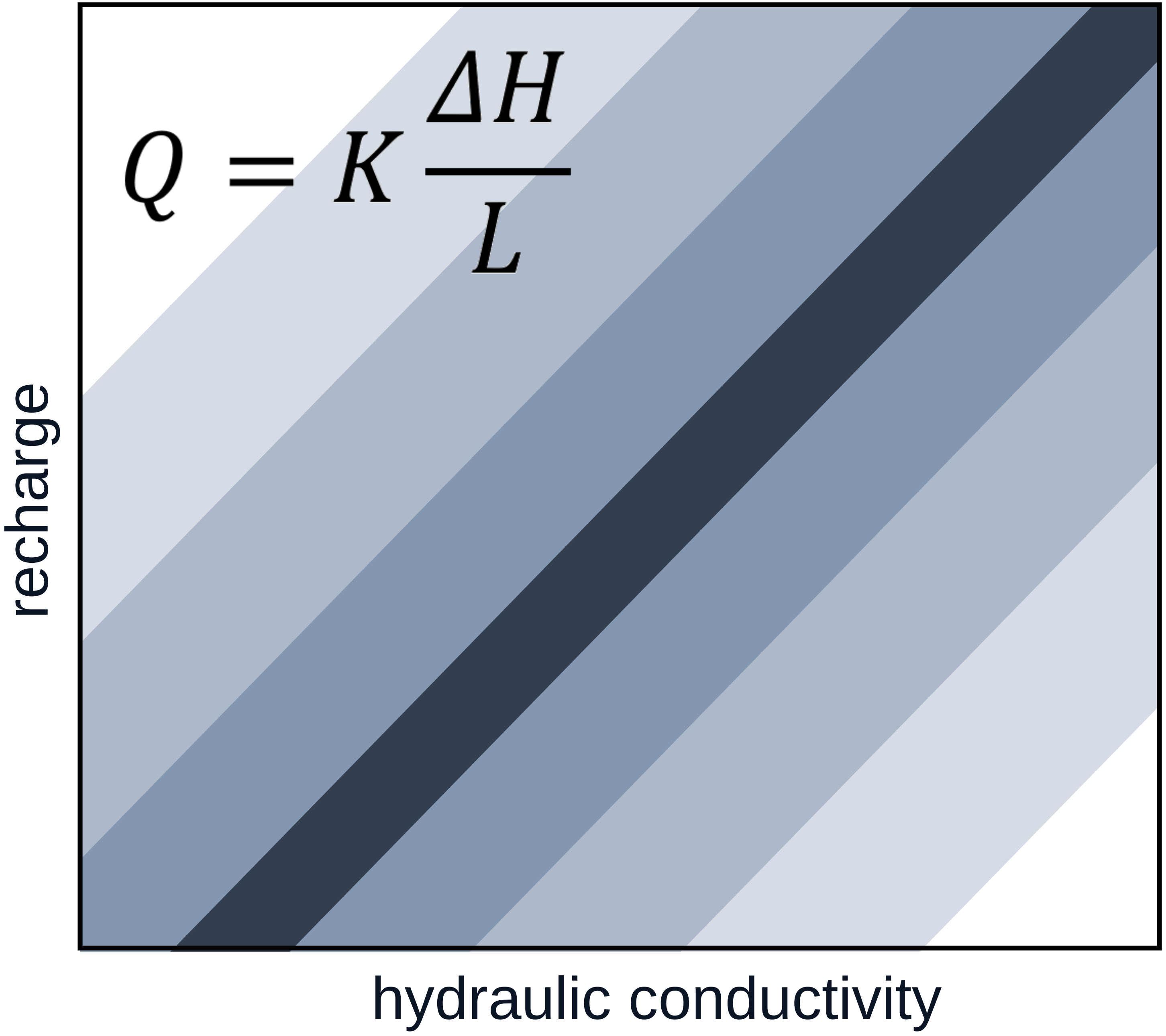
Common Challenges

parameter correlation



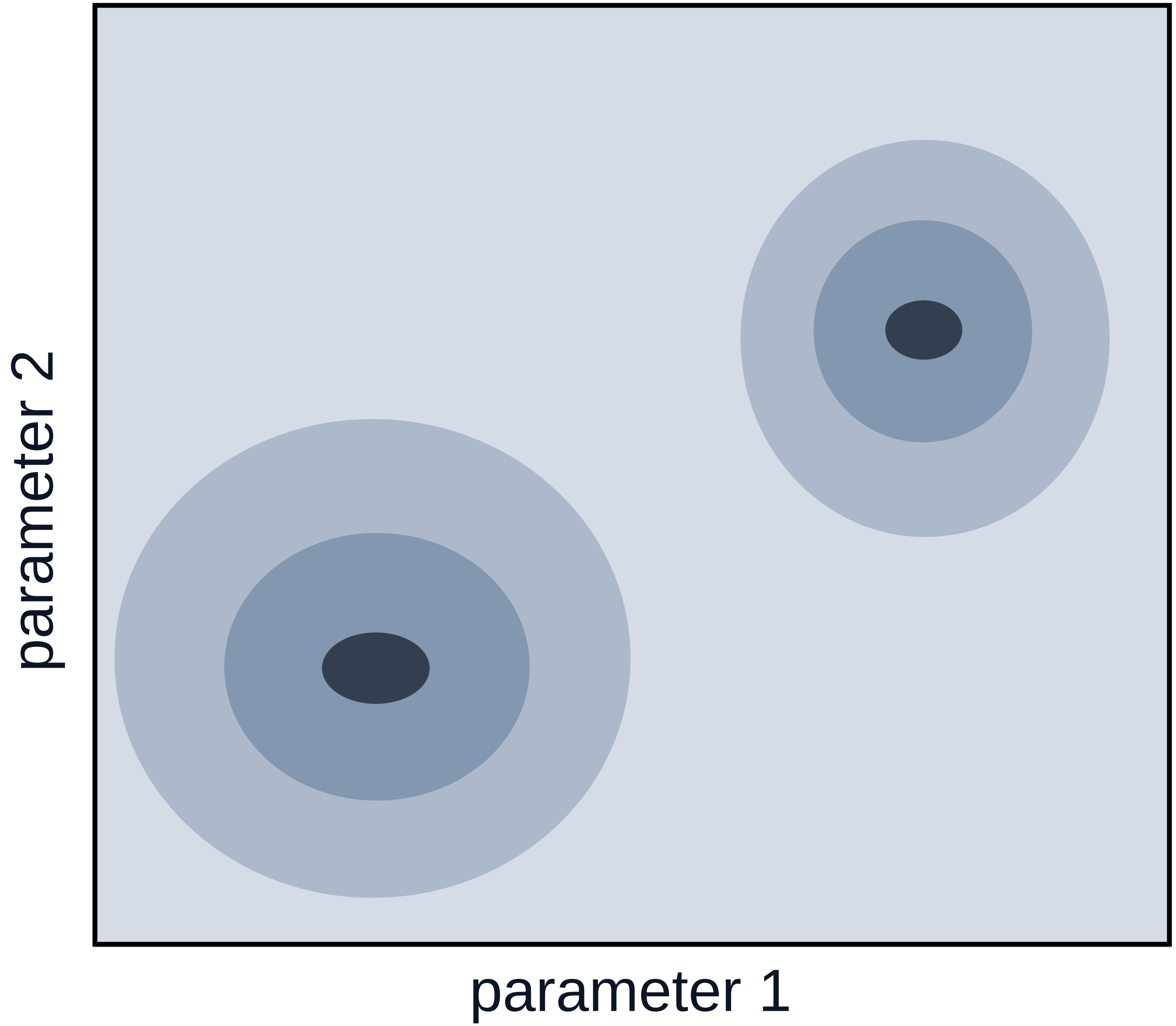
Common Challenges

non-uniqueness



Common Challenges

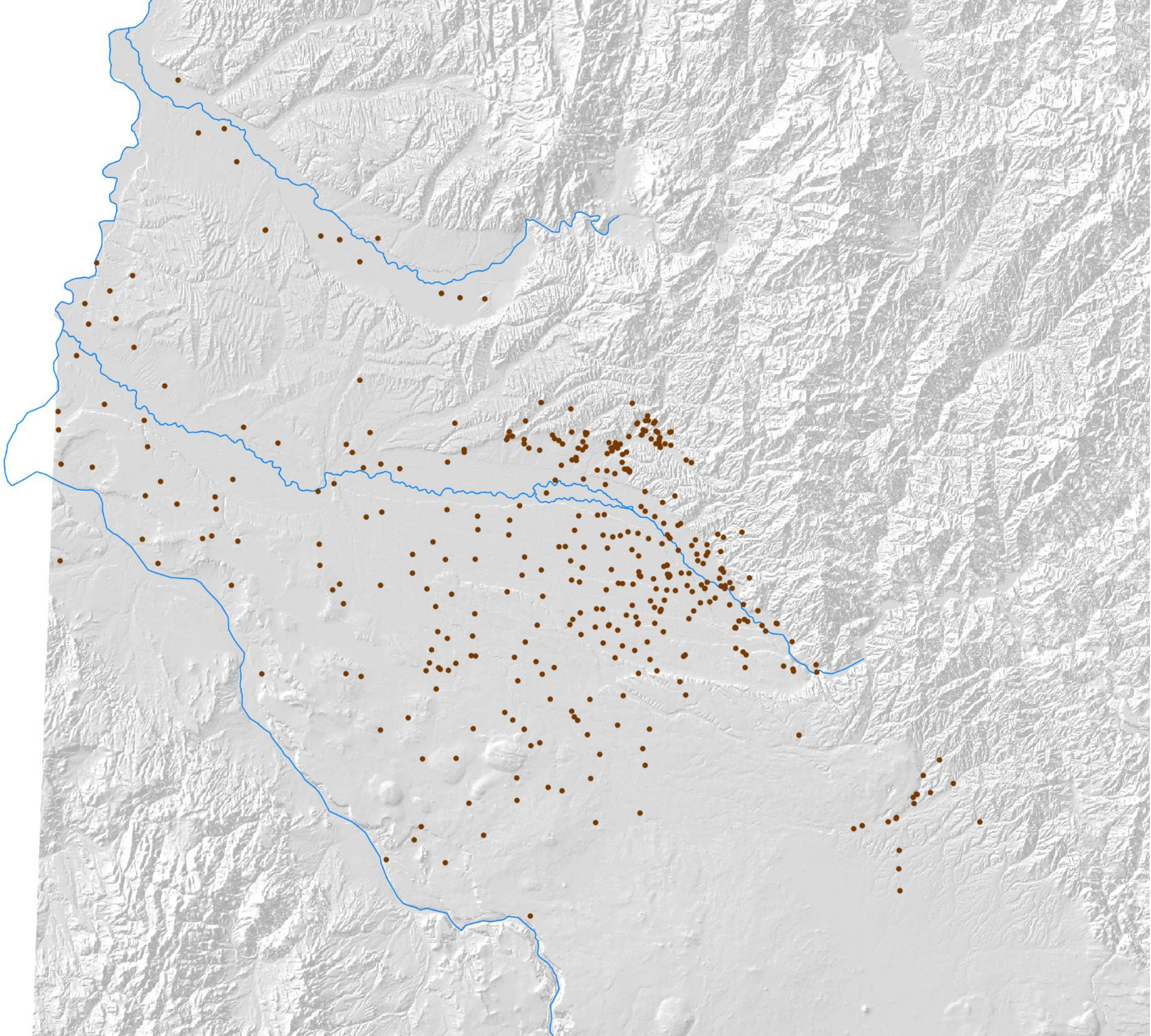
non-uniqueness



USGS
science for a changing world

