Brief introductions to

Gridding, Inversion, Tomography

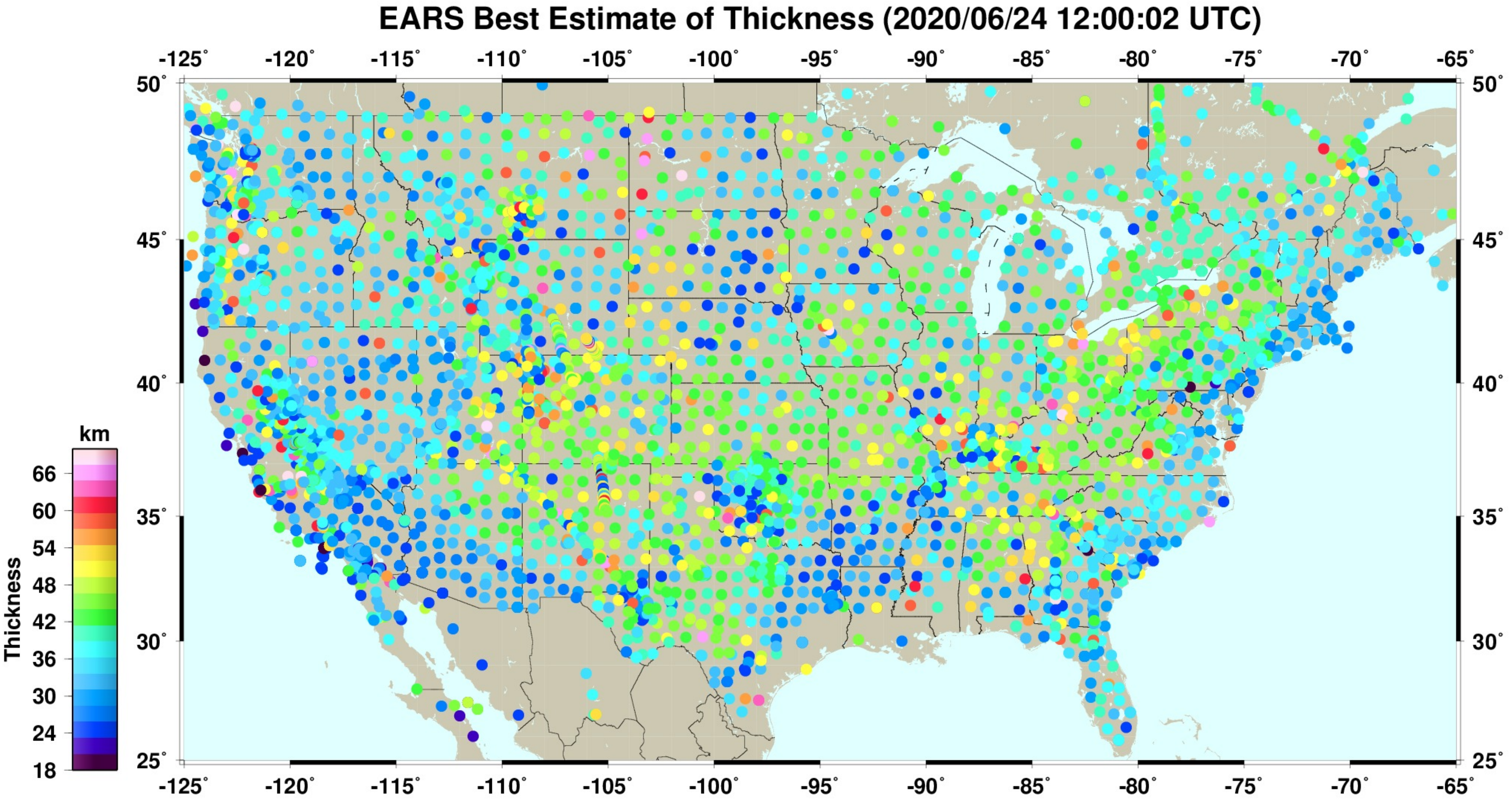
Seismology Summer School

ROSES 2020 -- unit 11 (last one)!

