SAND2024-01676C

Sandia National Laboratories

Exceptional service in the national interest

Assessing Model Prediction Trustworthiness in the Presence of Model-Form Uncertainty

Teresa Portone, Rebekah White, Joseph Hart

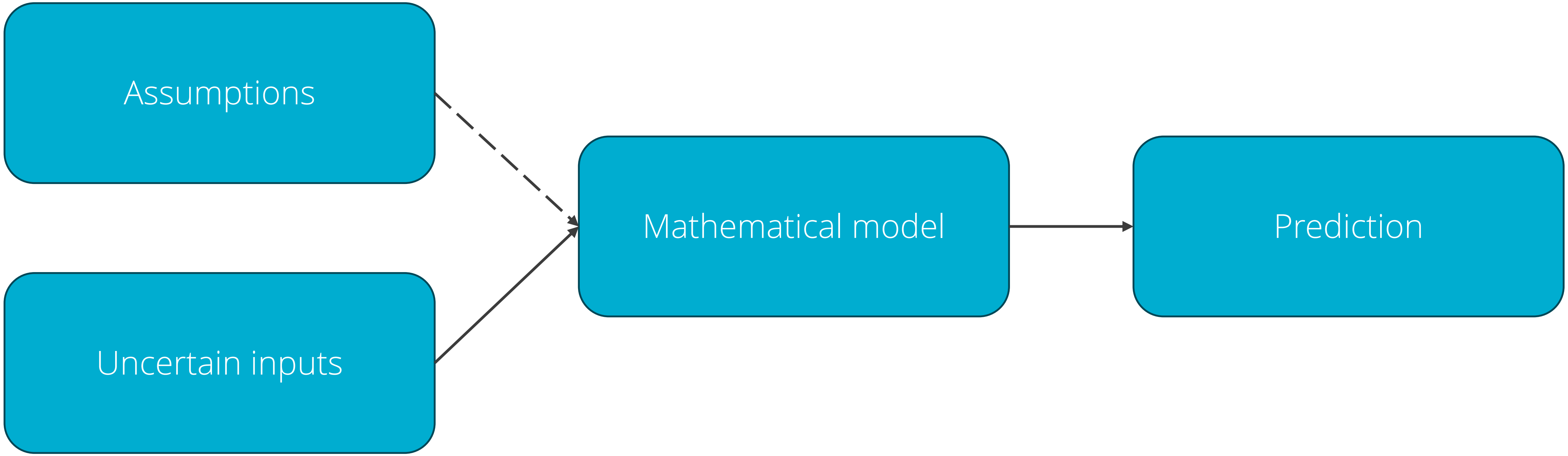
SIAM Conference on Uncertainty Quantification 2024

Trieste, Italy, 2/27-3/1, 2024

SAND2024-01676C

Assumptions affect trustworthiness of model predictions

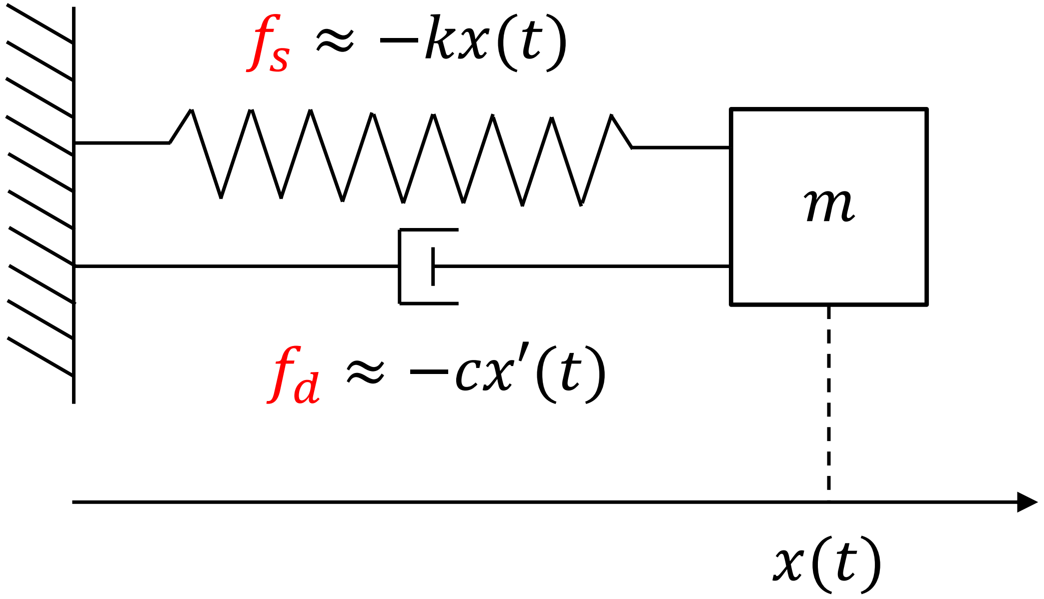
Assumptions in mathematical models not valid in all cases.



How can we quantify their influence on predictions?

Physics-based models = reliable theory + less-reliable embedded assumptions

Reliable theory → $\mathcal{R}(\boldsymbol{s};\boldsymbol{u}) = 0$ ← Unknowns



Newton's 2nd law (energy conservation):

$$f_s + f_d \equiv F = ma \equiv m\frac{d^2x(t)}{dt^2}$$

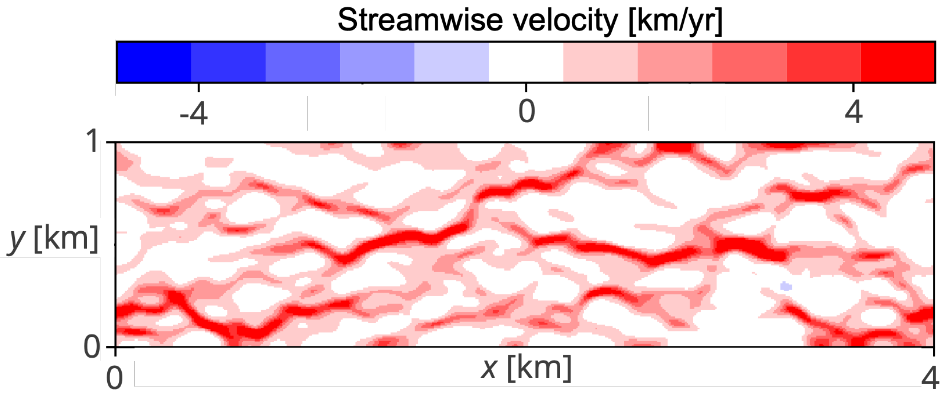
Hypothesis: assumptions reliable ⇒ model prediction reliable

