Automatic Calibration and Uncertainty Analysis of Groundwater Models

David Dempsey

Department of Engineering Science, University of Auckland

THE UNIVERSITY OF AUCKLAND
Te Whare Wananga o Tamaki Makaurau
NEW ZEALAND

ENGINEERING

Calibration

Making models look more like reality.

Two options:

- Fiddle with the parameters to match the data.
- Have a computer do it for you.

Objective function:

- A quantitative measure of misfit between model and data.
- E.g., sum of squares

$$S = \sum_i \left(f(x_i; \theta) - \tilde{y}_i \right)^2$$

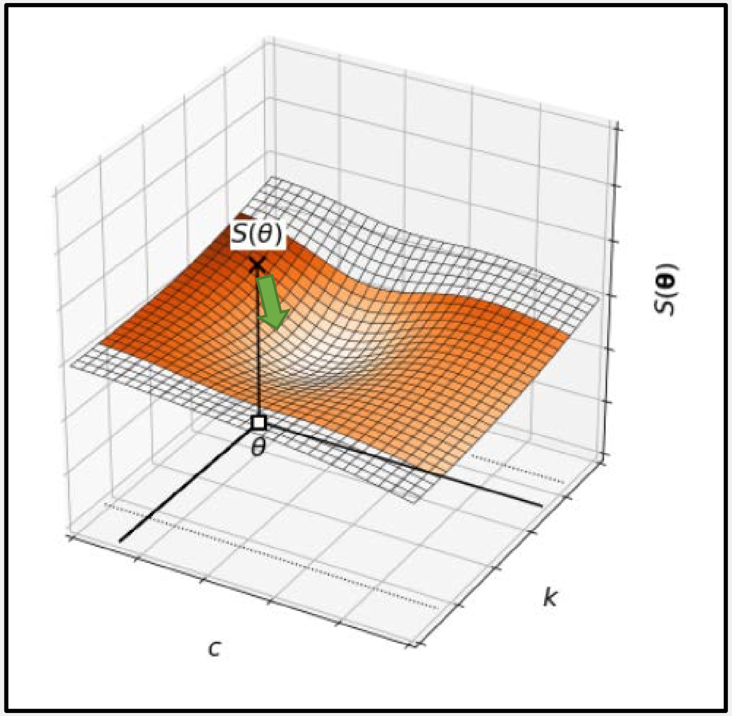
model (e.g., EMOD3D)

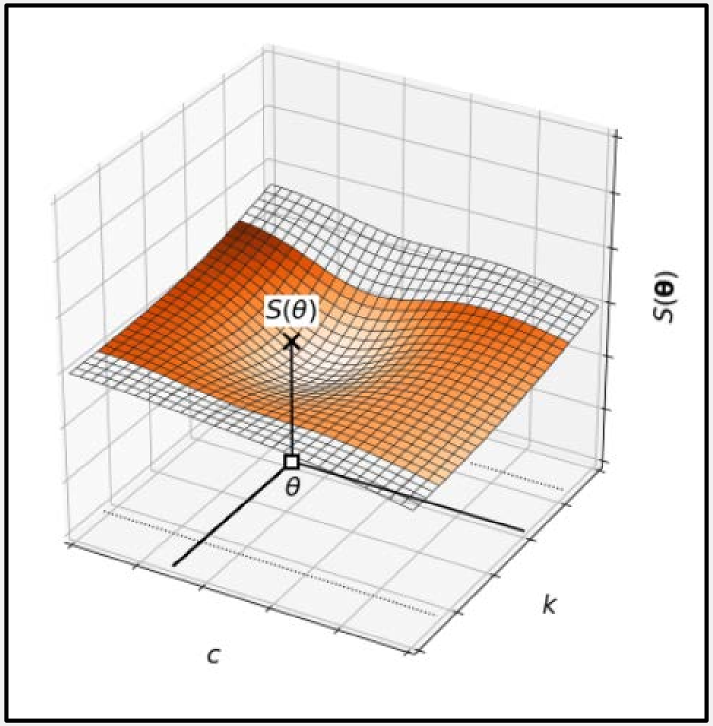
parameters (e.g., velocity)

data (e.g., arrival times)

Gradient Calibration

Algorithm automatically updates θ in a direction that causes misfit to get smaller (model to get “better”).





Automatic Calibration of Geothermal Models

Hundreds of parameters → "direction" in parameter space is "hundreds"-dimensional vector.

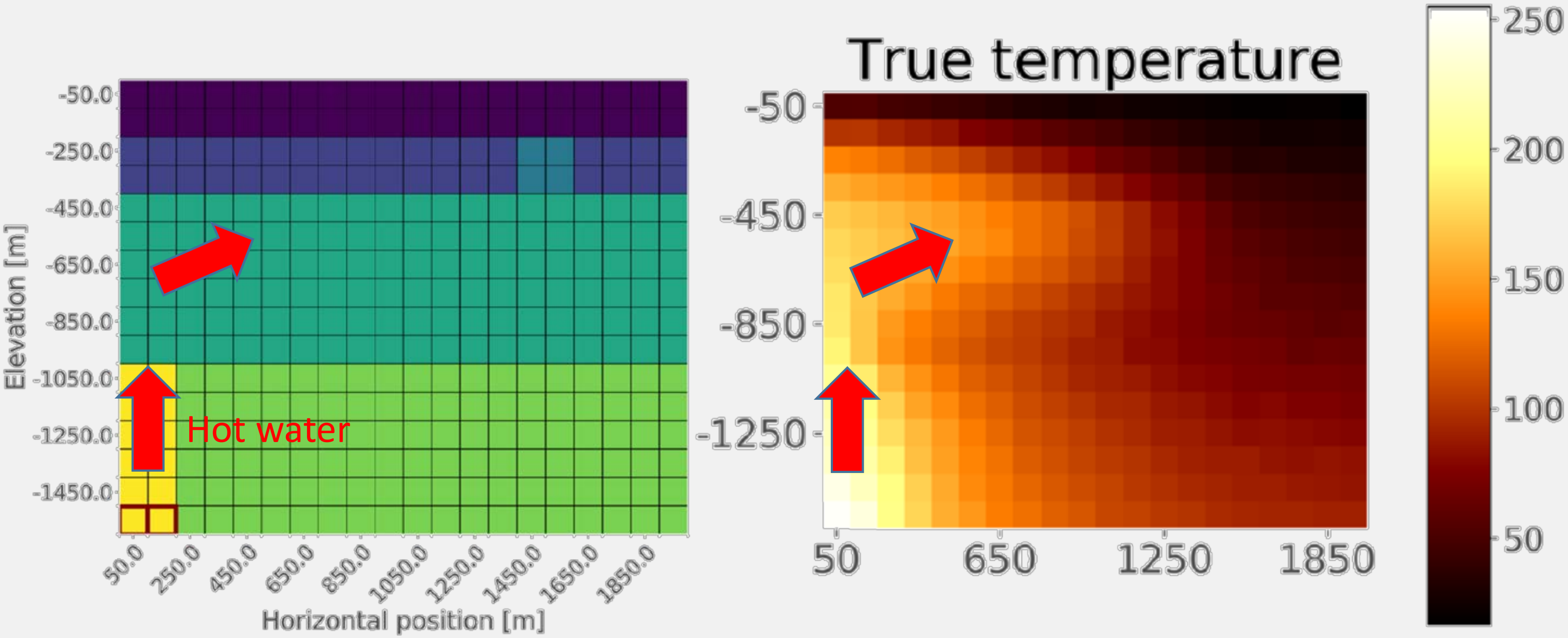
Each component requires a full run of the model to evaluate (numerical Jacobian).

E.g., 10 parameters ⇒ 11 ground motion simulations, just improve the model one time. Many iterations may be required.

Expensive. Speed this up using analytic Jacobian, or the adjoint.

Non-uniqueness. If we have more parameters than data, lots of different models can fit the data. Need regularisation.