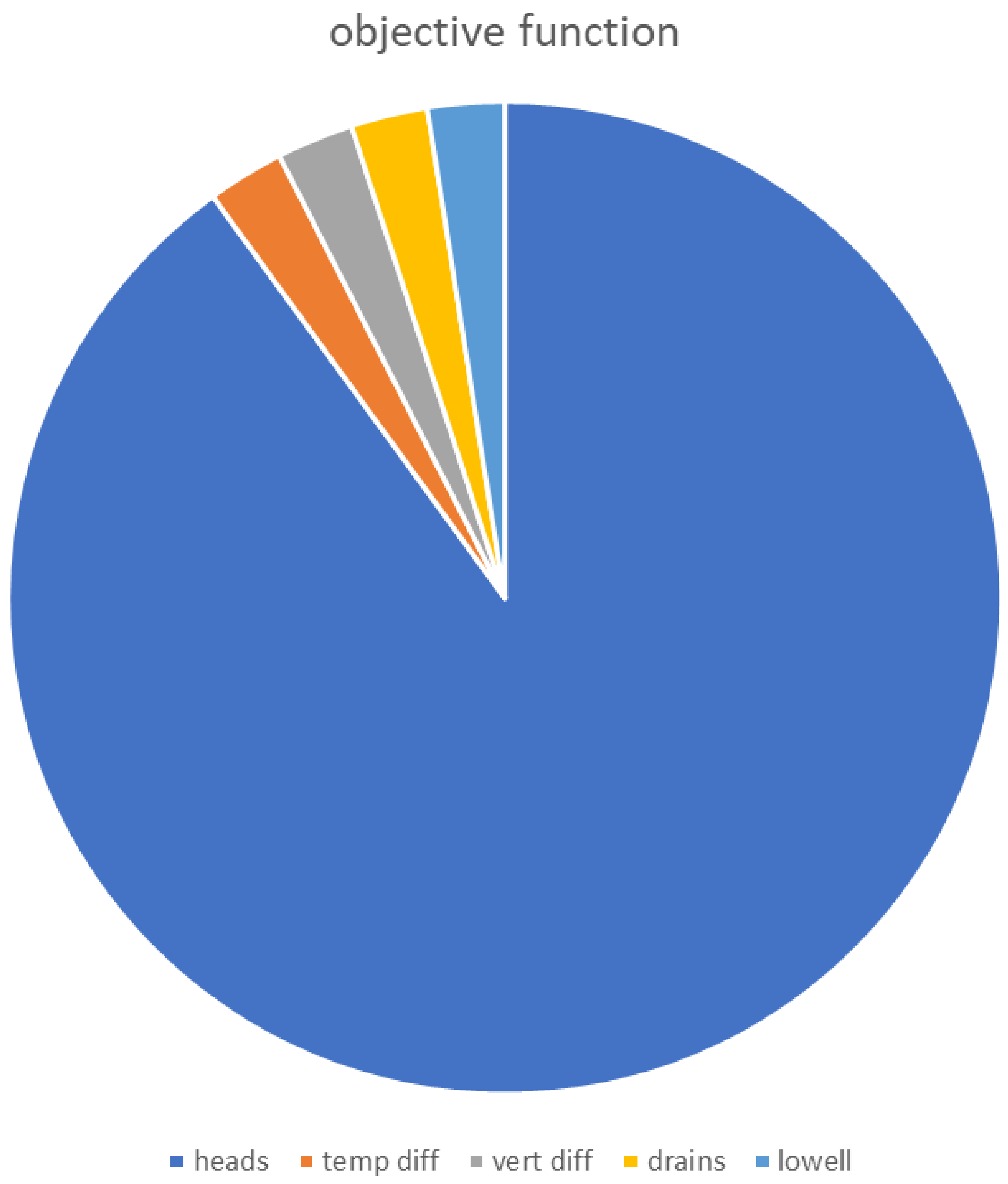
Common Challenges

weighting



Common Challenges

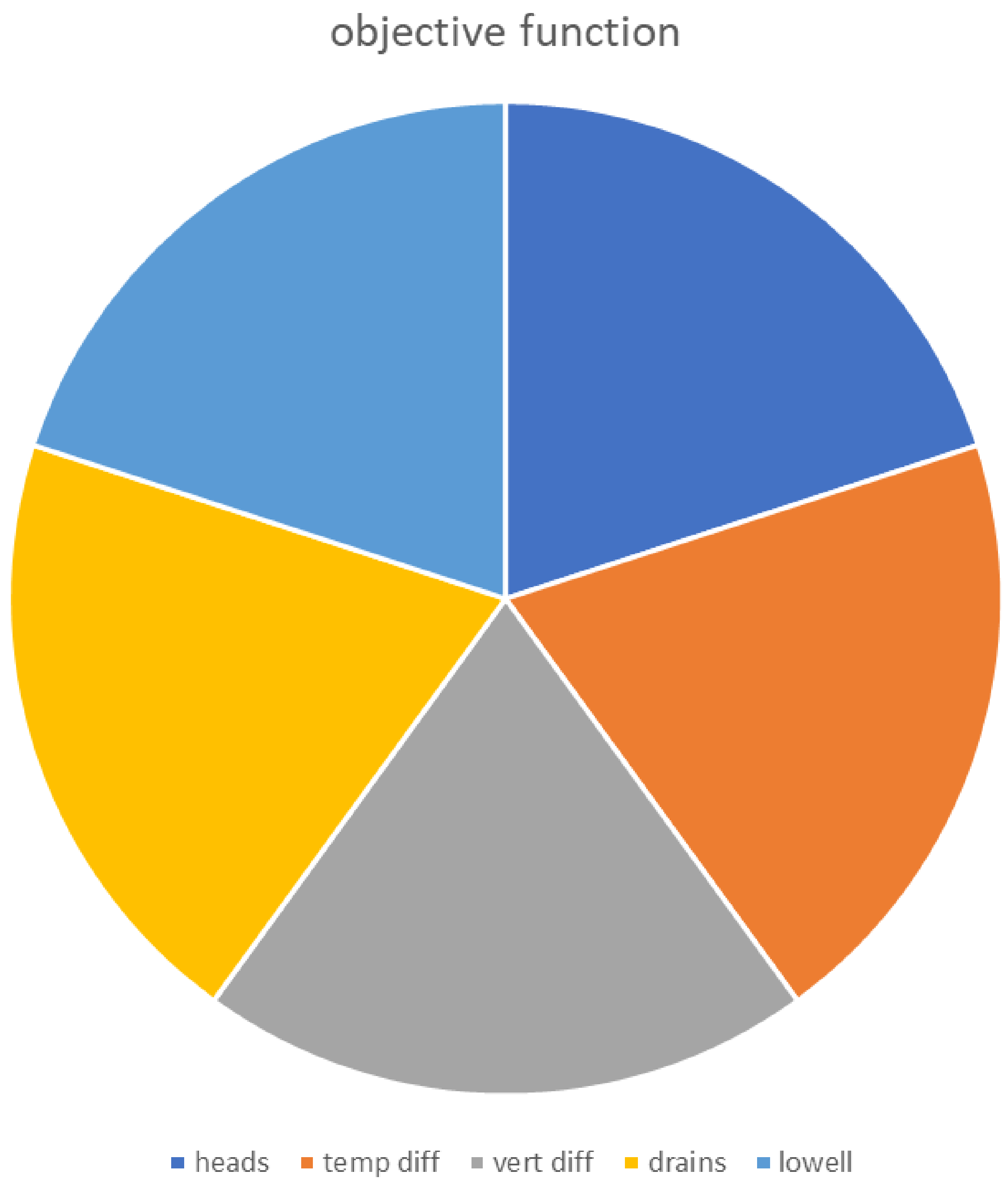
weighting



science for a changing world

Common Challenges

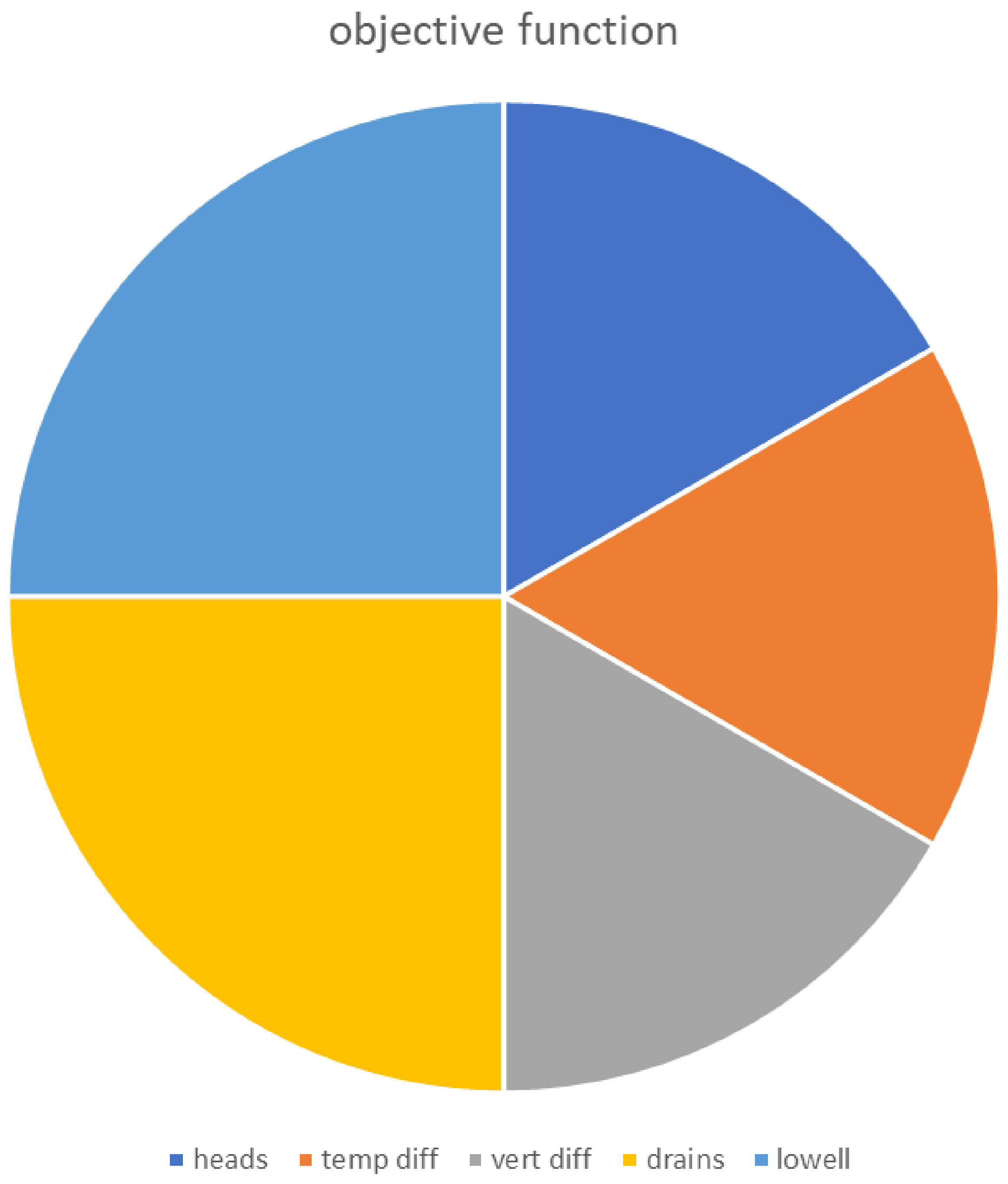
weighting



USGS
science for a changing world