For example: have to assume the form of unknown quantities

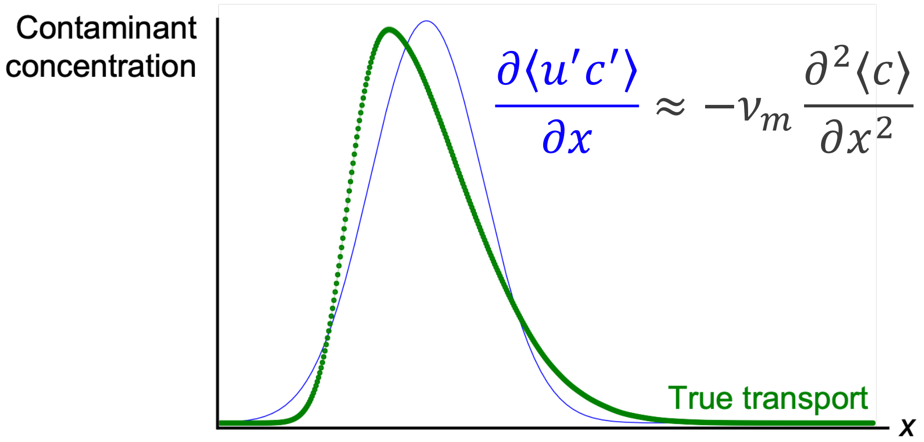
Upscaled subsurface contaminant transport:

$$\frac{\partial\langle c\rangle}{\partial t} + \langle u\rangle \frac{\partial\langle c\rangle}{\partial x} = \nu\Delta\langle c\rangle - \frac{\partial\langle u'c'\rangle}{\partial x}$$

Conservation of mass highly reliable, but assumption that dispersion term depends locally on concentration can be invalid.



High-velocity "streaks" cause nonlocality in contaminant transport.



Ingredients to assess prediction trustworthiness based on assumptions

Assumption important to prediction?

+

Assumption stress-tested in validation?

We measure assumption importance by combining model-form uncertainty representations with grouped sensitivity analysis

Model-form uncertainty representation: parameterized modification to model assumption

Grouped sensitivity analysis: Measures importance of group of parameters to model output

↓

Importance measure for assumption

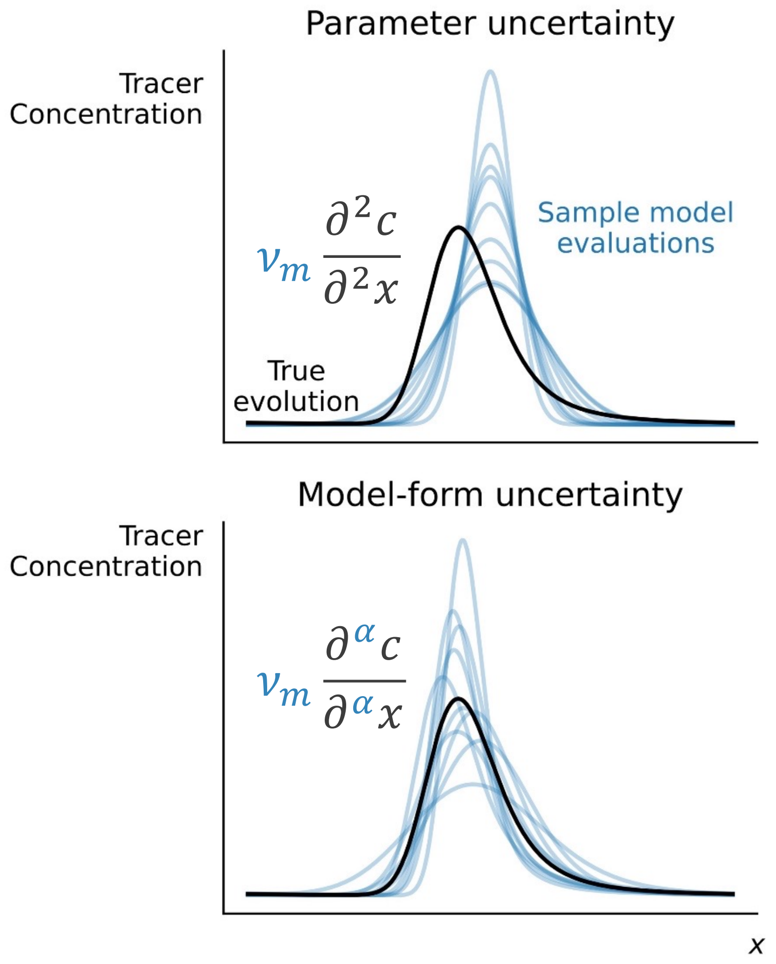
MFU representations reflect a range of plausible assumption forms

$$\frac{\partial \langle c \rangle}{\partial t} + \langle u \rangle \frac{\partial \langle c \rangle}{\partial x} = \nu_p \Delta \langle c \rangle - \frac{\partial \langle u'c' \rangle}{\partial x}$$

$$c_0(x) = \exp(-(x-s)^2/l^2)$$

MFU representation:

$$\frac{\partial \langle u'c' \rangle}{\partial x} \approx -\nu_m \frac{\partial^\alpha \langle c \rangle}{\partial^\alpha x}$$



MFU representations should be parameterized to

- respect relevant physics
- reflect range of plausible behavior
- maintain randomness after inference
  - MFU representation's form doesn't perfectly capture true behavior

Bernstein von-Mises theorem[1]: $$p(\theta|\boldsymbol{d}) \xrightarrow{N_{obs}\to\infty} \delta(\theta - \theta_{MLE})$$

Data should not inform MFU parameters directly.

[1]B.J.K. Kleijn and A.W. van der Vaart. "The Bernstein-Von-Mises Theorem under Misspecification." *Electronic Journal of Statistics*, vol. 6, no. none, Jan. 2012, pp. 354–81, https://doi.org/10.1214/12-EJS675.