Adjoint example - geothermal



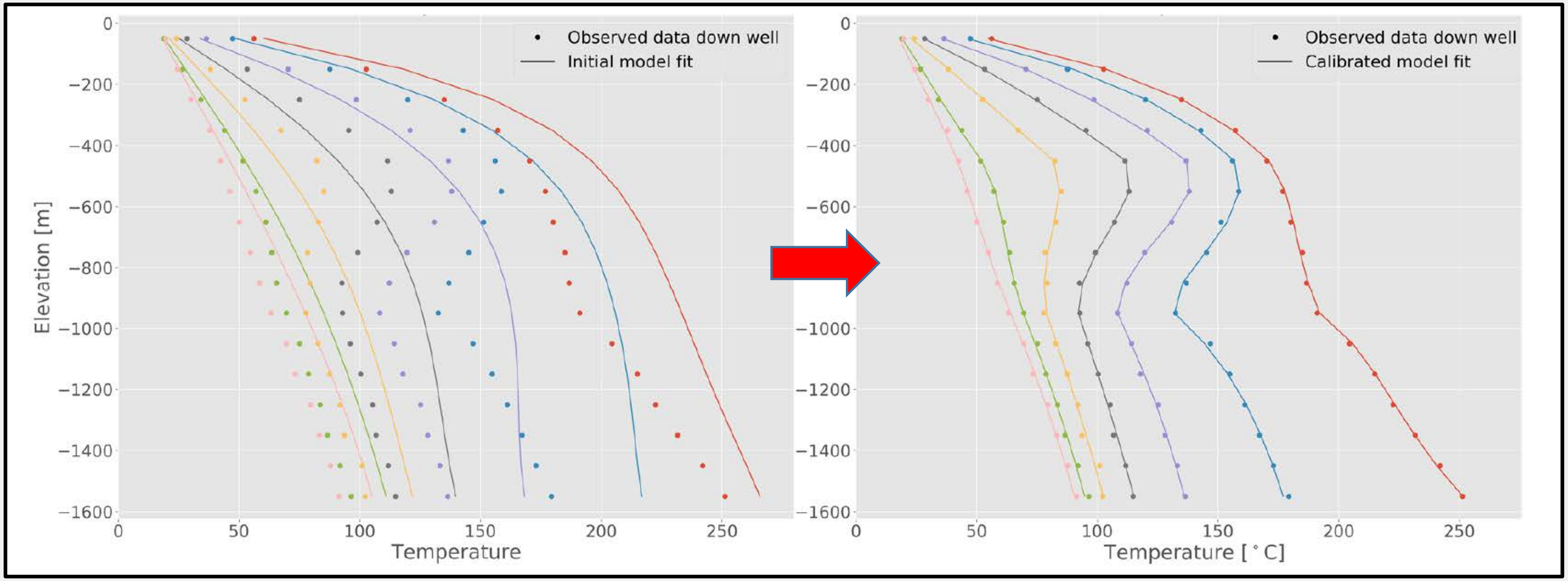
"True" model – invert permeabilities by matching temperatures

Gonzalez-Guiterrez et al. (2018)

Adjoint example - geothermal

Initial model

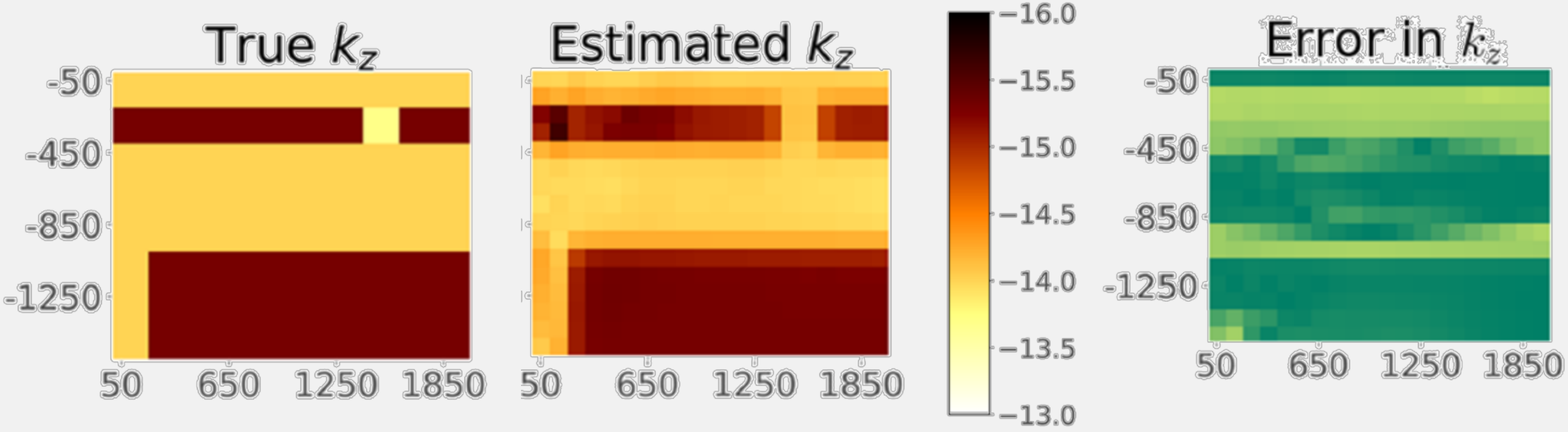
Calibrated model



"True" model – invert permeabilities by matching temperatures

Gonzalez-Guiterrez et al. (2018)

Adjoint example - geothermal



"True" model – invert permeabilities by matching temperatures

Gonzalez-Guiterrez et al. (2018)

Uncertainty

We use models to make predictions. These predictions are always wrong.

BUT it is possible to make two predictions, and have reality fall between.

Two predictions, requires two models, requires two parameter sets, θ_1 and θ_2.

What can we say about θ_1 and θ_2?

- Both should give models that fit the data "well enough".
- Neither are "the best fitting" model.
- Models between θ_1 and θ_2 may be okay as well.

Uncertainty

How to find the okay fitting parameters, θ_1 and θ_2?

Need a likelihood, LK, e.g.,

$$LK = e^{-S/2}, \qquad S = \frac{1}{\sigma^2}\sum_i (f(x_i;\theta) - \tilde{y}_i)^2$$

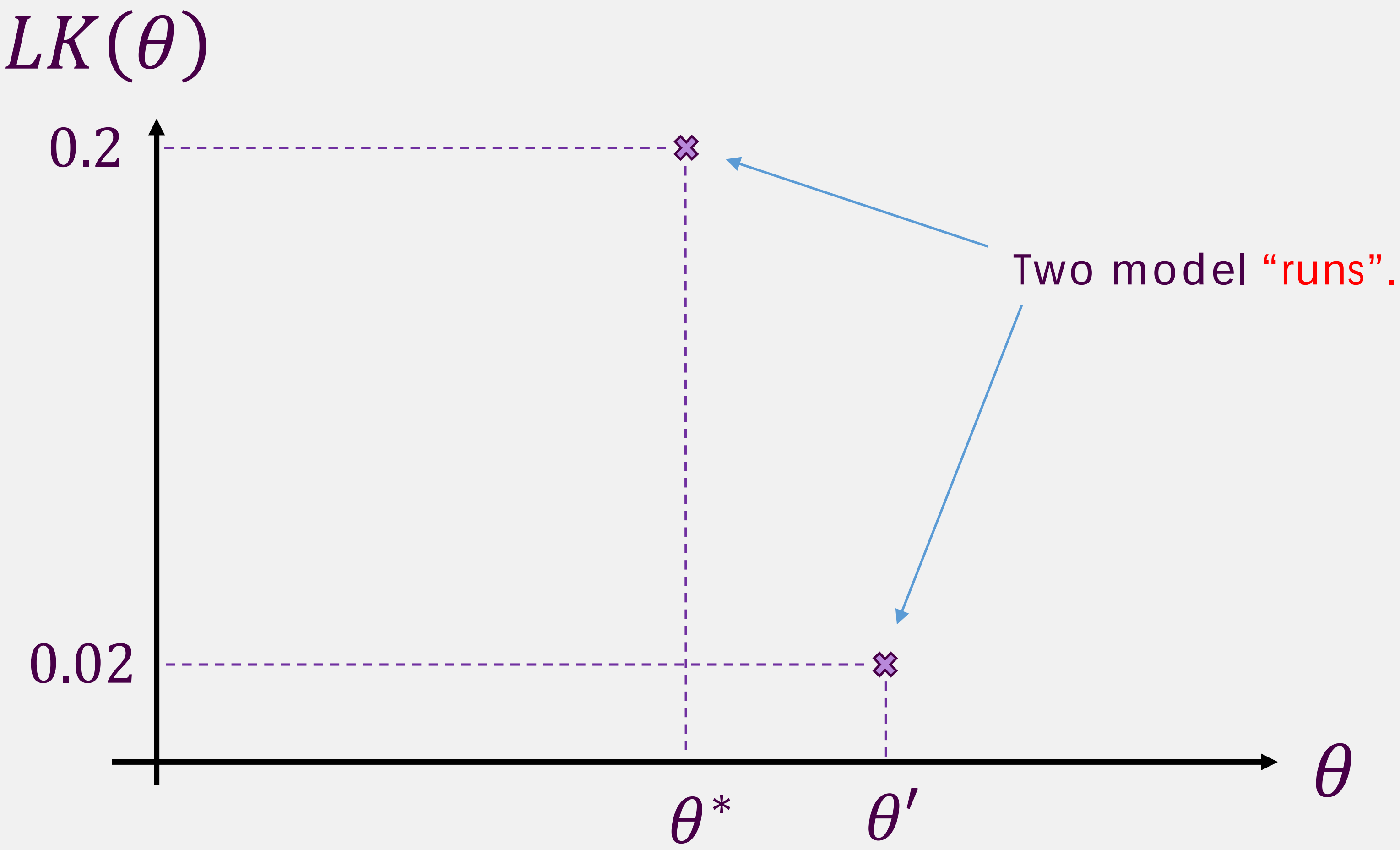
(the above only true if $\tilde{y}_i$ has normally distributed errors, σ^2)