Suzan van der Lee – *Northwestern University*

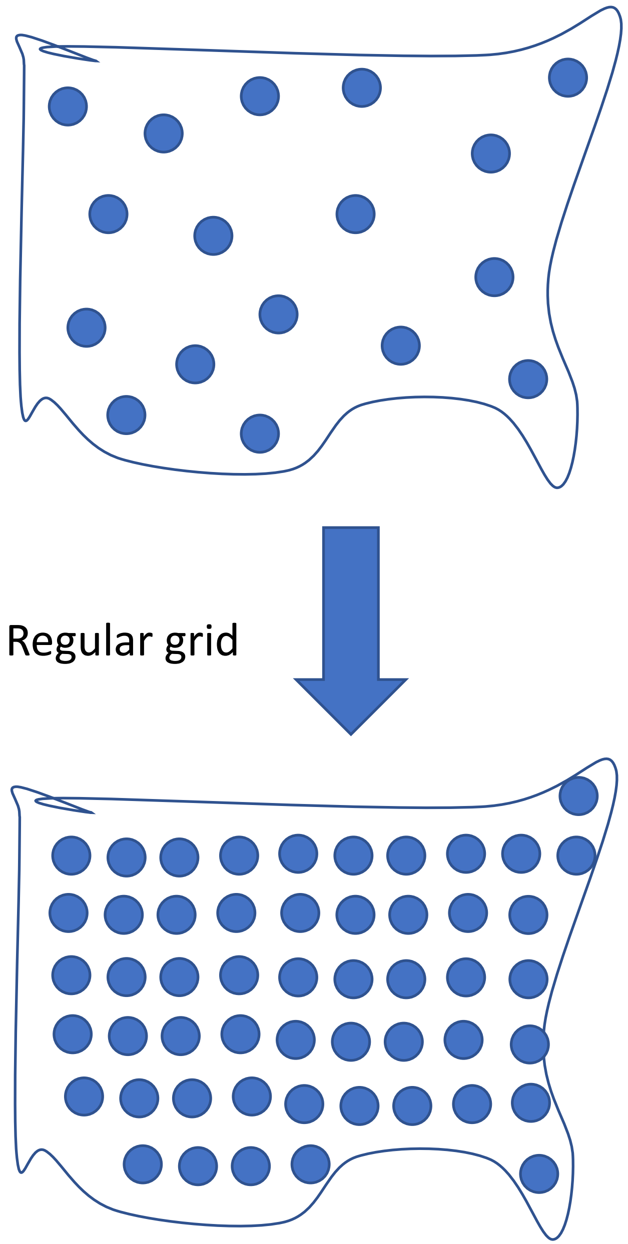
Please ask questions at any time.

Specific application to the depth of the Moho beneath 48 states of the USA



The Intro

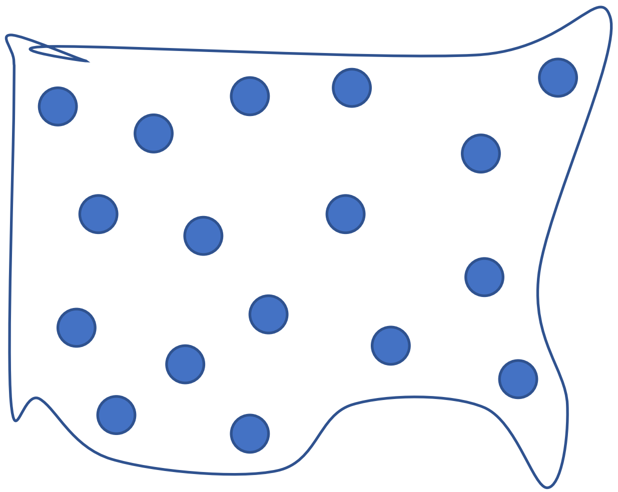
Distribution of points for we have data may be irregular.

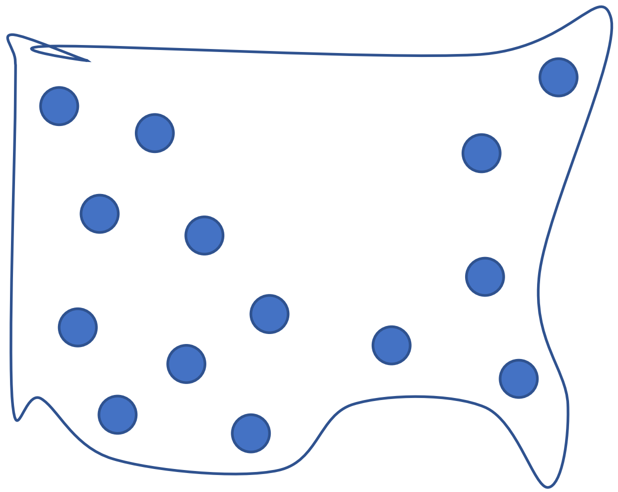


What are some reasons for one or more of us wanting to grid a quantity (e.g. Moho depth) for which we have irregular measurements?