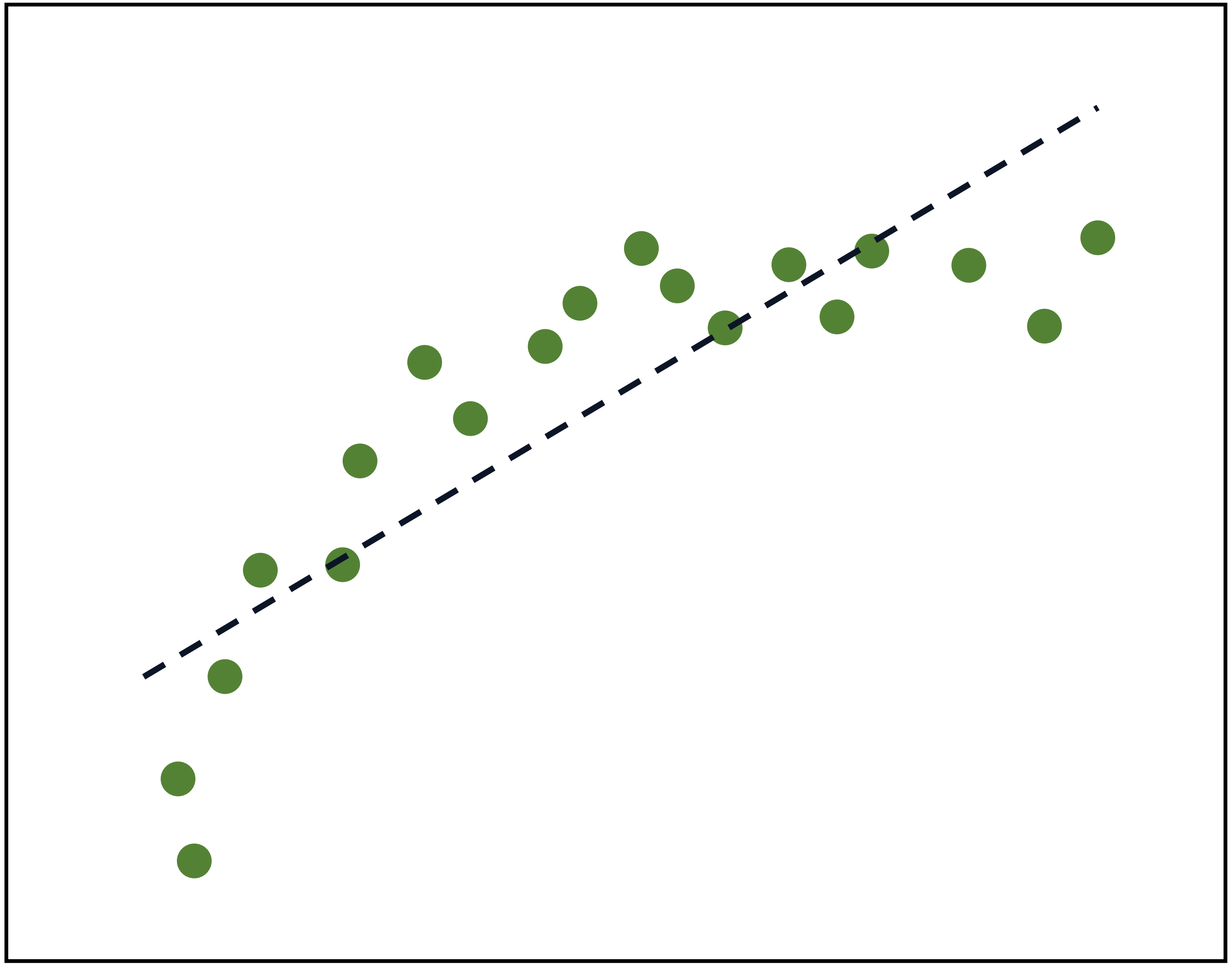
Common Challenges

weighting



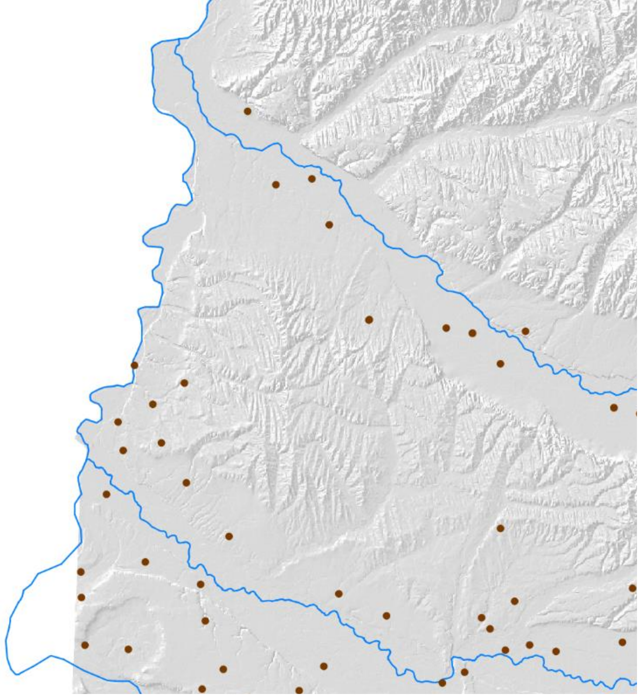
Common Challenges

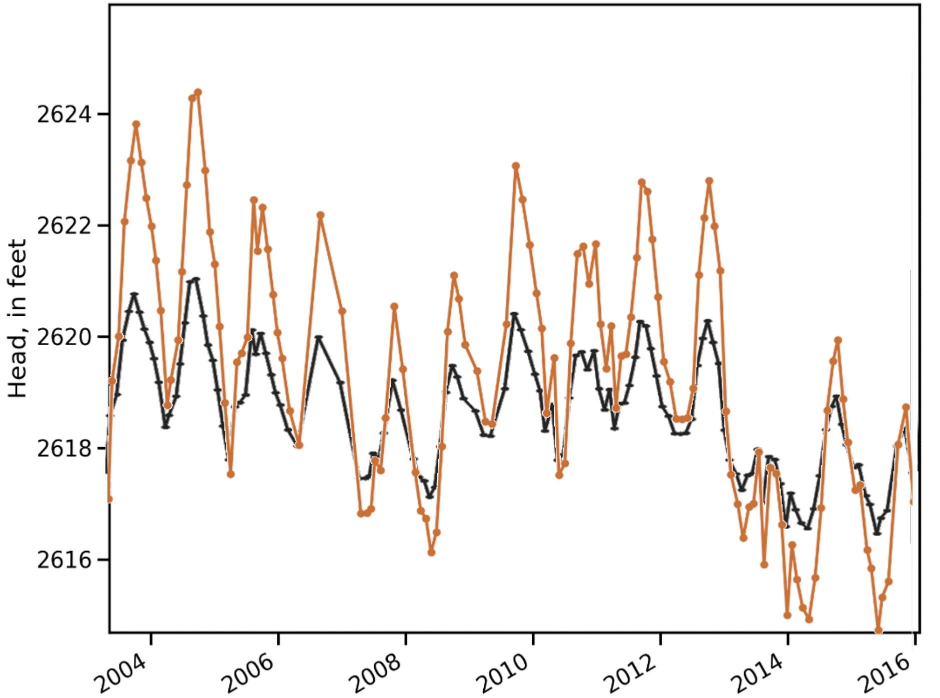
model inadequacy

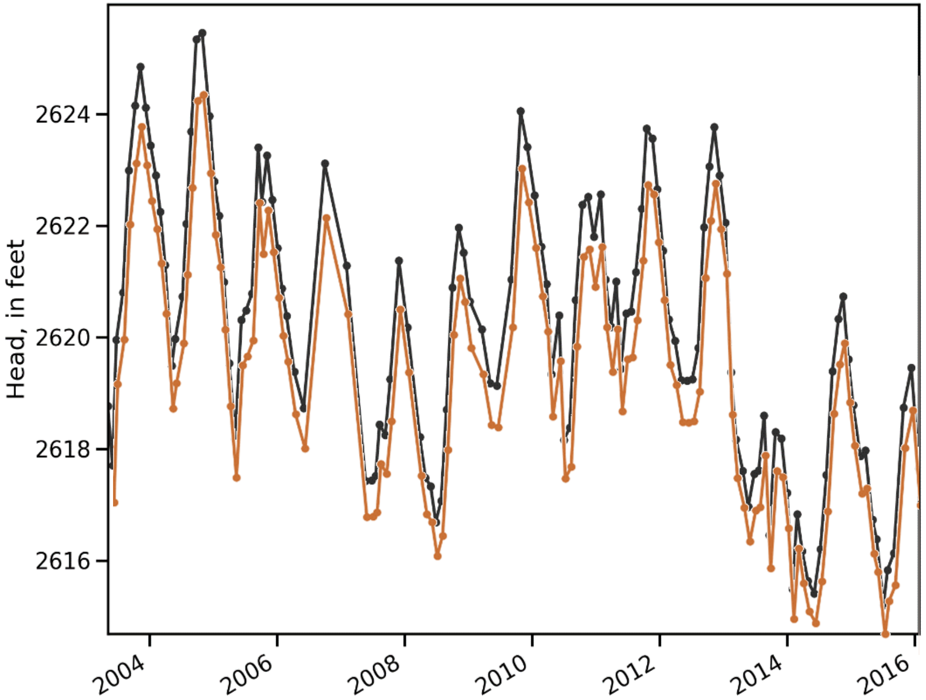


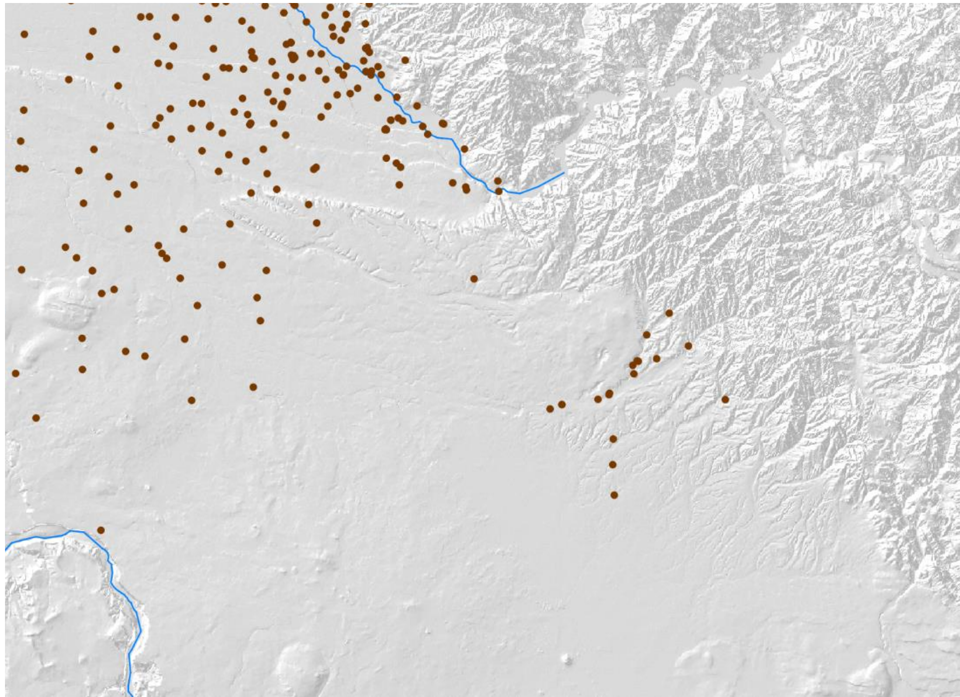