The likelihood is "what is the relative probability this model is correct".

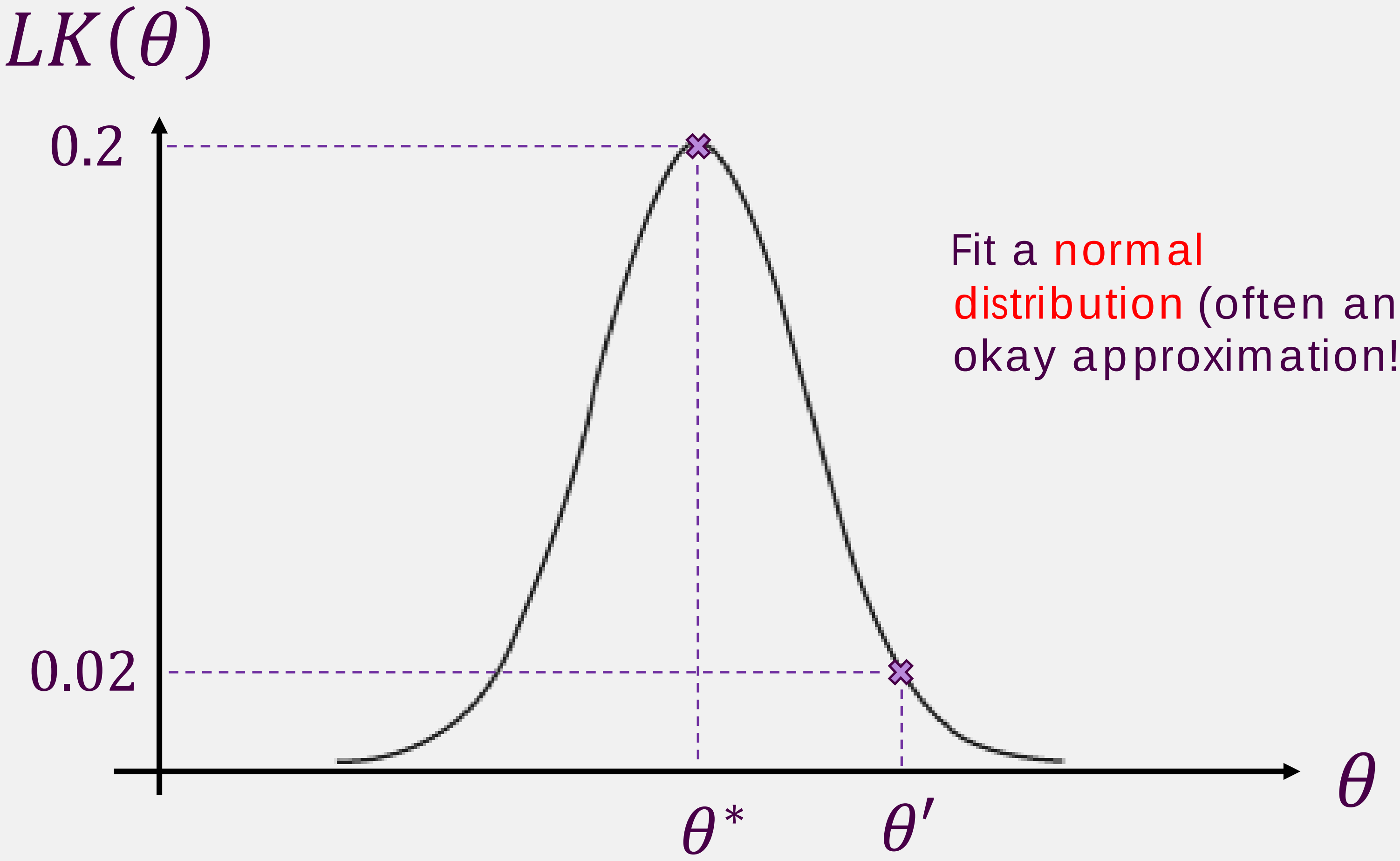
Start with "best-fitting" model, θ^*, which has likelihood, $LK(\theta^*) = 0.2$ – this is not a useful number on it's own...

But what about the different model θ'? It has $LK(\theta') = 0.02$. This model is 10 times less likely to be "correct" than our best model, θ^*.