Grouped Sobol' indices measure model output sensitivity to a group of parameters and their interactions

$$\boldsymbol{X} = [\boldsymbol{u}, \boldsymbol{u_c}]$$

Main effect index

$$S_{\boldsymbol{u}} = \frac{\mathbb{V}_{\boldsymbol{u}}\left(\mathbb{E}_{\boldsymbol{u_c}}[\,f(\boldsymbol{X})|\boldsymbol{u}]\right)}{\mathbb{V}\left(f(\boldsymbol{X})\right)}$$

Total effect index

$$T_{\boldsymbol{u}} = \frac{\mathbb{E}_{\boldsymbol{u_c}}(\mathbb{V}_{\boldsymbol{u}}[\,f(\boldsymbol{X})|\boldsymbol{u_c}])}{\mathbb{V}\left(f(\boldsymbol{X})\right)}$$

If $\boldsymbol{u} = [x_1, x_2]$,

$$S_{\boldsymbol{u}} = S_1 + S_2 + S_{12}$$

Assumption importance measured by grouped Sobol' index for MFU parameters

$$\frac{\partial \langle u'c' \rangle}{\partial x} \approx -\nu_m \frac{\partial^{\alpha} \langle c \rangle}{\partial^{\alpha} x}$$

$$S_{\{\nu_m, \alpha\}}$$

Numerical example

Goal: measure importance of dispersion assumption relative to other sources of uncertainty in the system

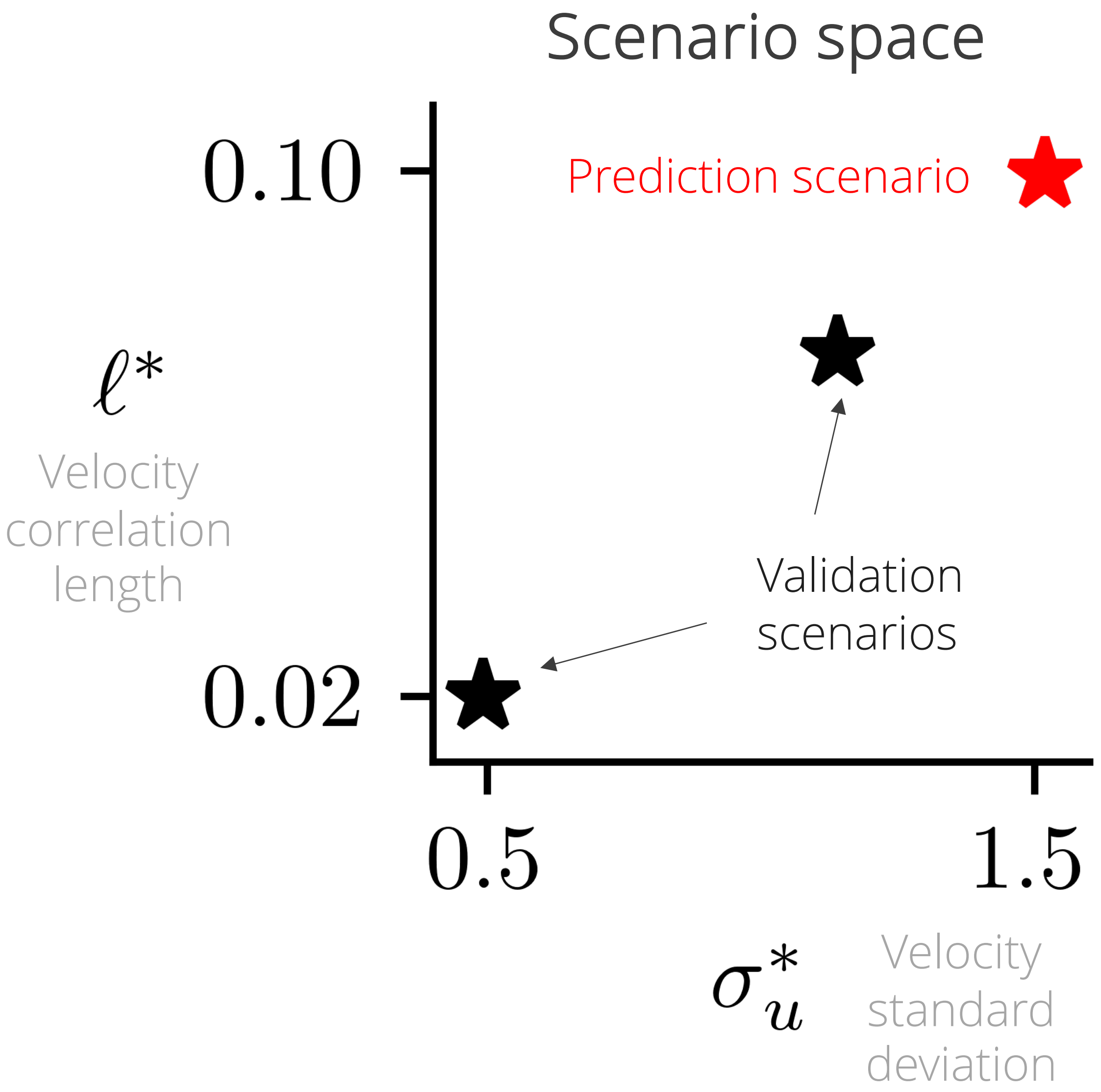
$$\frac{\partial \langle c \rangle}{\partial t} + \langle u \rangle \frac{\partial \langle c \rangle}{\partial x} = \nu_p \Delta \langle c \rangle - \frac{\partial \langle u'c' \rangle}{\partial x}$$