Common Challenges

systematic misfits or tradeoffs





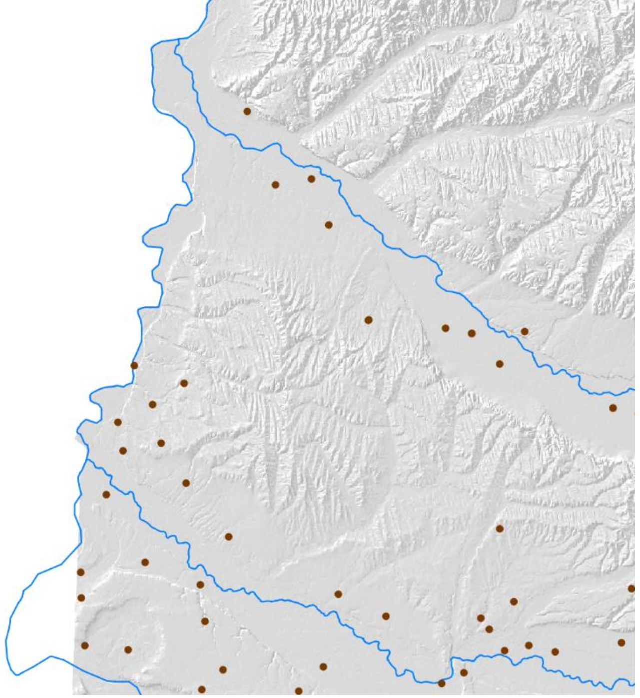


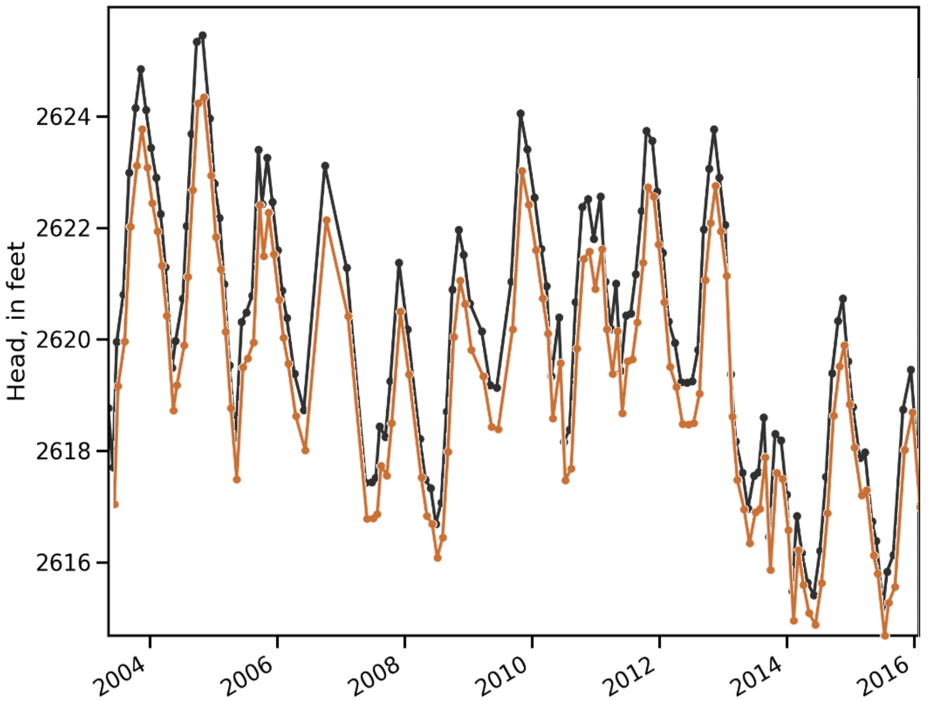


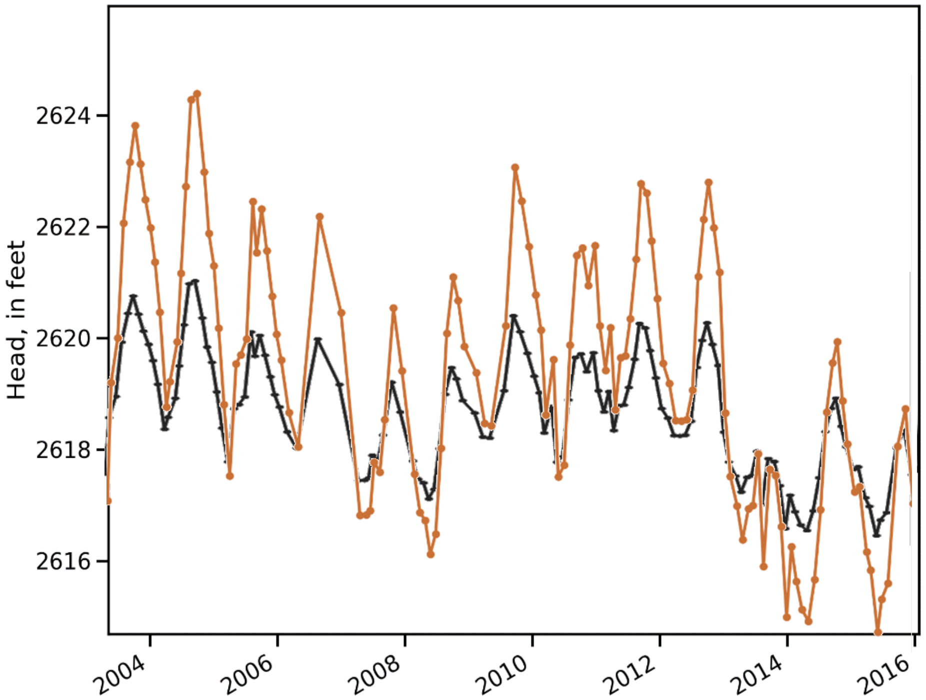
USGS
science for a changing world

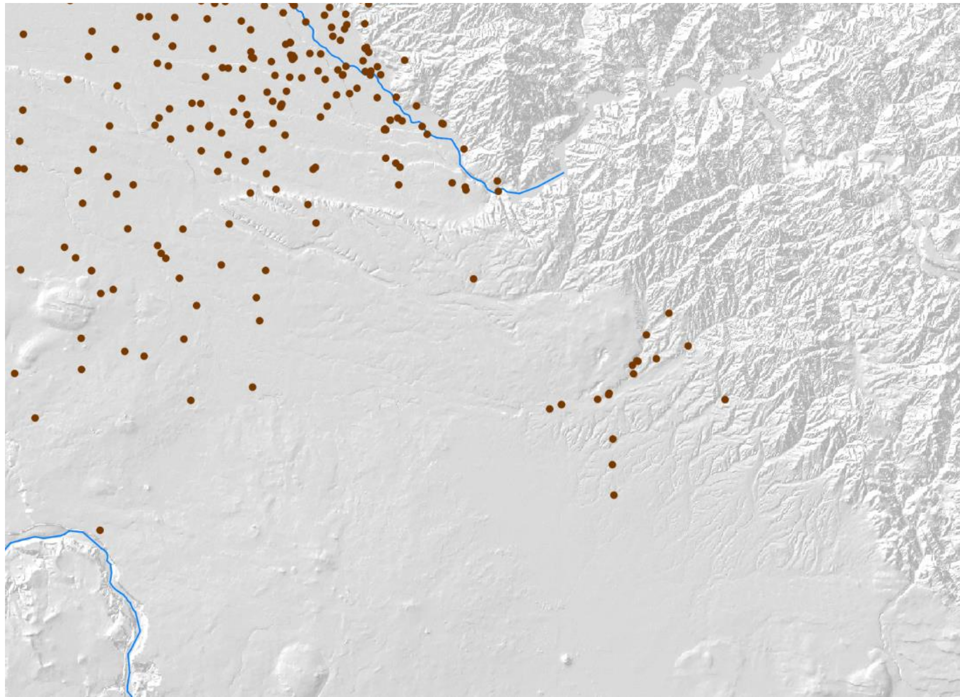
Common Challenges