Linear Sensitivity analysis

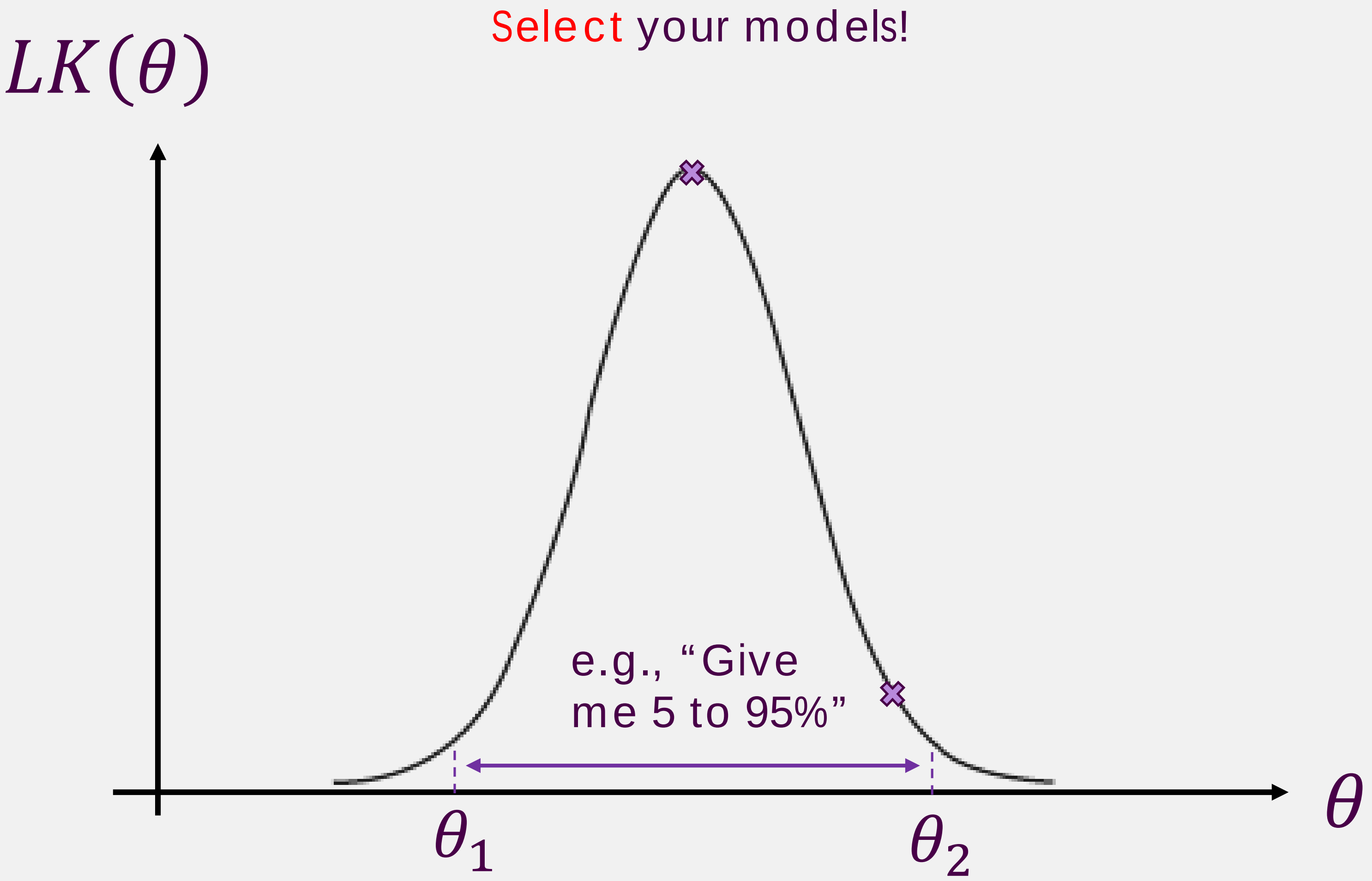


Linear Sensitivity analysis



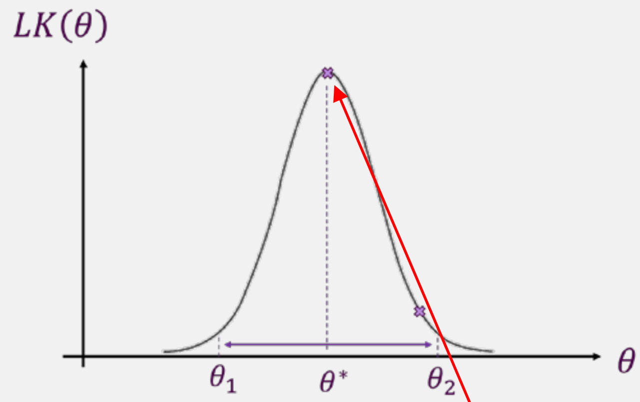
Fit a normal distribution (often an okay approximation!).

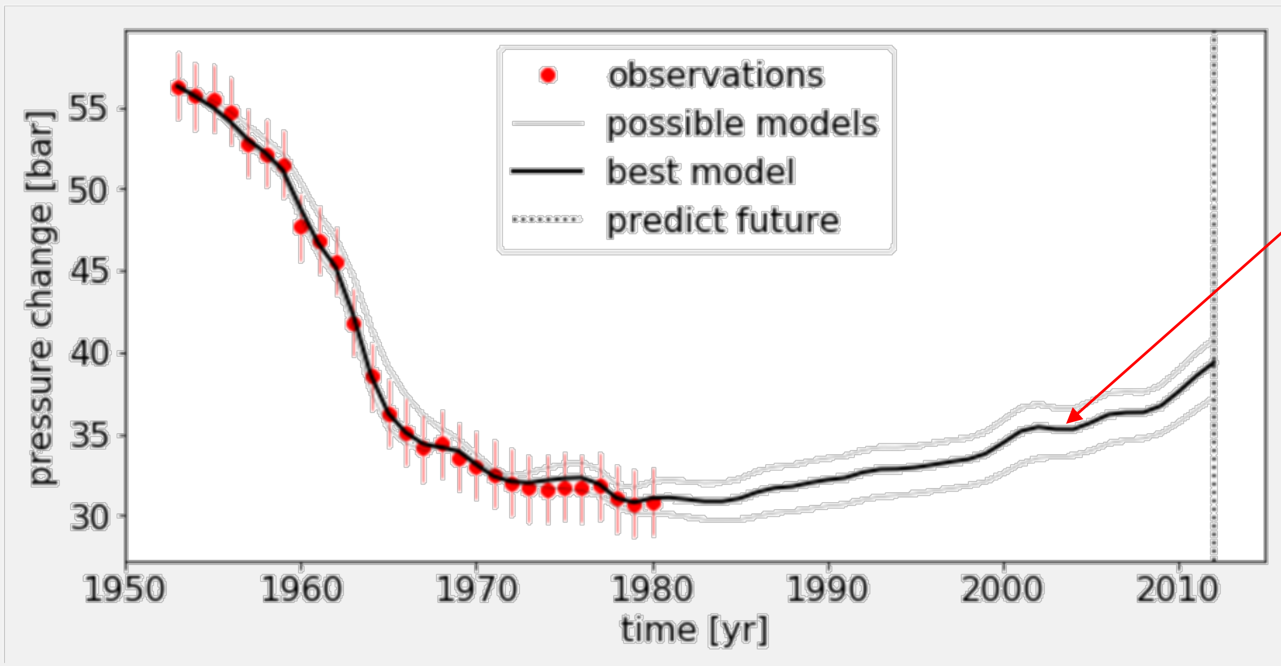
Linear Sensitivity analysis



Linear Sensitivity analysis

E.g., a simple geothermal model.

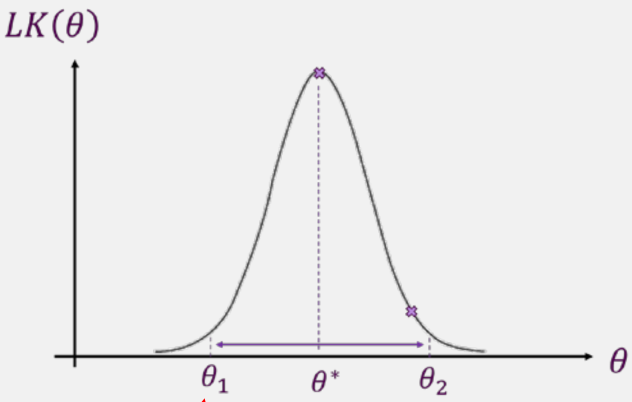




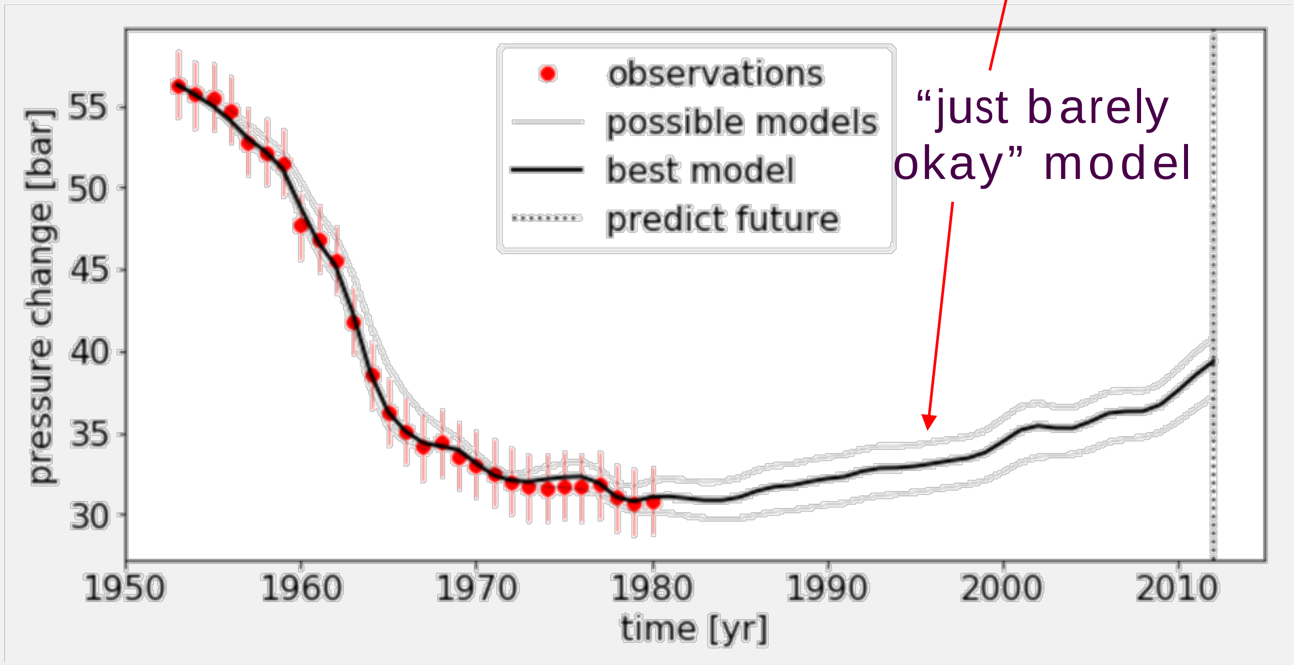
"best fit" model

Linear Sensitivity analysis

E.g., a simple geothermal model.