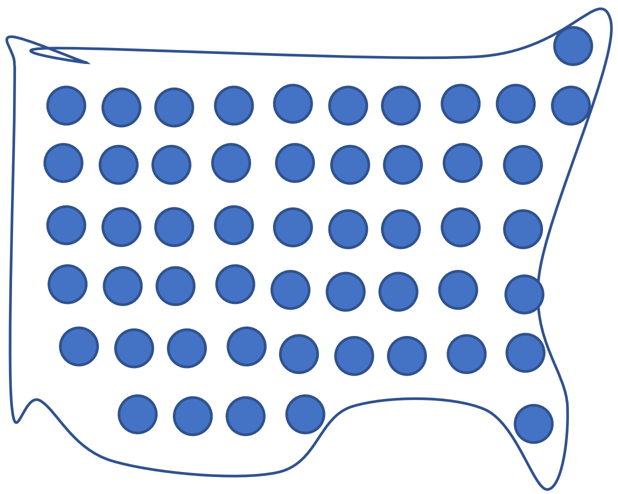
Type answer in zoom chat

Distribution of points for we have data may be irregular, have gaps,

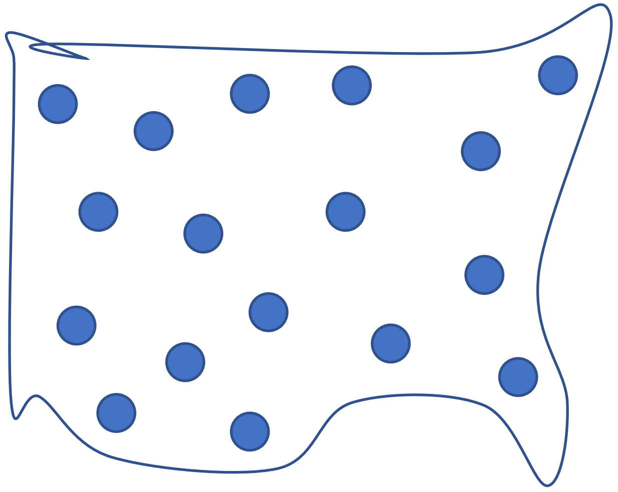


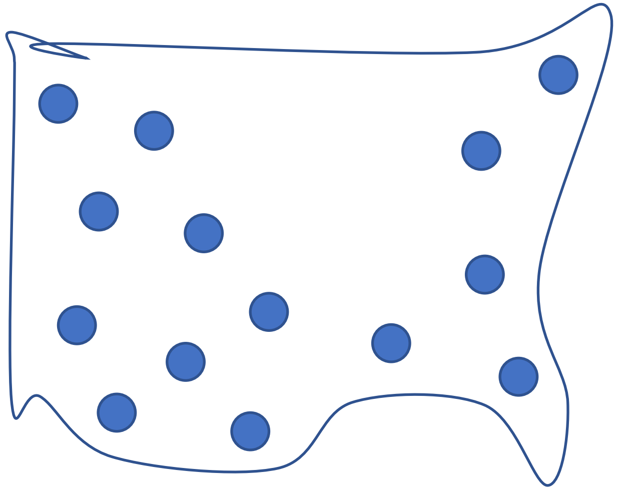


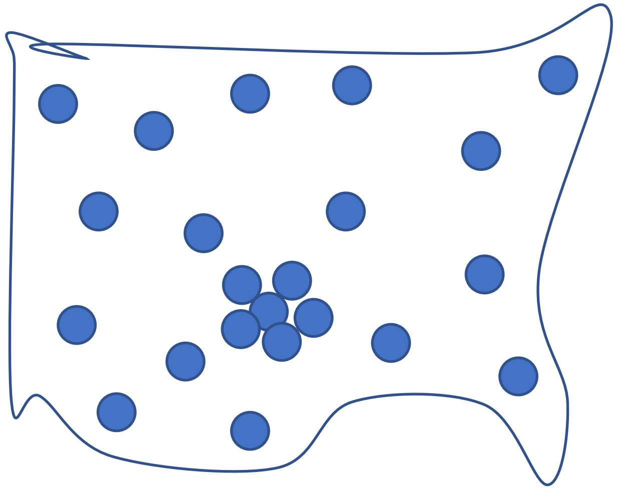
Regular grid



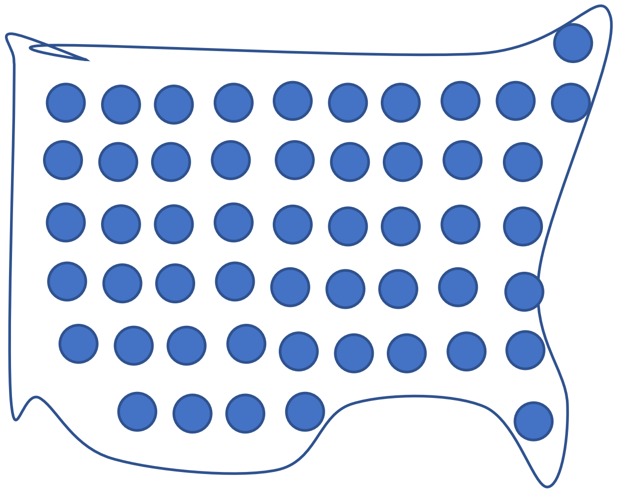
Distribution of points for we have data may be irregular, have gaps, or have concentrations. Data may be inconsistent.