systematic misfits or tradeoffs





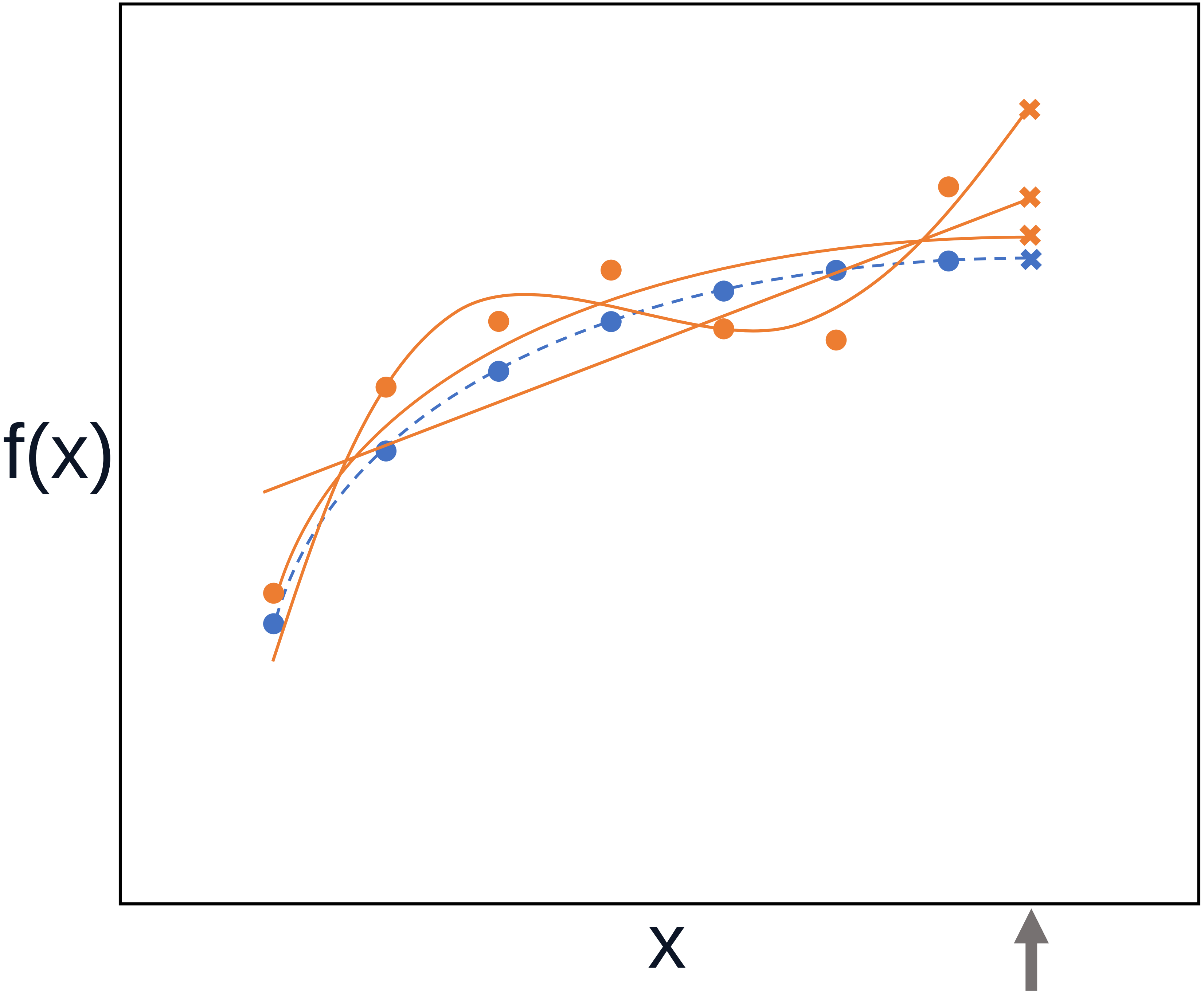




USGS
science for a changing world

Common Challenges

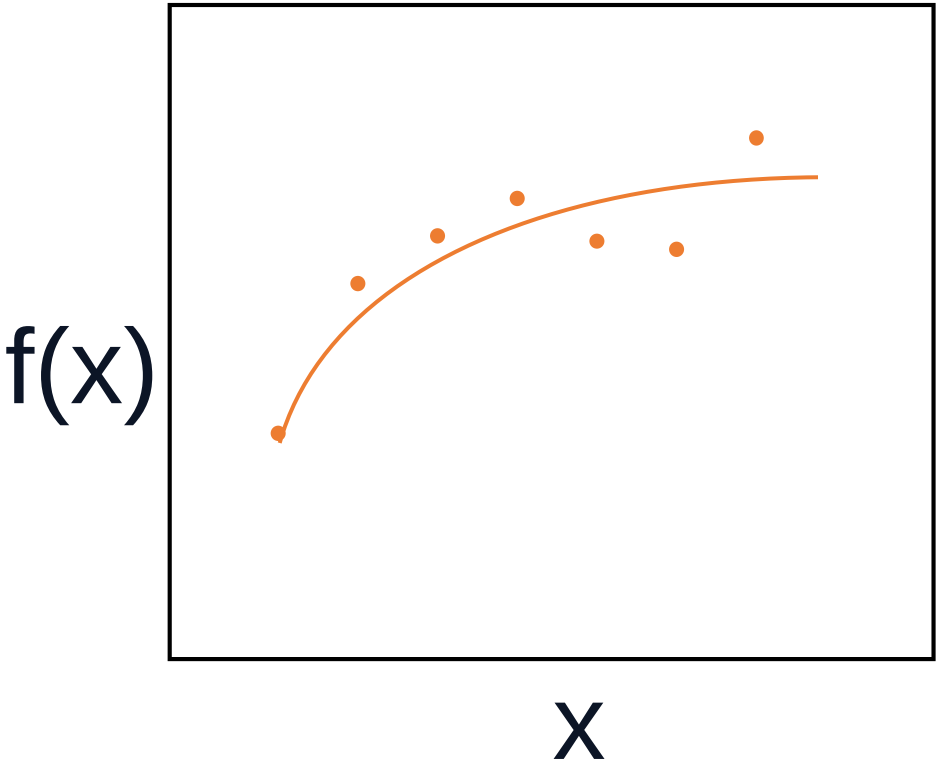
overfitting

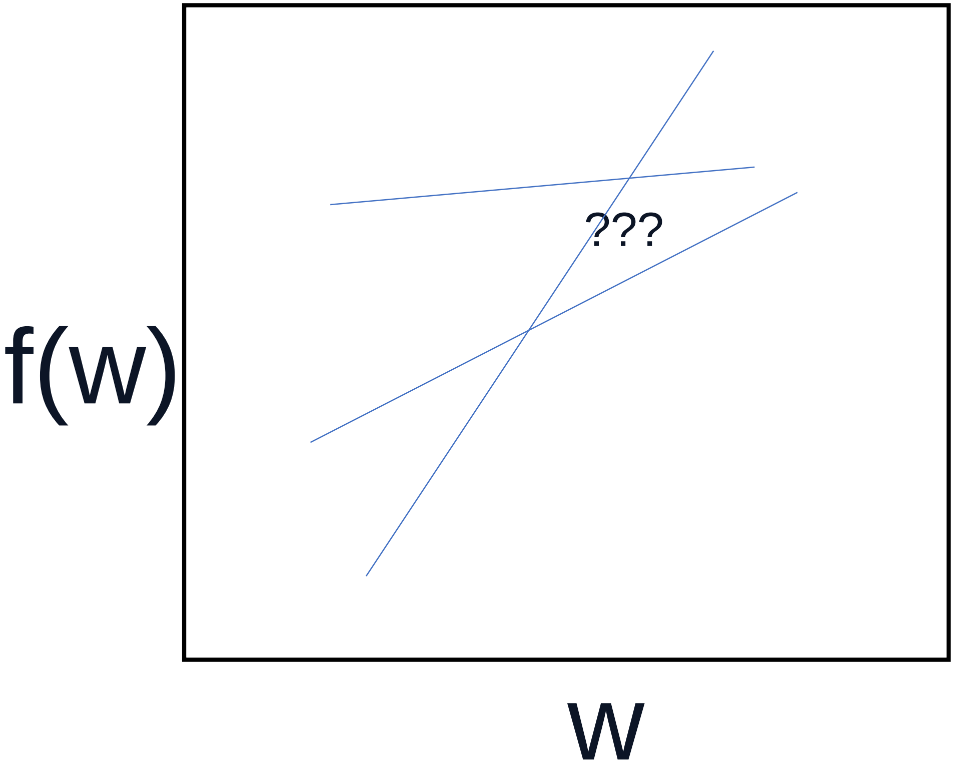


![USGS science for a changing world]