$$c_0(x) = \exp\left(-(x - s)^2 / l^2\right)$$

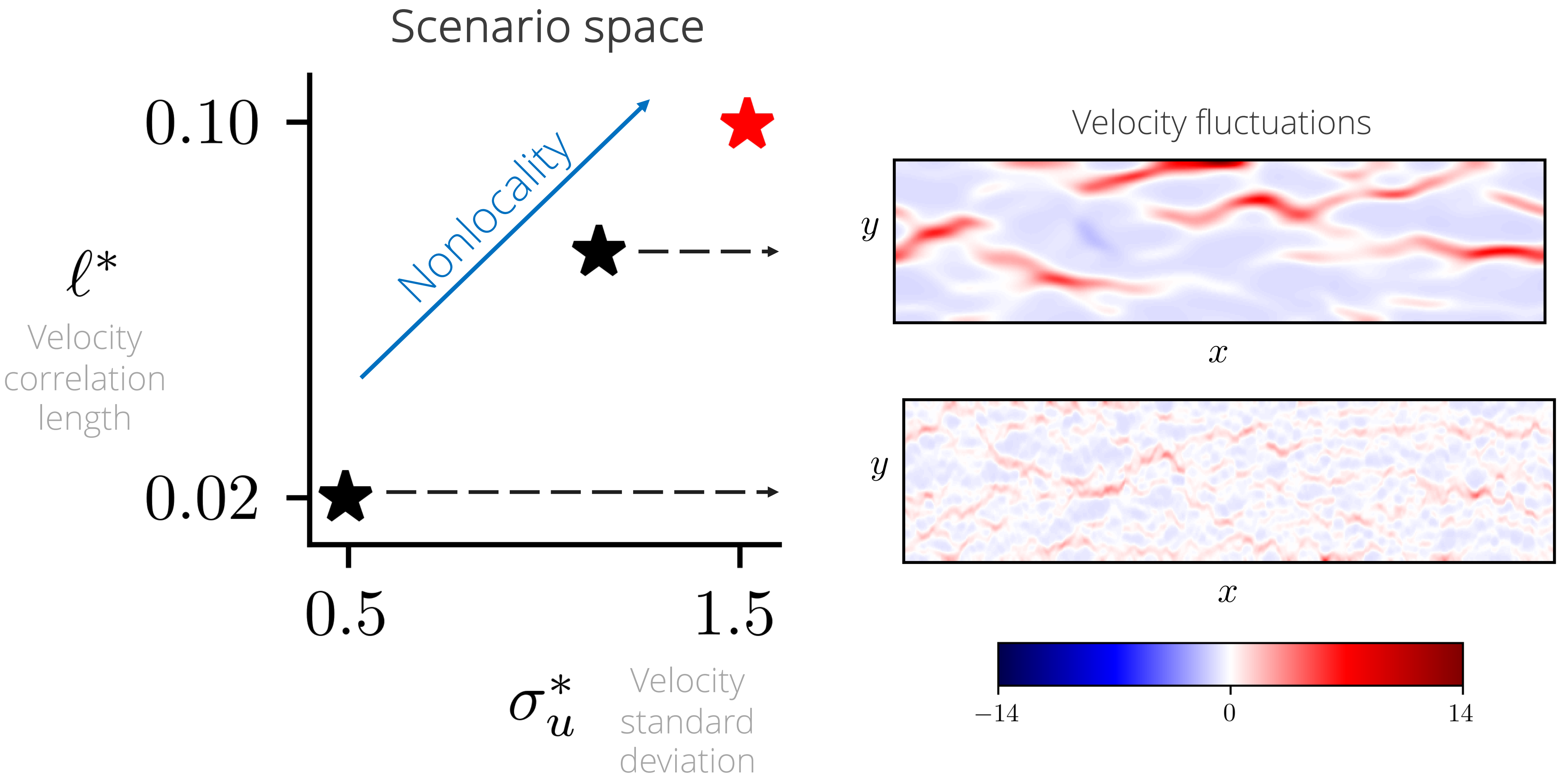
MFU representation: $$\frac{\partial \langle u'c' \rangle}{\partial x} \approx -\nu_m \frac{\partial^\alpha \langle c \rangle}{\partial^\alpha x}$$

Sensitivity measures for: $s, \langle u \rangle, \nu_p, \{\nu_m, \alpha\}$

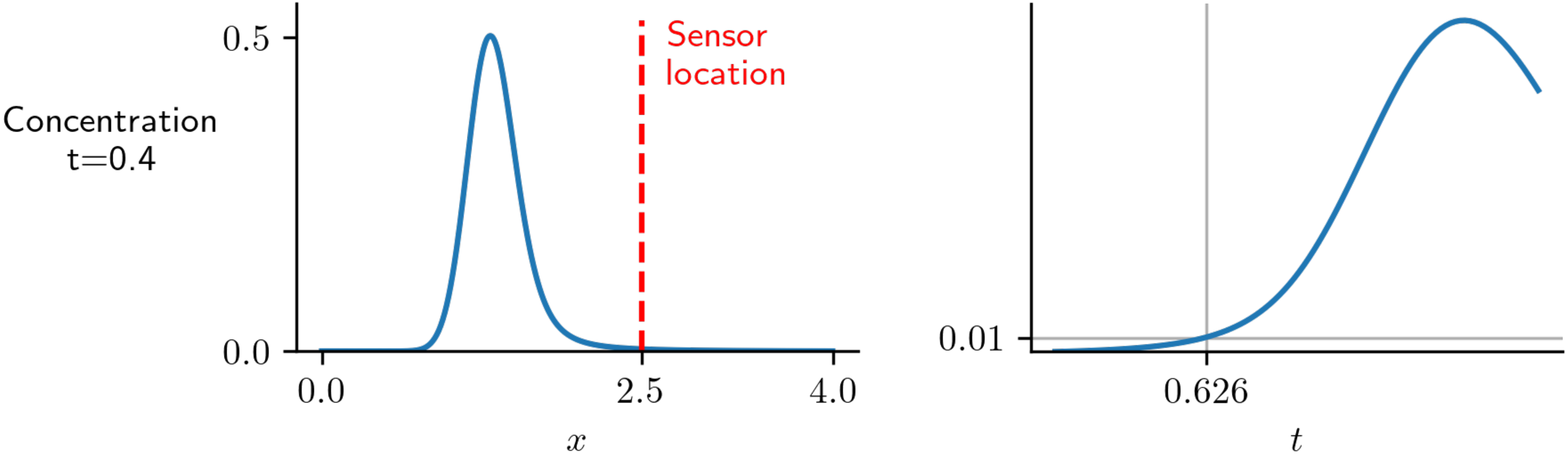
How important is assumption to prediction vs. validation scenarios?



How important is assumption to prediction vs. validation scenarios?



Quantity of interest (QoI): breakthrough time at x=2.5



Find first time concentration exceeds 10^{-2} *at* $x = 2.5$.