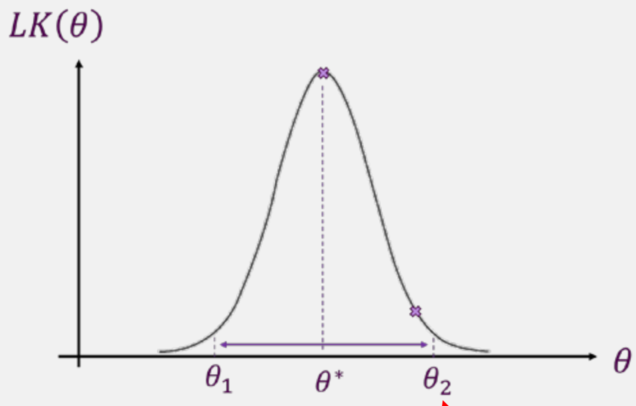


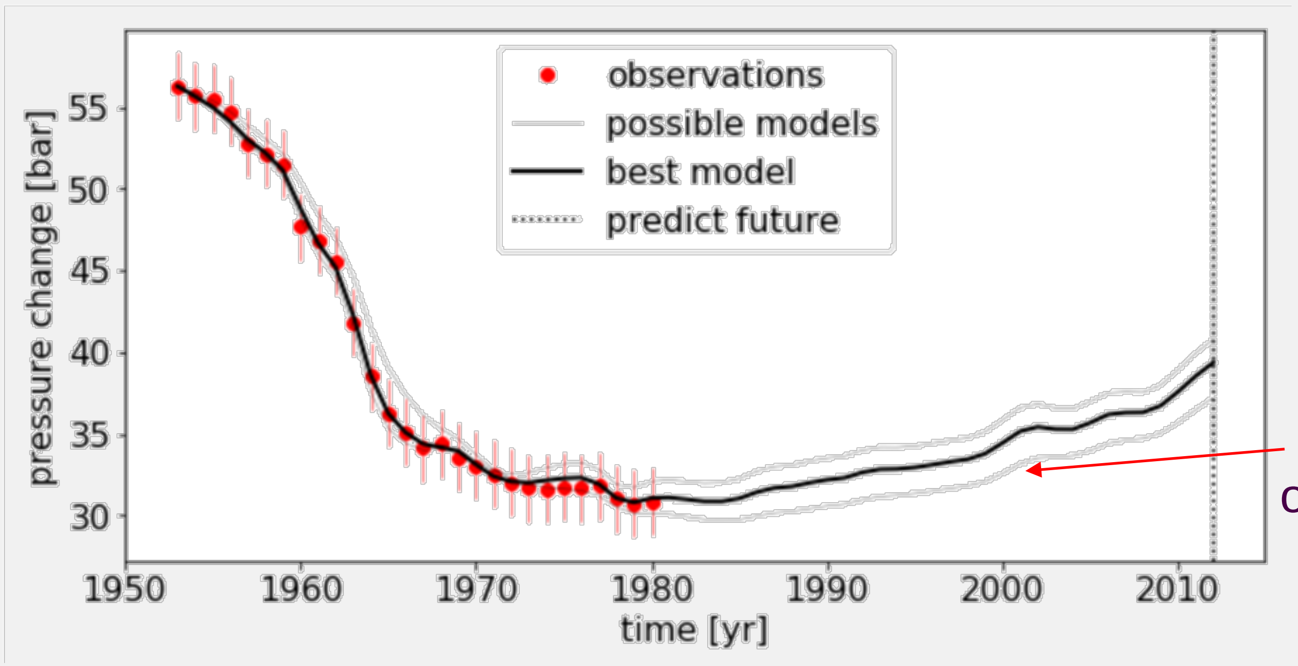
"just barely okay" model



Linear Sensitivity analysis

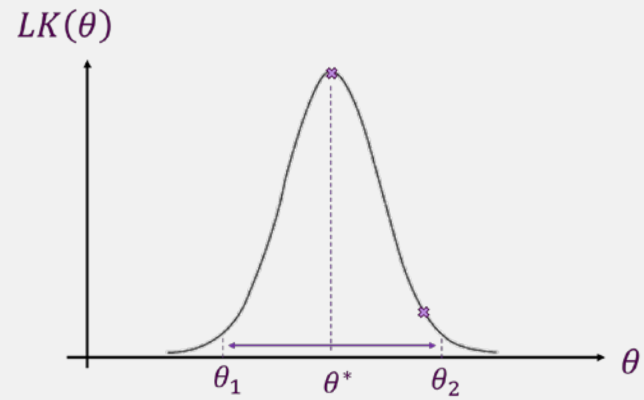
E.g., a simple geothermal model.