Common Challenges

scenario forecasts uninformed by data

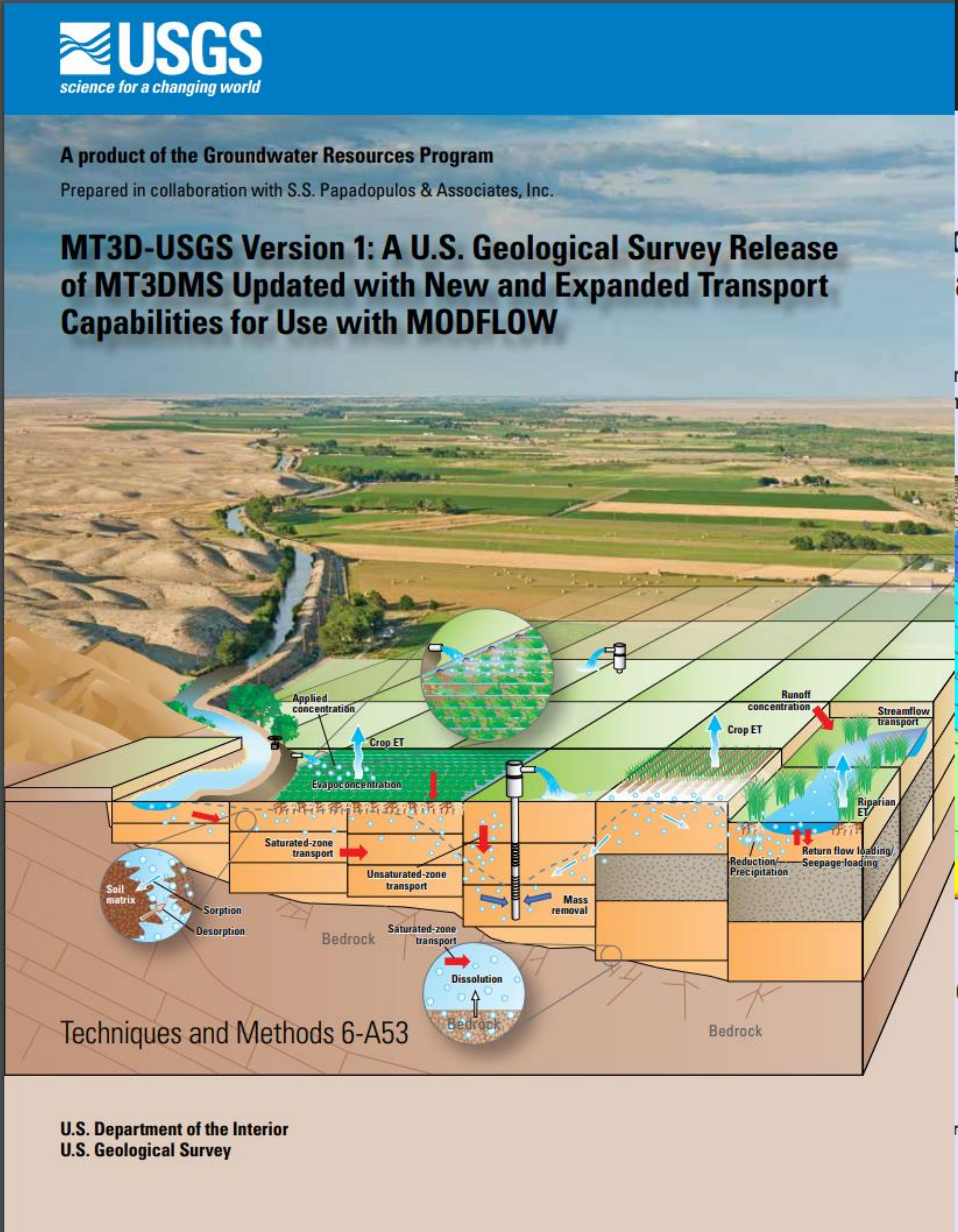


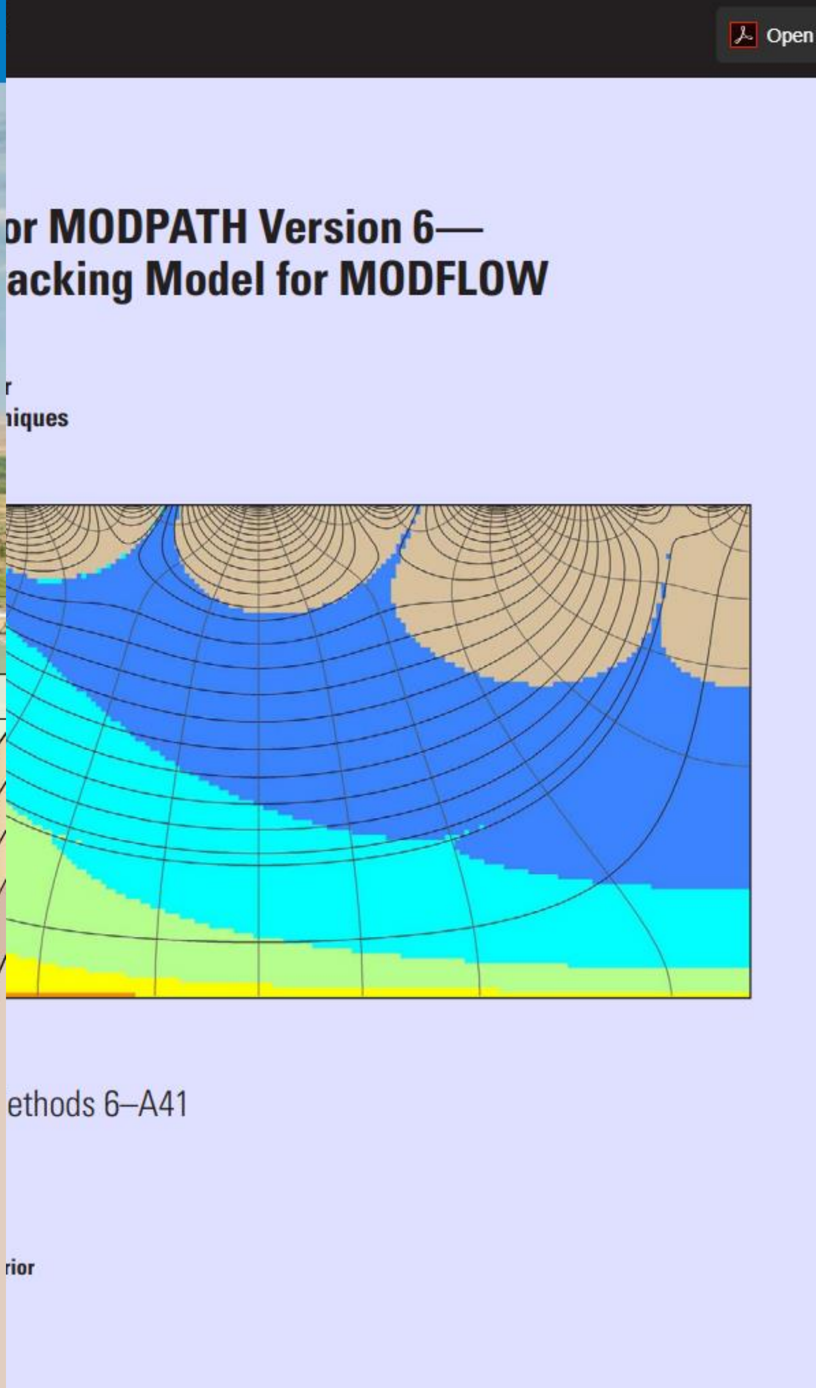


USGS
science for a changing world

Common Challenges

scenario forecasts uninformed by data





Implementation

PEST(++)

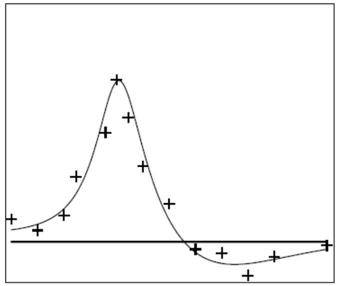
PEST

Model-Independent Parameter Estimation

User Manual Part I:

PEST, SENSAN and Global Optimisers

(See part II for documentation of PEST support and uncertainty analysis utilities.)



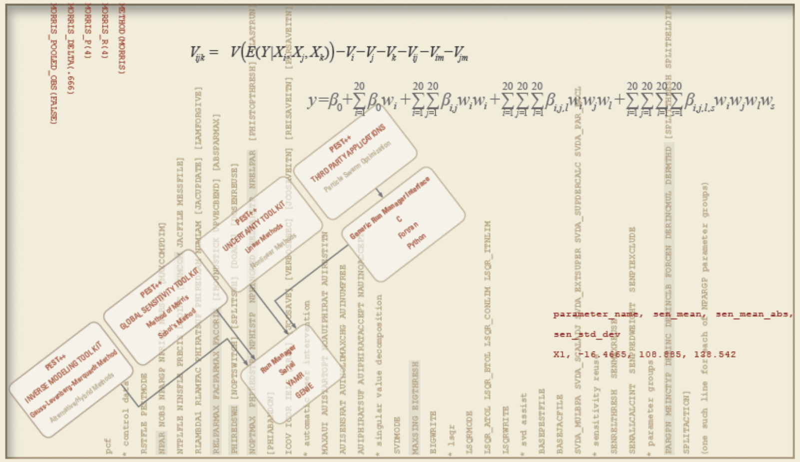
7th Edition published in 2018
Latest additions: May 2019

Watermark Numerical Computing

USGS
science for a changing world

Groundwater Resources Program
Prepared in cooperation with U.S. Environmental Protection Agency,
Great Lakes Restoration Initiative

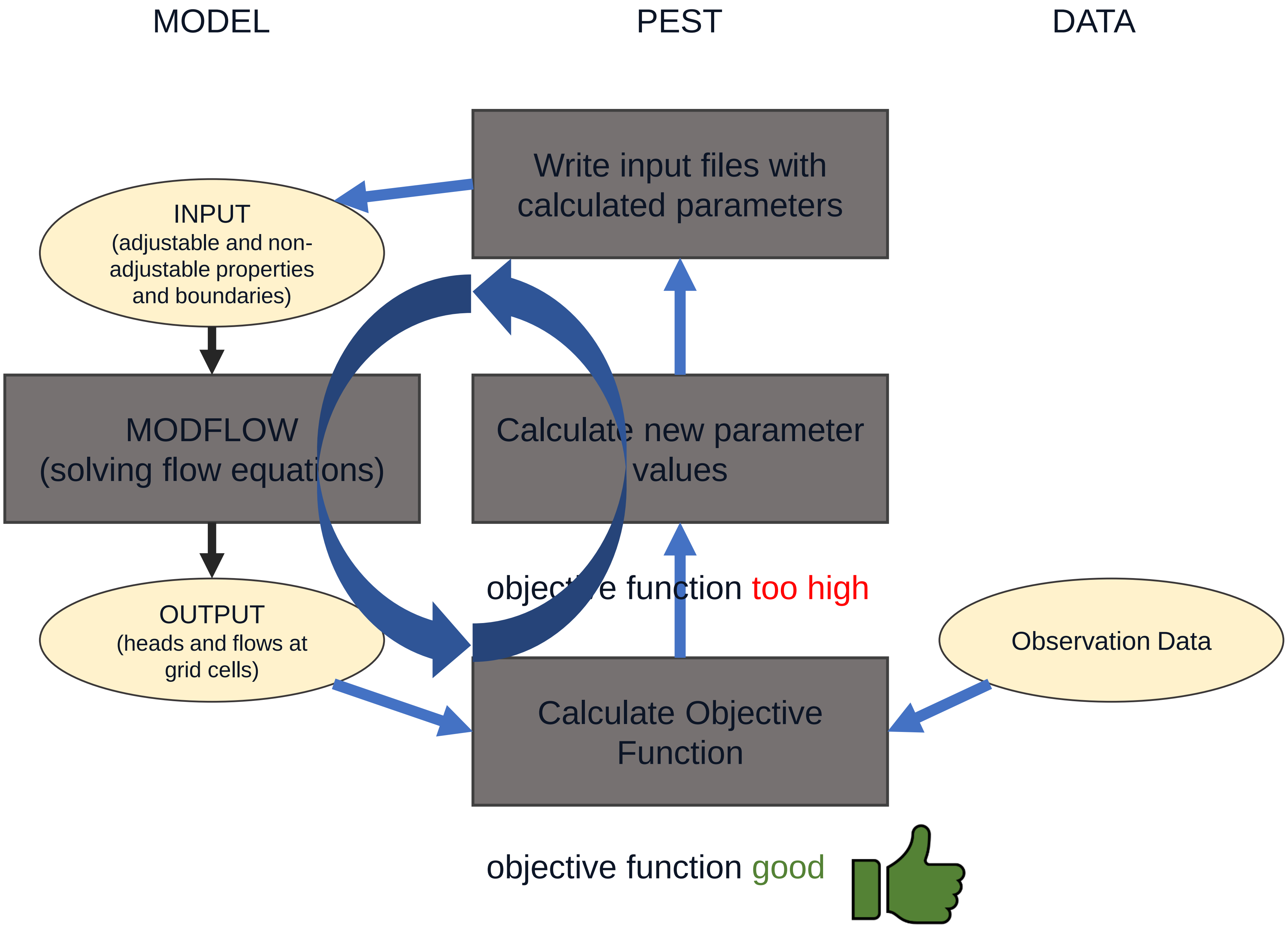
Approaches in Highly Parameterized Inversion: PEST++ Version 3, A Parameter ESTimation and Uncertainty Analysis Software Suite Optimized for Large Environmental Models



Techniques and Methods 7–C12

U.S. Department of the Interior
U.S. Geological Survey

Wrapping Model



USGS
science for a changing world

Techniques

many settings for tweaking parameter estimation

Tikhonov Regularization

Singular Value Decomposition

SVD-Assist

Iterative Ensemble Smoother

USGS
science for a changing world

High-throughput computing

```
                   Generalized Run Manager Interface
                                   by:

                     The PEST++ Development Team


             starting PANTHER master...

IP addresses:
  0.0.0.0:4004 (IPv4)

PANTHER master listening on socket: 0.0.0.0:4004 (IPv4)


COMPUTING JACOBIAN:

  Iteration type: base parameter solution
  calculating jacobian...     running model 1488 times

      waiting for agents to appear...


PANTHER progress
   runs(C = completed | F = failed | T = timed out)
   agents(R = running | W = waiting | U = unavailable)
------------------------------------------------------------------------------
05/28 12:26:25 runs(C=240   | F=0    | T=0    ): agents(R=16   | W=0    | U=0    )
```

science for a changing world

High-throughput computing

HPC SYSTEMS

Machine Access

General Purpose HPC
- Good place to start
- CPU and GPU
- 143 nodes
- 3,728 CPU, 56,596 CUDA
- 100 Tflop/s