MFU representation parameterized to…

$$\frac{\partial \langle u'c' \rangle}{\partial x} \approx -\nu_m \frac{\partial^\alpha \langle c \rangle}{\partial^\alpha x}$$

respect relevant physics

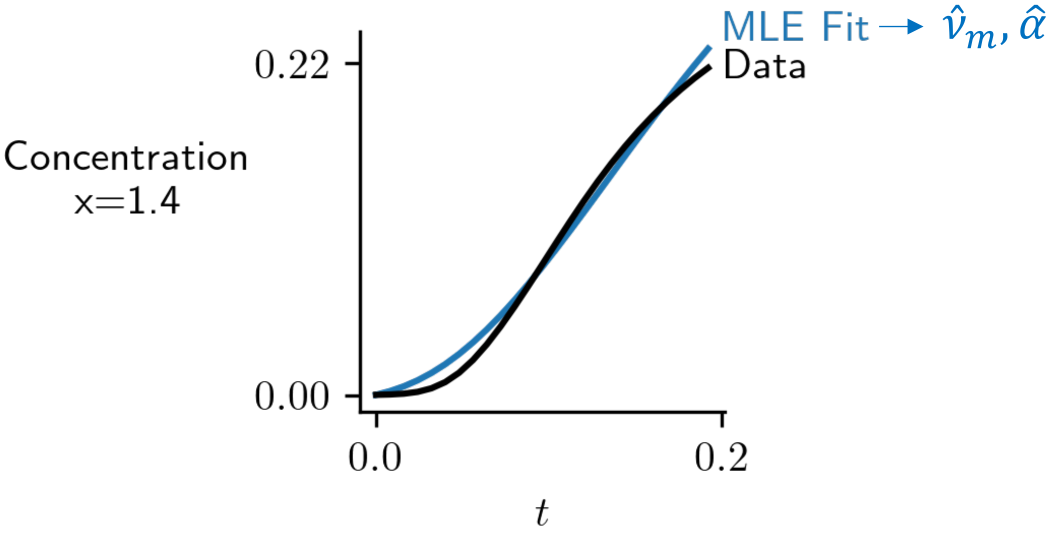
- *conserve mass*
- *maintain positive concentration*

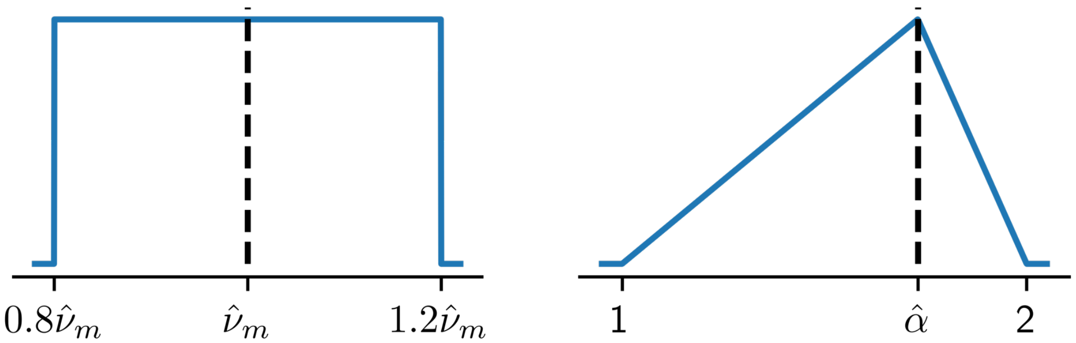
reflect range of plausible behavior

- *vary influence and degree of nonlocality through* ν_m, α

maintain randomness after inference

- *data informs hyperparameters of parameter distributions vs. informing parameter values directly*





Data from ensemble of detailed contaminant transport simulations

We measured sensitivity of MFU parameters relative to other uncertainties in the model

$$\frac{\partial\langle c\rangle}{\partial t} + \langle u\rangle \frac{\partial\langle c\rangle}{\partial x} = \nu_p \Delta\langle c\rangle - \frac{\partial\langle u'c'\rangle}{\partial x}$$

$$c_0(x) = \exp(-(x-s)^2/l^2)$$

Parameter type	Parameter	Distribution
Other model parameters	IC mode (s)	$\mathcal{U}[0.8s_n\ 1.2s_n]$, $s_n = 1$
	$\langle u\rangle$	$\mathcal{U}[0.8u_n, 1.2u_n]$, $u_n = 1$
	ν_p	$\mathcal{U}[0.8\nu_{p,n}, 1.2\nu_{p,n}]$, $\nu_{p,n} = 0.01$
Fractional derivative MFU	ν_m	$\mathcal{U}[0.8\hat{\nu}_m, 1.2\hat{\nu}_m]$
	α	Triangular([1,2], mode=$\hat{\alpha}$)

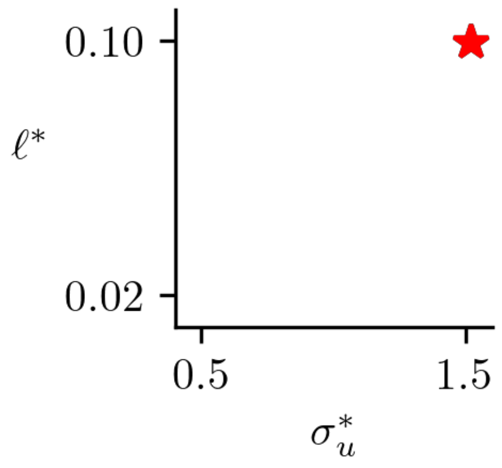
Ingredients to assess prediction trustworthiness based on assumptions