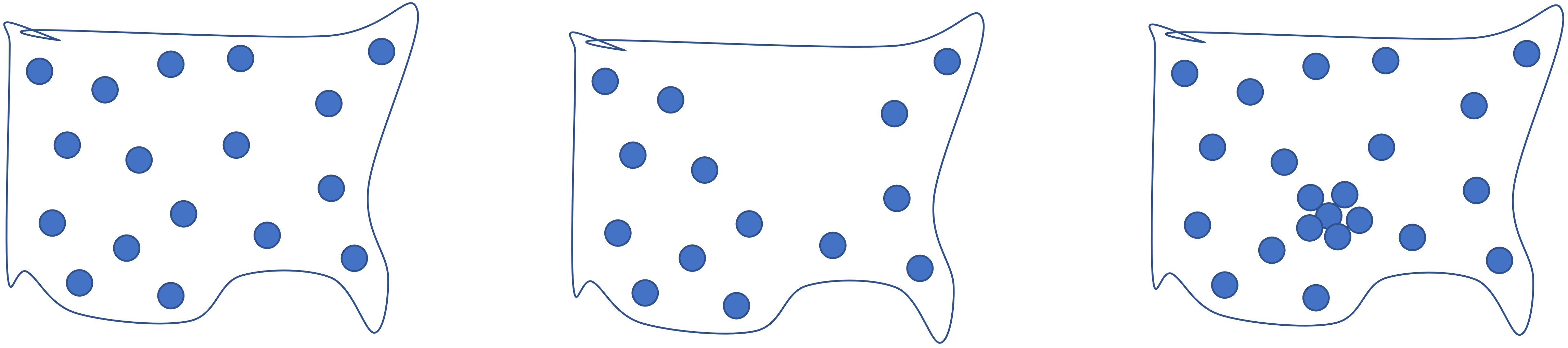




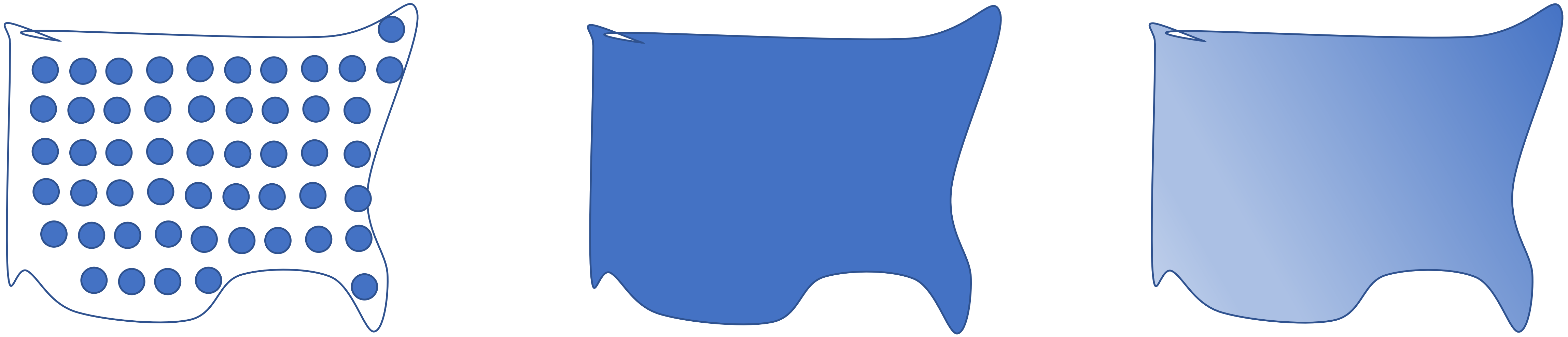
Regular grid



Distribution of points for we have data may be irregular, have gaps, or have concentrations. Data may be inconsistent.



Regular grid can function as a set of local supports