Yeti

Flagship System
- Large-scale Models
- CPU Only
- 232 nodes
- 9,280 CPU (18,560 hyper-threads)
- 448 Tflop/s

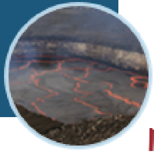
Denali

Prototype System
- ML and Analytics at scale
- Built-in Software Stack
- 22 nodes, 792 CPU
- 122, 800 CUDA
- 15,360 Tensor

Tallgrass

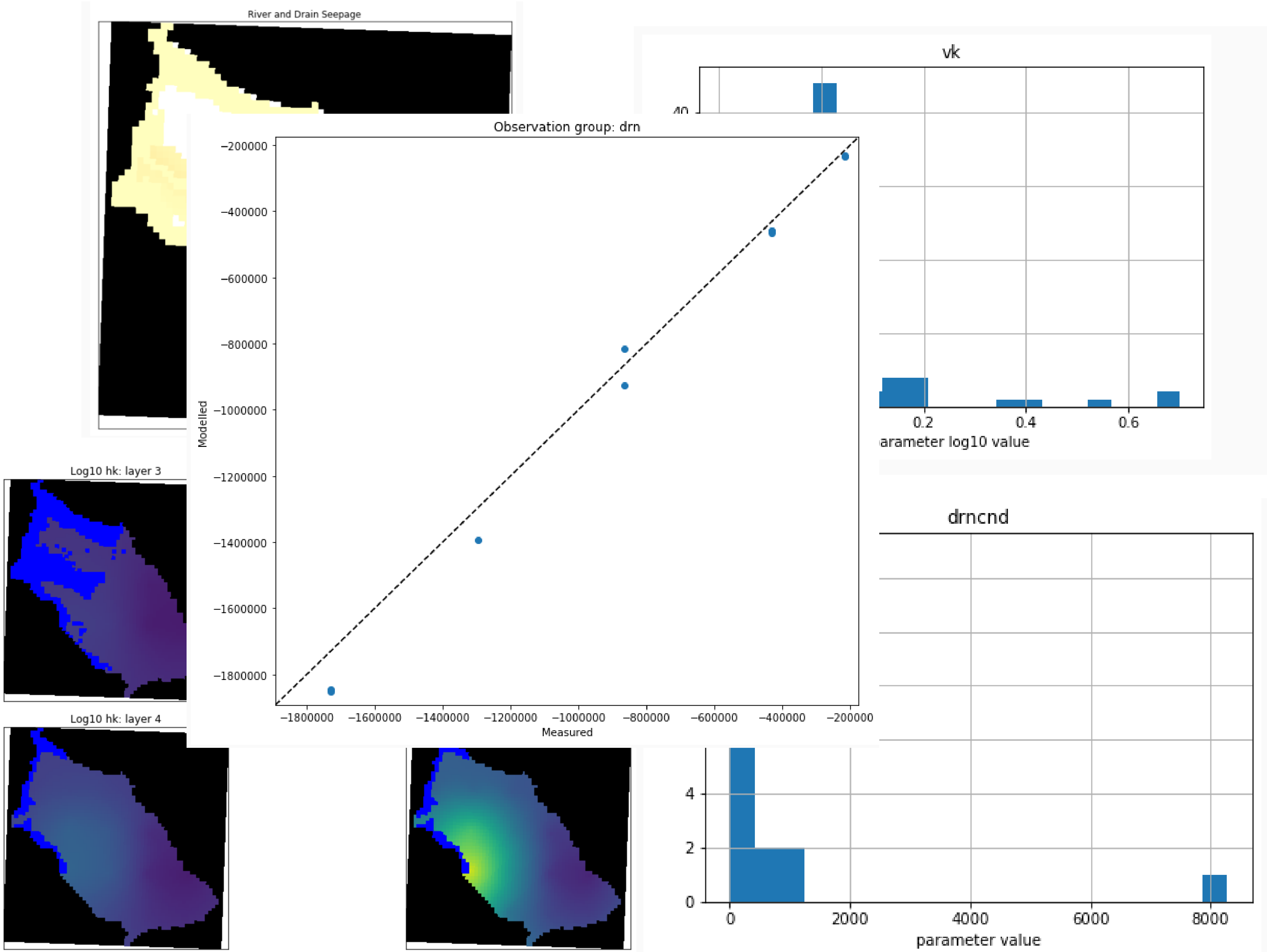
Tiered Storage System
- 4.2 PB Total
- Connects to Denali & Tallgrass
- Provide tiers of storage (high performance & object)

Caldera



USGS
science for a changing world

Iterate


River and Drain Seepage
Observation group: drn
Modelled
Measured
−200000
−400000
−600000
−800000
−1000000
−1200000
−1400000
−1600000
−1800000
−1800000 −1600000 −1400000 −1200000 −1000000 −800000 −600000 −400000 −200000
vk
0.2 0.4 0.6
arameter log10 value
drncnd
0 2000 4000 6000 8000
parameter value
0 2 4
Log10 hk: layer 3
Log10 hk: layer 4

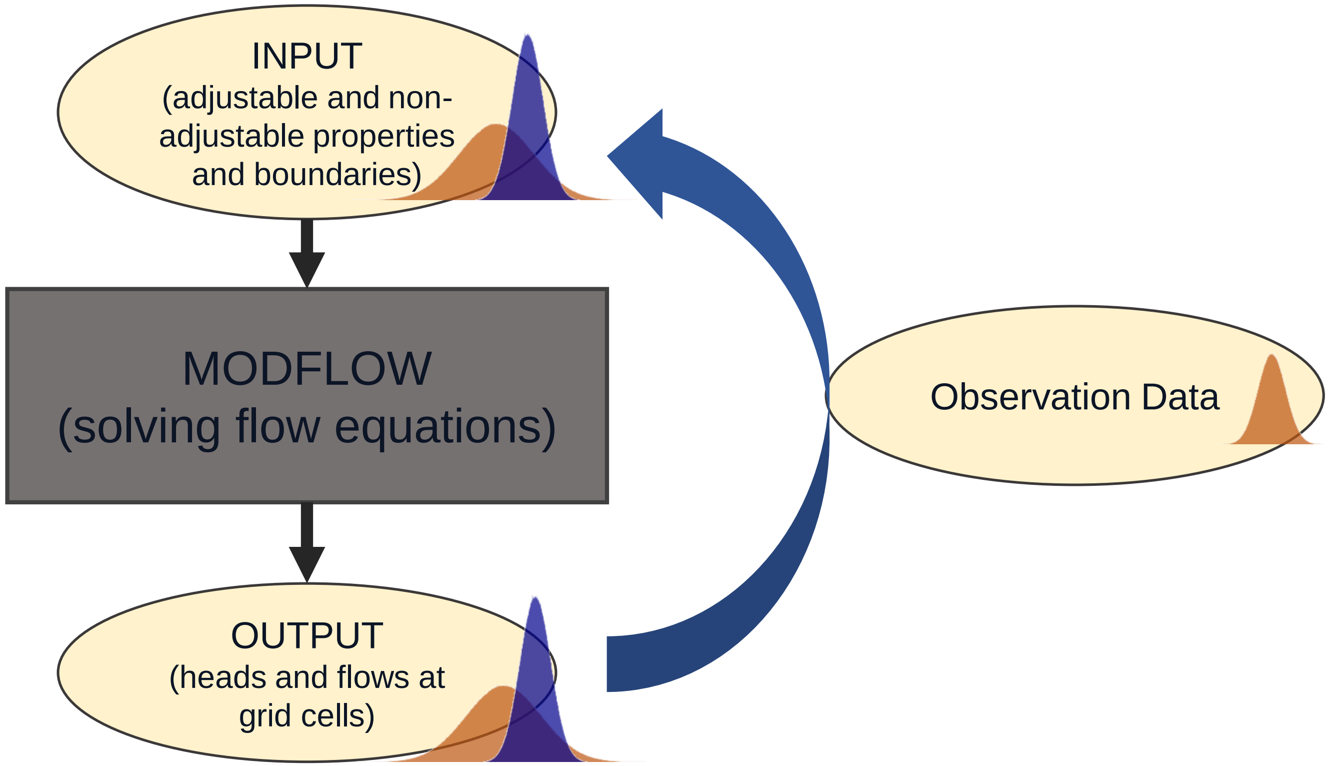
Iterate

```
# settings
# model
sim_workspace = 'base'
model_name = 'mf6-tv_hist_ss'
ss_years = list(range(1998, 2007))
# pest
read_parameters_from_file = False
read_parameters_group_mean = False
balance_group_phi = True
tie_groups = ['leakdistprop', 'leakfac']
# tie_groups = ['drncnd', 'hk', 'infilfac', 'leakdistprop', 'leak
man_obs_group_weight_mult = pd.DataFrame({
    'obgnme' : ['obs_well', 'drn', 'lowell', 'vert_head_diff'],
    # 'weight_mult' : [1, 1e-2, 5e-4, 1]
    'weight_mult' : [1, 3e2, 1e-1, 1e1]
})
```

USGS
science for a changing world

What else?

Uncertainty Analysis



Thanks for listening!

USGS
science for a changing world