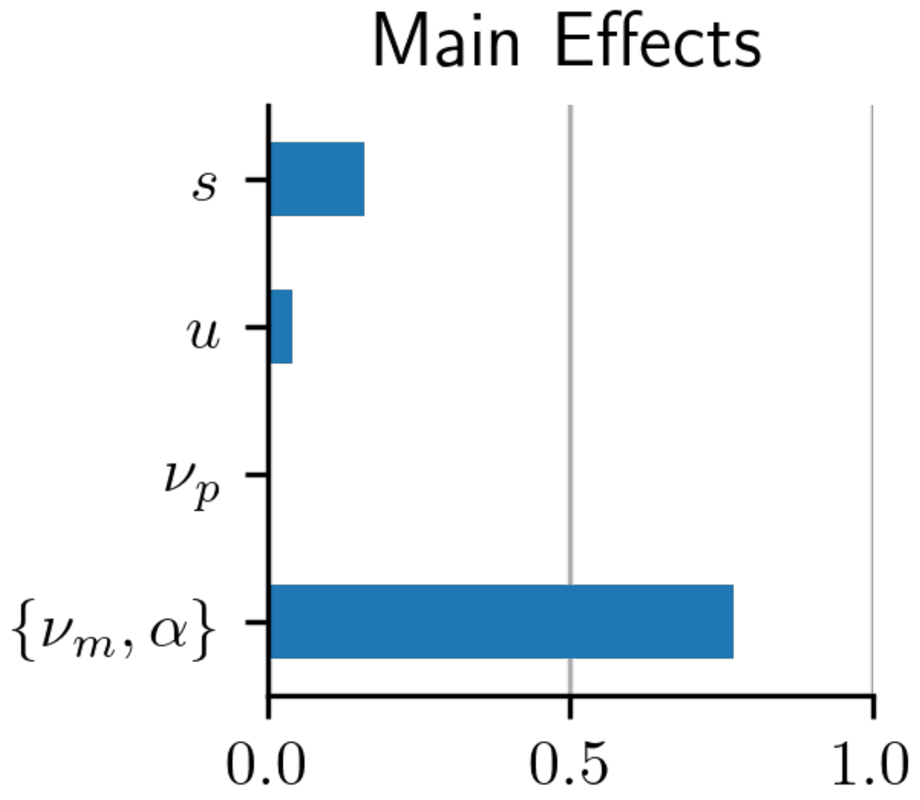
Assumption important to prediction?

+

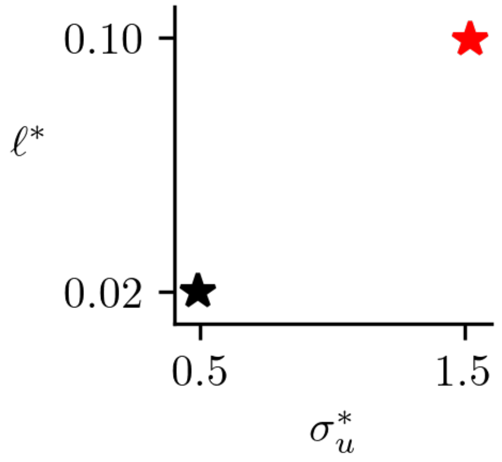
Assumption stress-tested in validation?

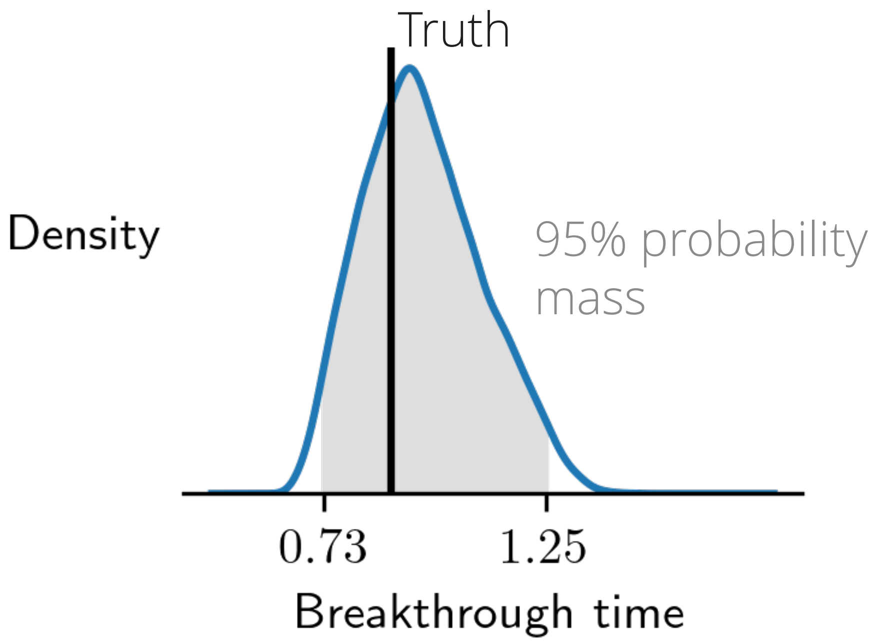
Dispersion assumption very important to prediction

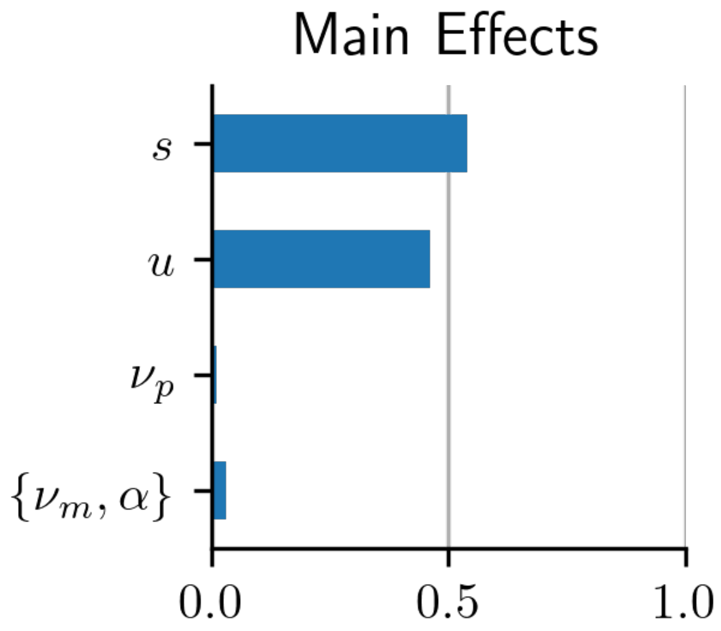




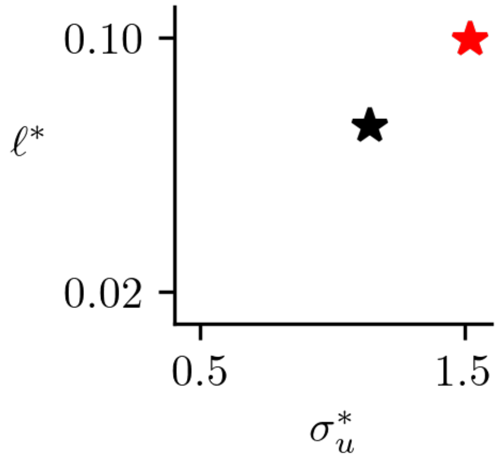
Model valid but assumption not important → doesn't confer confidence