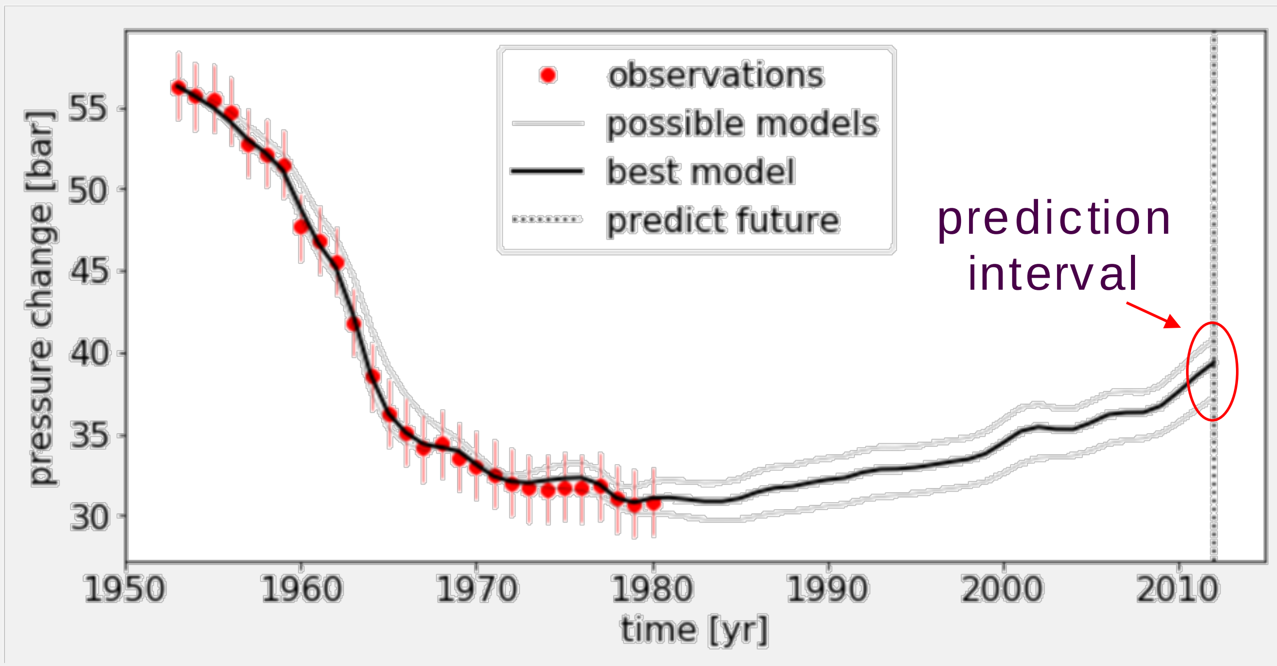


other “just barely okay” model

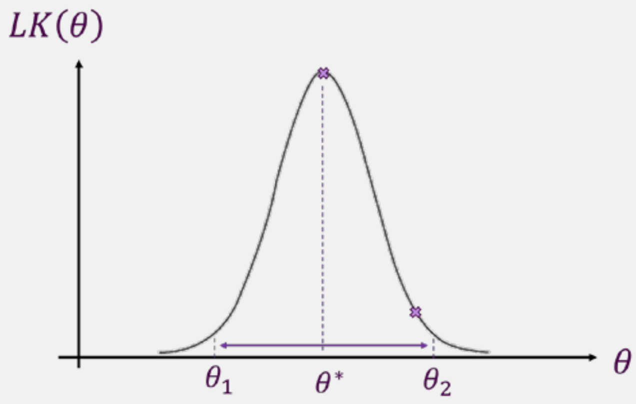
Linear Sensitivity analysis



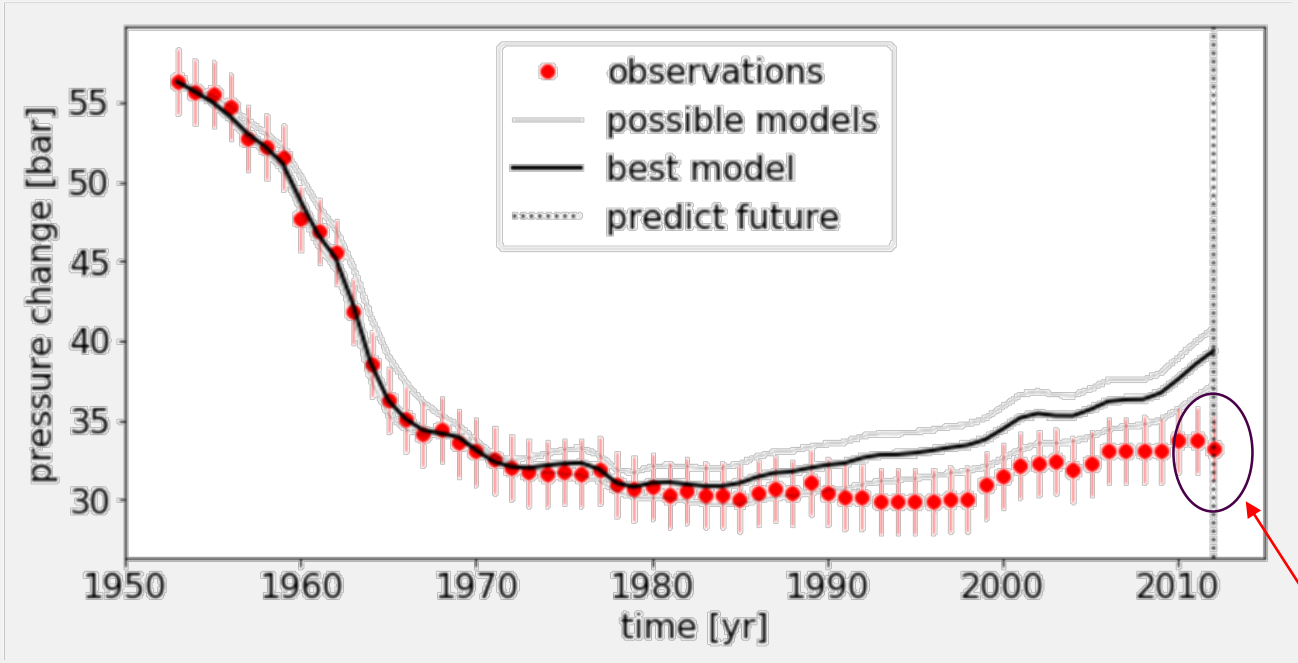
E.g., a simple geothermal model.



Linear Sensitivity analysis



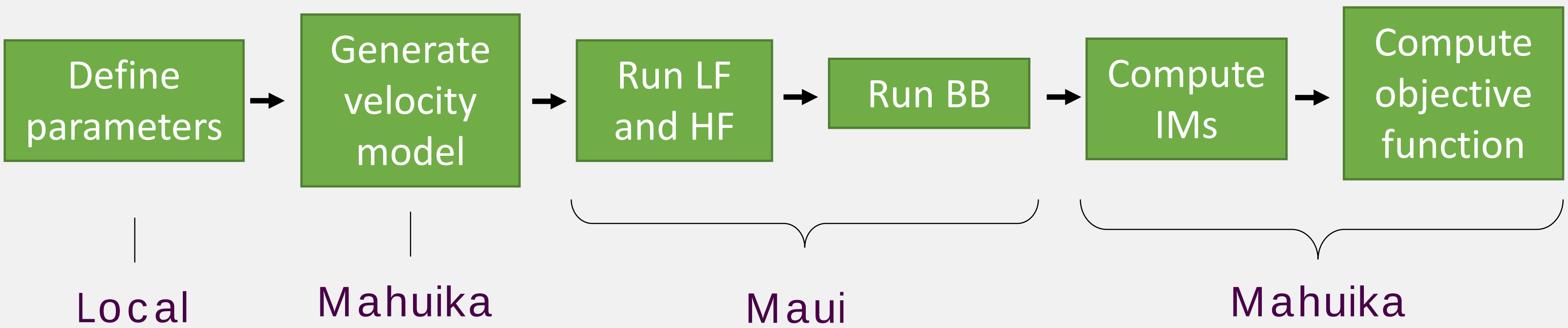
E.g., a simple geothermal model.



Damn!

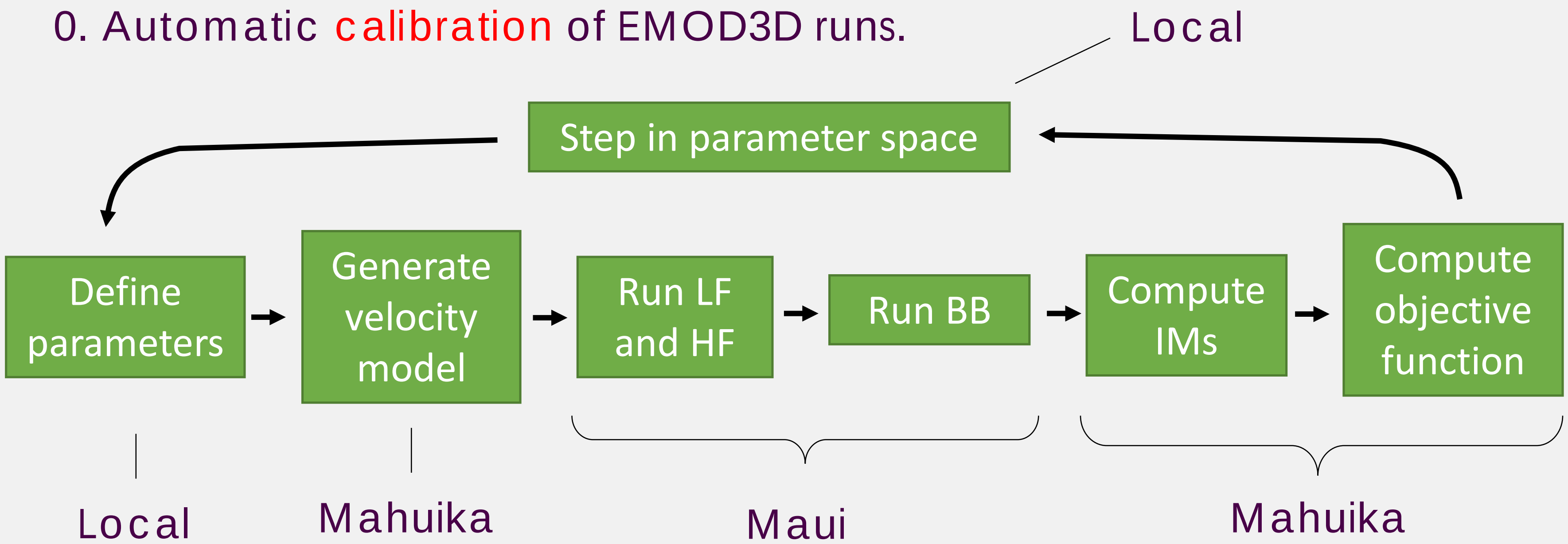
What’s next?

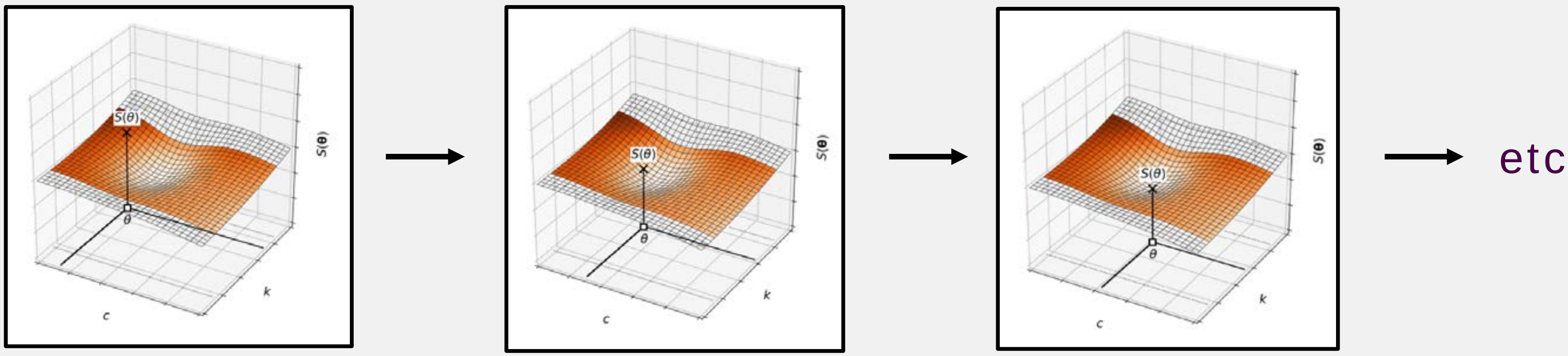
0. Automatic generation, submission, evaluation of EMOD3D runs.



What's next?

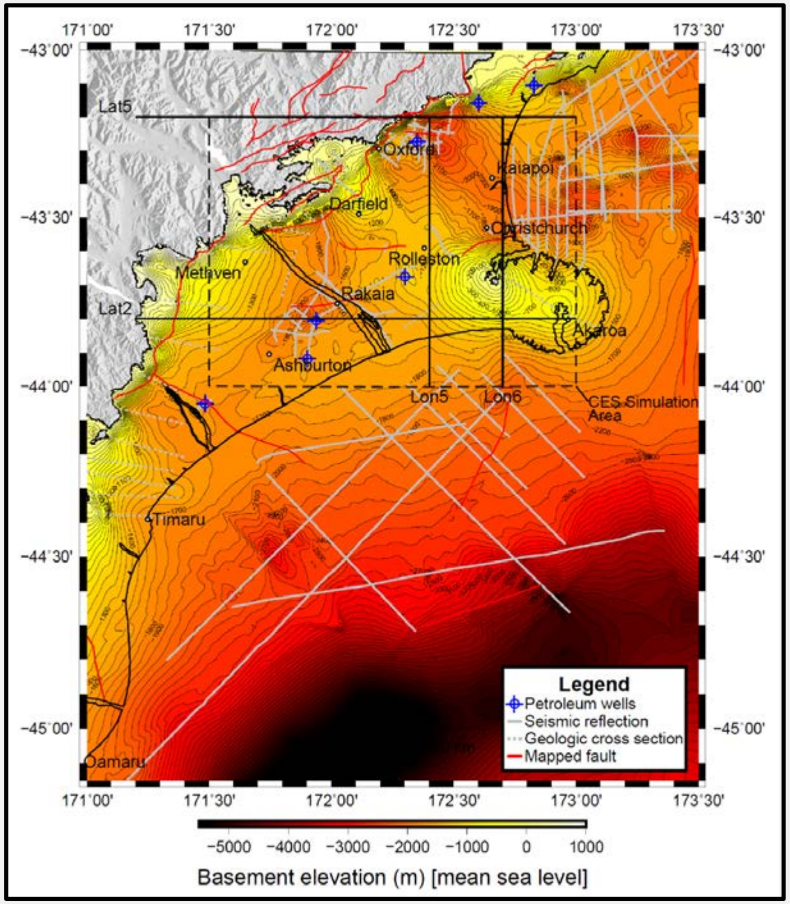
0. Automatic calibration of EMOD3D runs.

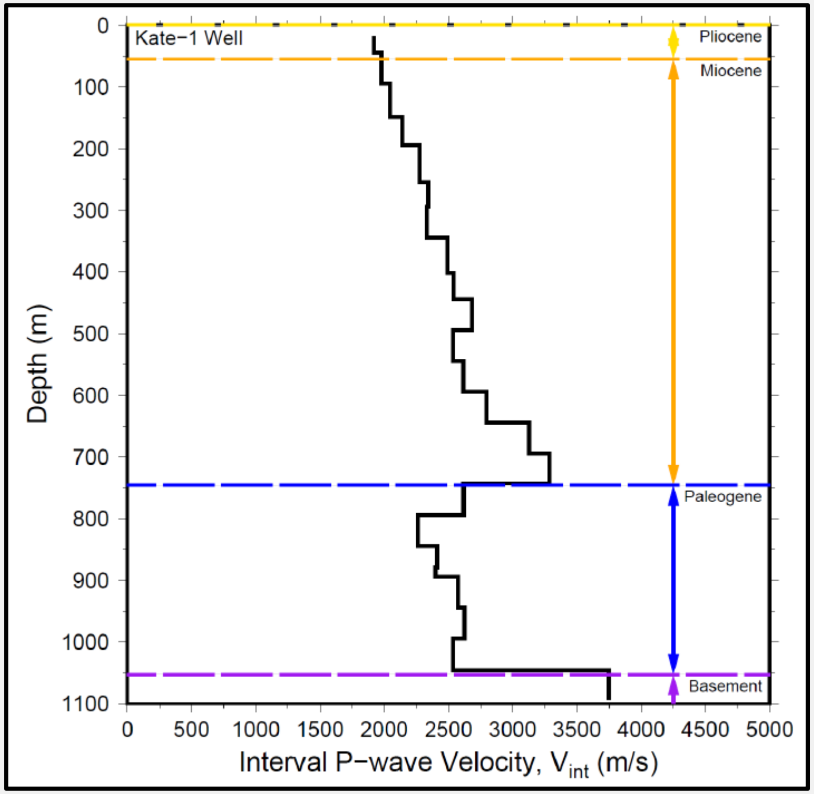




What’s next?

1. Uncertainty of basement contact in Canterbury region.

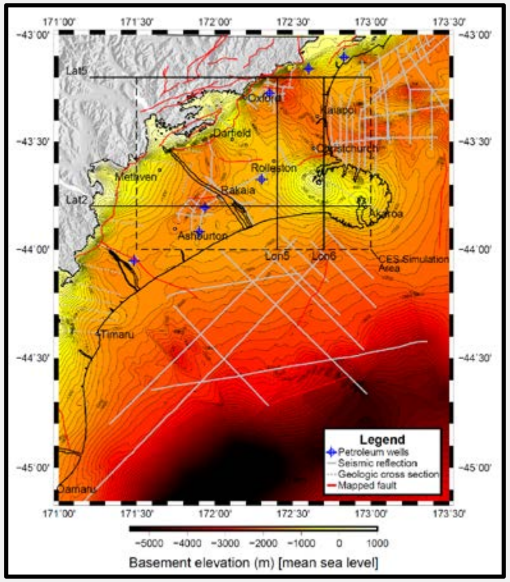




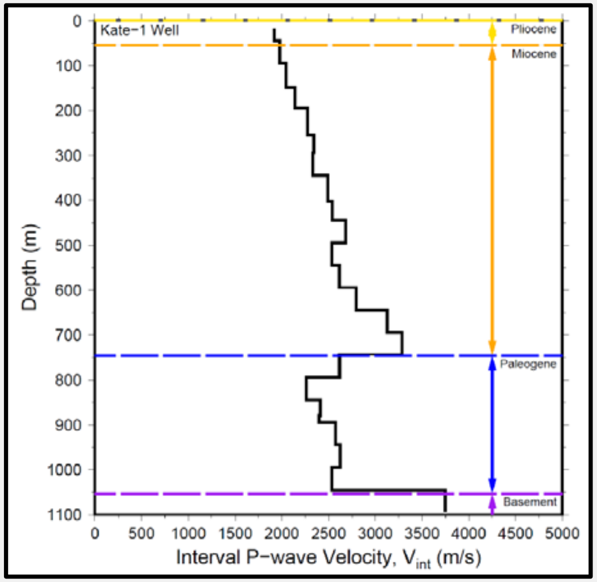
Lee (2017)

What’s next?

1. Uncertainty of basement contact in Canterbury region.



Perturb these.