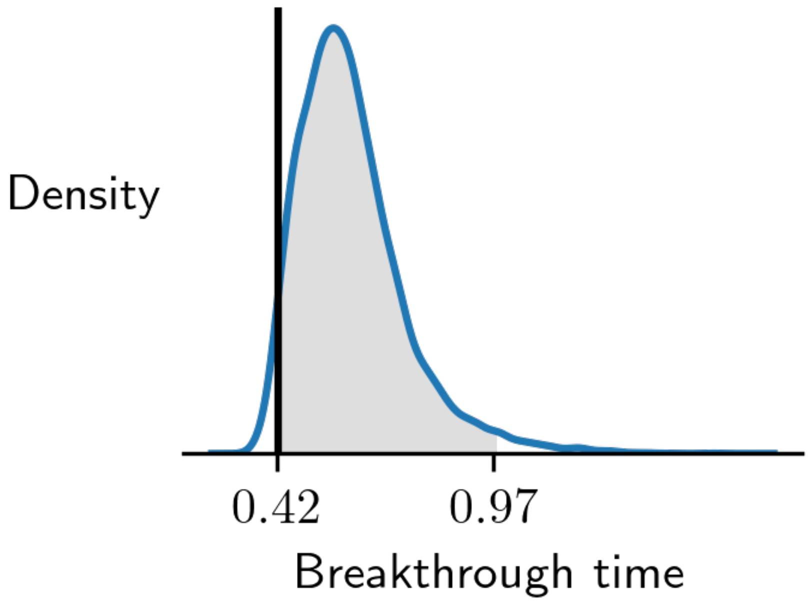


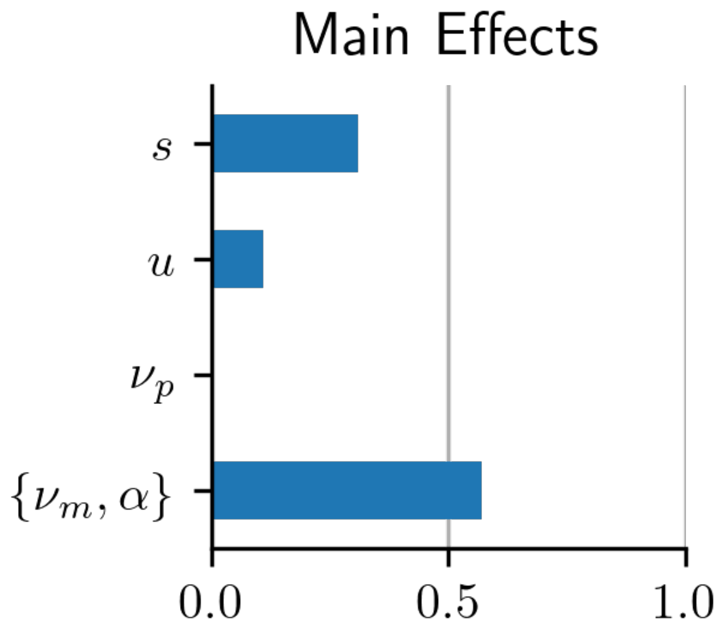




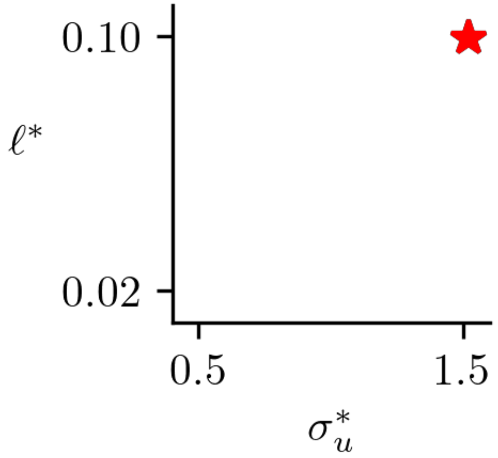
Model valid + assumption important → confers confidence for prediction