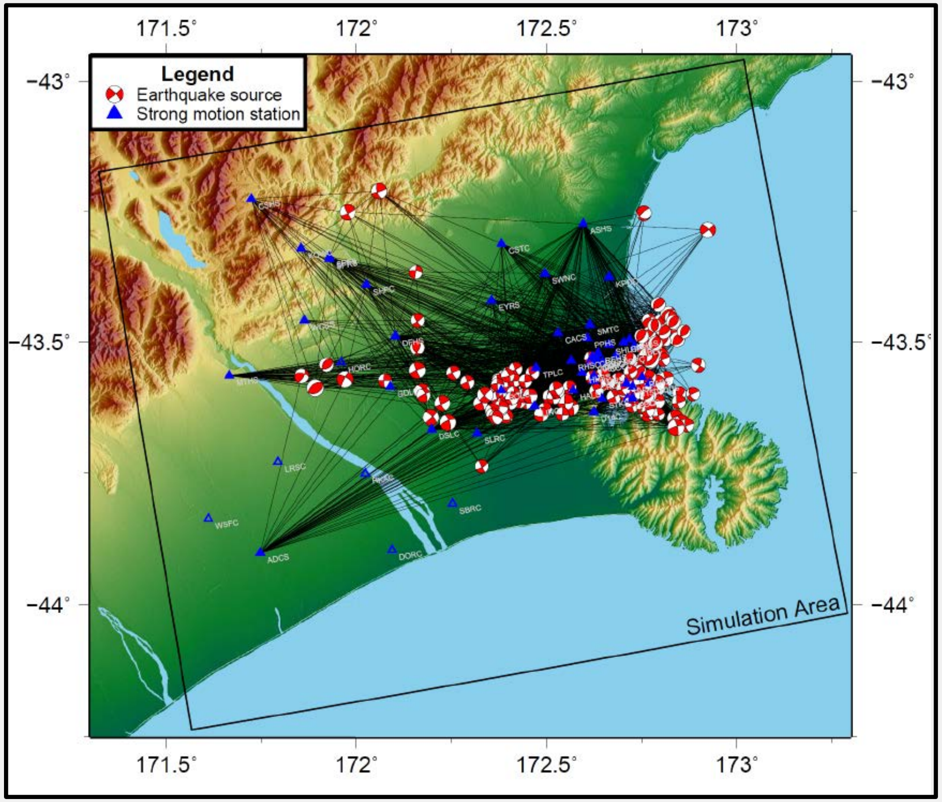


Lee (2017)

What's next?

1. Uncertainty of basement contact in Canterbury region.

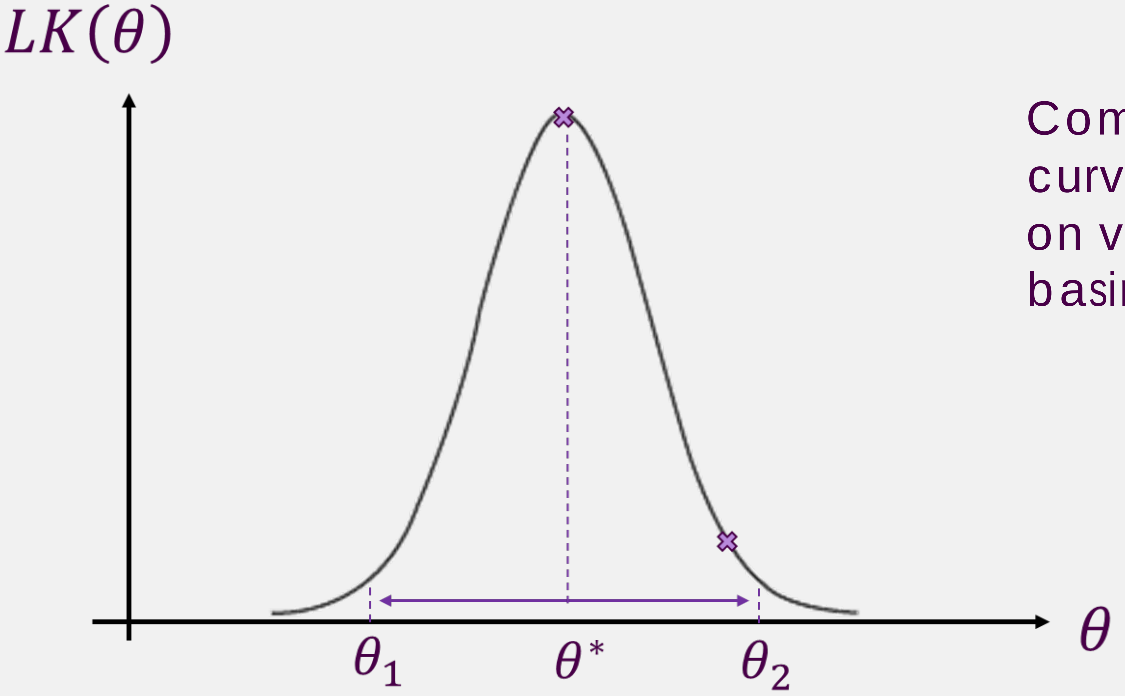
Use small-to-moderate event arrival times as objective function.



Lee (2017)

What's next?

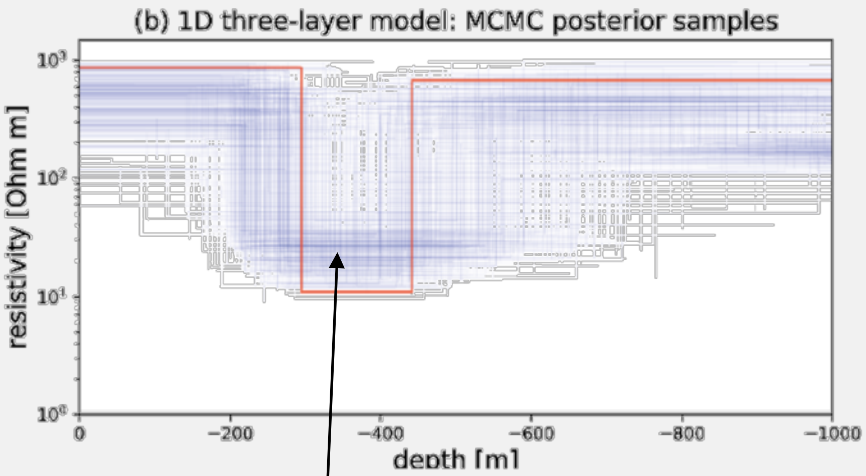
1. Uncertainty of basement contact in Canterbury region.

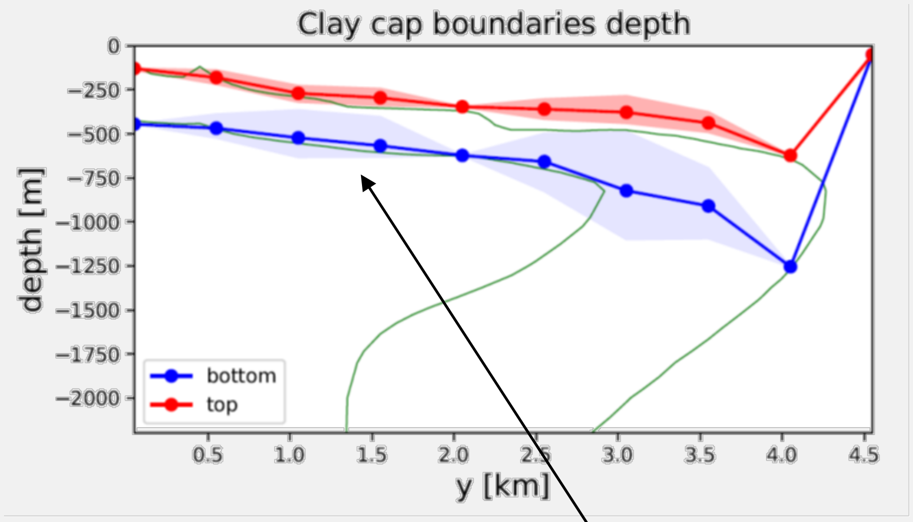


Compute sensitivity curve – obtain bounds on velocity model, basin depth.

Can it work?

Not sure. But we have had luck applying this approach to uncertain inversion of clay cap depth and thickness in geothermal systems.





Using MCMC to invert a series of 1D MT models - obtain probabilistic estimate of conductor depth.*

*MT = magnetotelluric