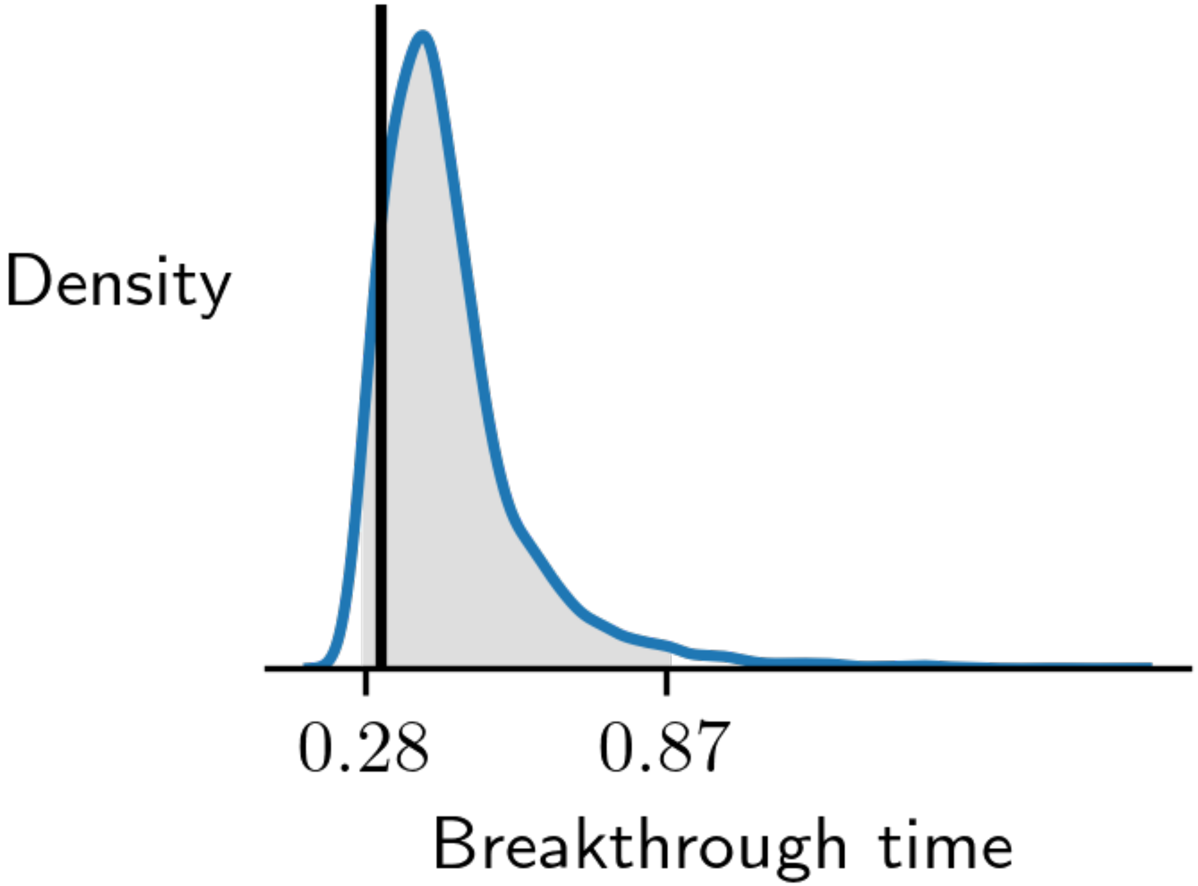






Prediction valid

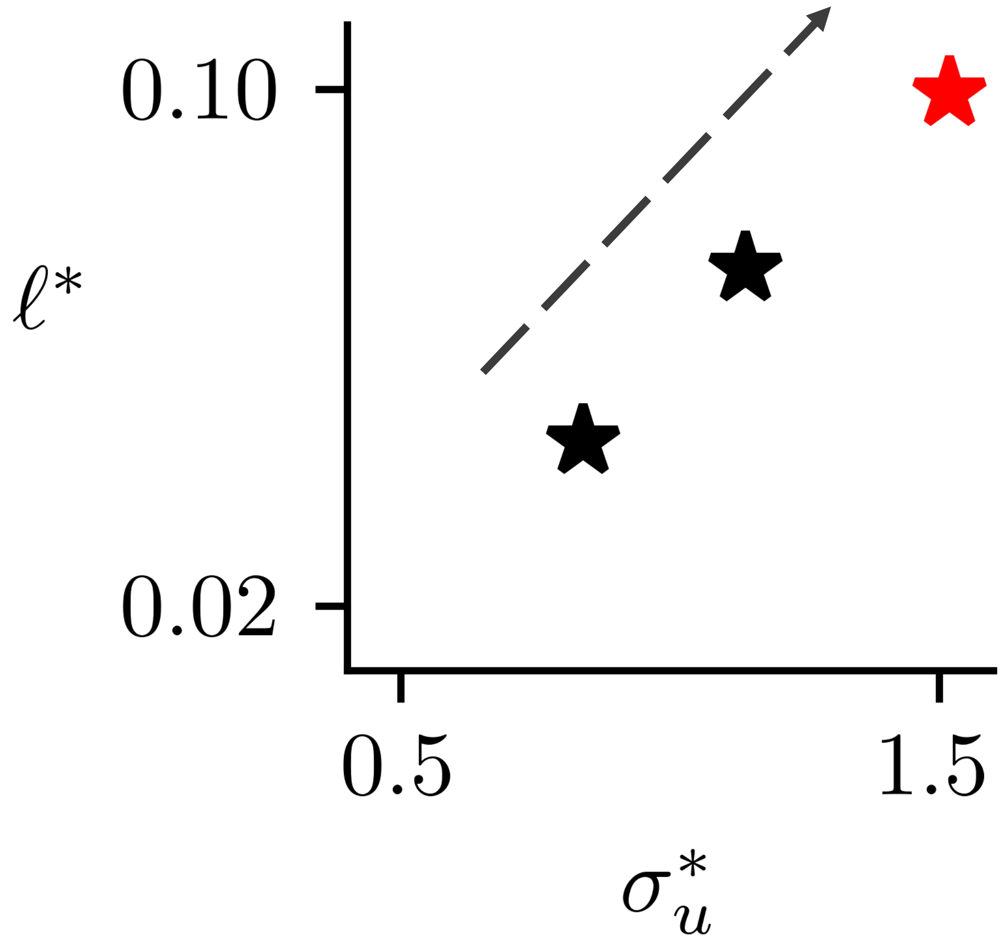




Validation scenarios: how close is “close enough”?

How much does assumption’s parameterization change with scenario?

(Especially in direction of prediction)



Conclusions and future work

- Stress testing model assumptions is critical to assess confidence in predictions
- Combined MFU representations & grouped GSA to quantify importance of assumptions to model outputs
- Lets us identify which assumptions are important to prediction
- Lets us check if those assumptions were important to validation

Future work: develop a method to measure if a validation test scenario is “close enough” to the prediction scenario to confer confidence

Thanks!

Portone, Teresa, et al. "Quantifying Model Prediction Sensitivity to Model-Form Uncertainty." *In Preparation*.

tporton@sandia.gov

Bernstein-von Mises Theorem

$$p(\theta|\boldsymbol{d}) \xrightarrow{N_{obs}\to\infty} \mathcal{N}\left(\hat{\theta}, \left(N_{obs} I(\hat{\theta})\right)^{-1}\right)$$

As number of observations approaches infinity, the variance of the normal approaches zero, approximating a Dirac delta

References

Grouped Sobol' Indices:

Prieur, Clémentine, and Stefano Tarantola. "Variance-Based Sensitivity Analysis: Theory and Estimation Algorithms." *Handbook of Uncertainty Quantification*, edited by Roger Ghanem et al., Springer International Publishing, 2017, pp. 1217–39, https://doi.org/10.1007/978-3-319-12385-1_35.

Model-form uncertainty/inadequacy/model-form error representations:

Morrison, Rebecca E., et al. "Representing Model Inadequacy: A Stochastic Operator Approach." *SIAM/ASA Journal on Uncertainty Quantification*, vol. 6, no. 2, Jan. 2018, pp. 457–96, https://doi.org/10.1137/16M1106419.

Oliver, Todd A., et al. "Validating Predictions of Unobserved Quantities." *Computer Methods in Applied Mechanics and Engineering*, vol. 283, Jan. 2015, pp. 1310–35, https://doi.org/10.1016/j.cma.2014.08.023.

Portone, Teresa. *Representing Model-Form Uncertainty from Missing Microstructural Information*. 2019. University of Texas at Austin, http://dx.doi.org/10.26153/tsw/10112.