for the data as a continuous function of position.



Moho depth $= y$

$$y = f(\underline{x}) + \varepsilon$$

$$f(\underline{x}_k) = \sum_{i=1}^{m} w_i \cdot h_i(\underline{x}_k)$$

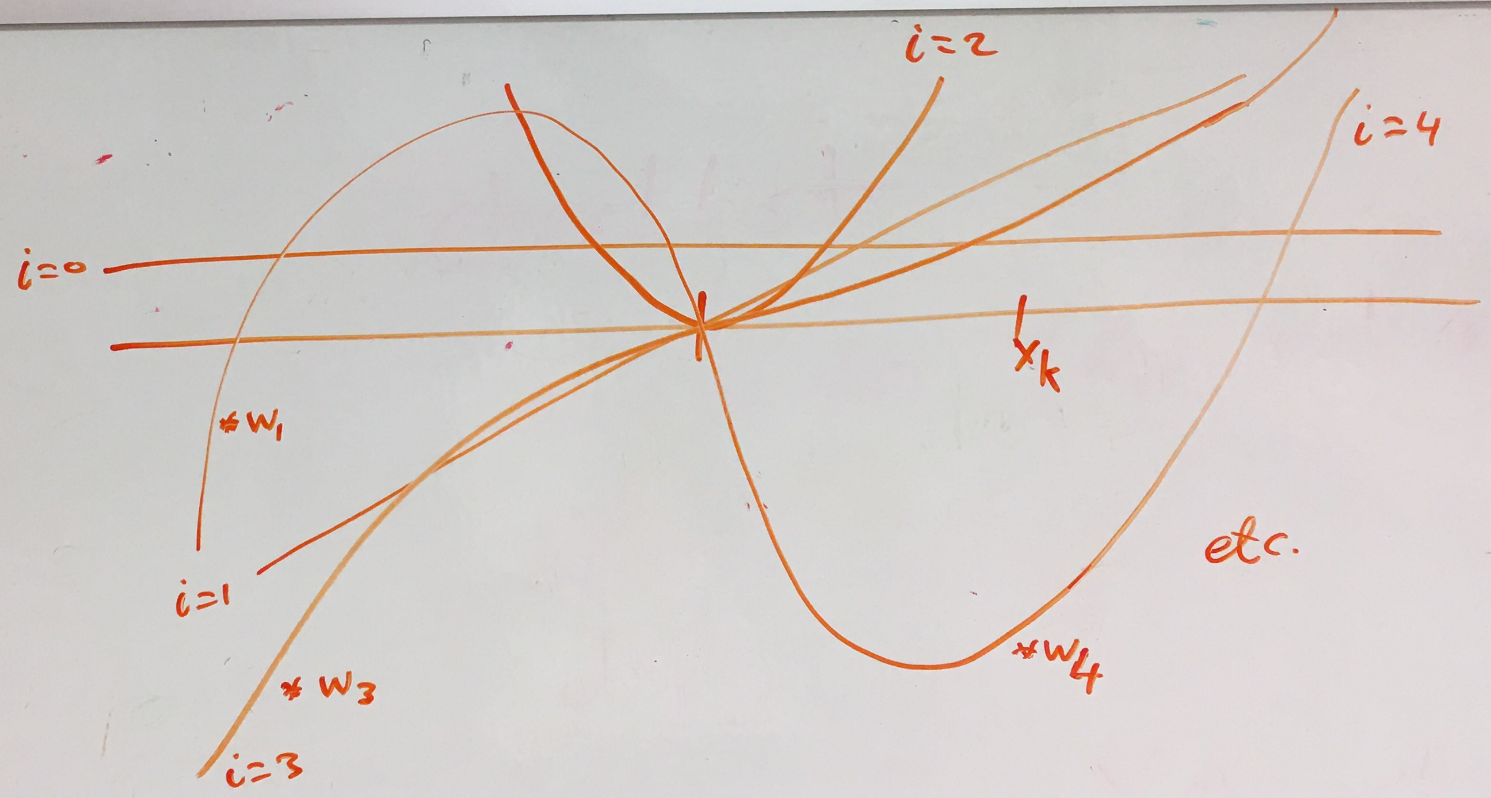
Possible choices for $h_i(\underline{x}_k)$:

$$= \sum_{i=1}^{m} w_i \cdot p_i(\underline{x}_k)$$ polynomials

$$= \sum_{i=1}^{m} w_i \cdot e^{i\, \underline{k}_i \cdot \underline{x}_k}$$ harmonics

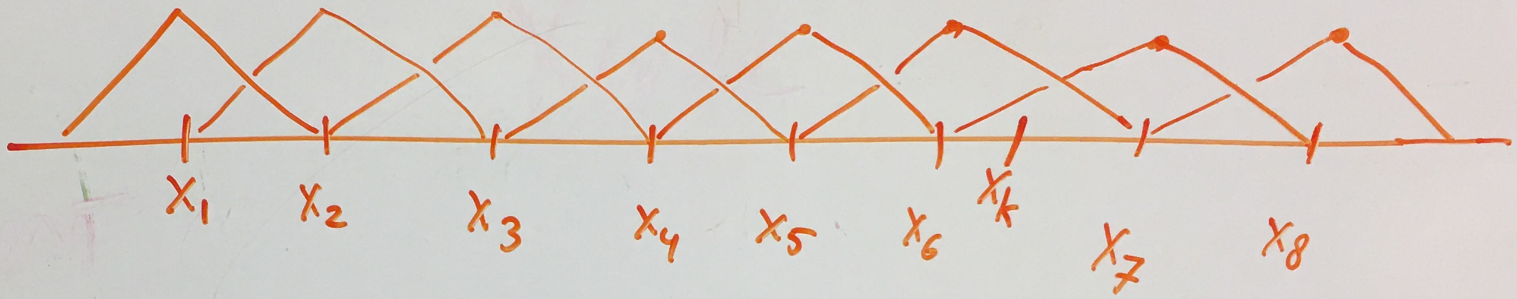
$$= \sum_{i=1}^{m} w_i \cdot h(\underline{x}_i - \underline{x}_k)$$ local basis functions

polynomials:



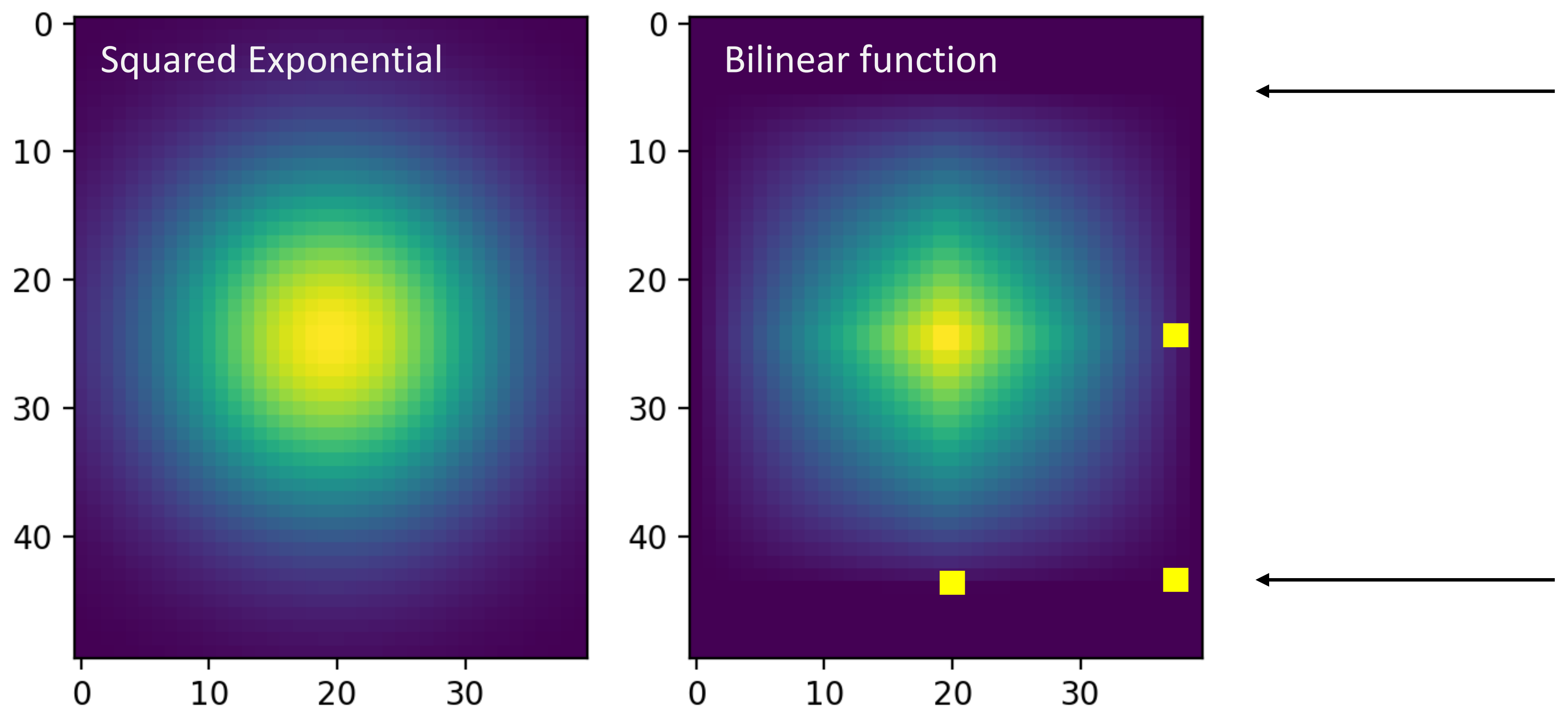
harmonics

local basis functions:

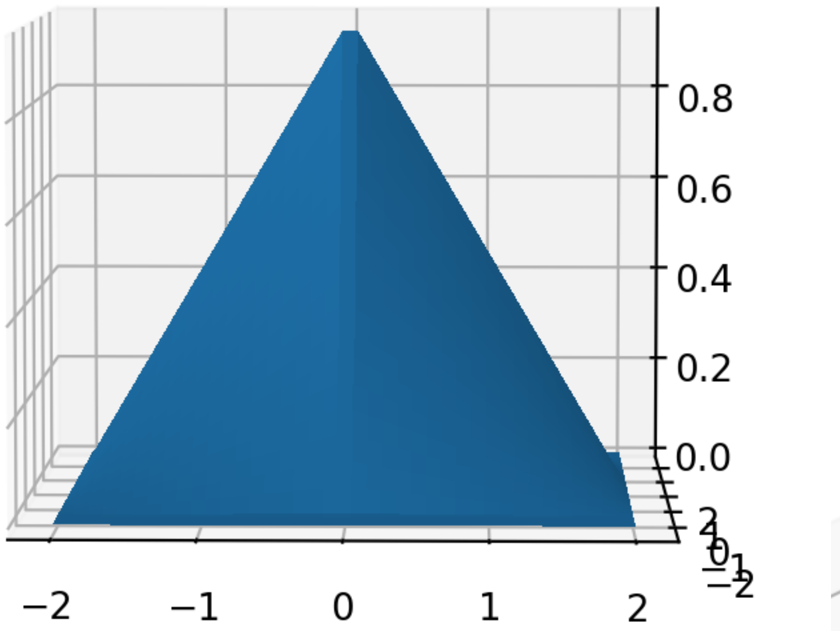


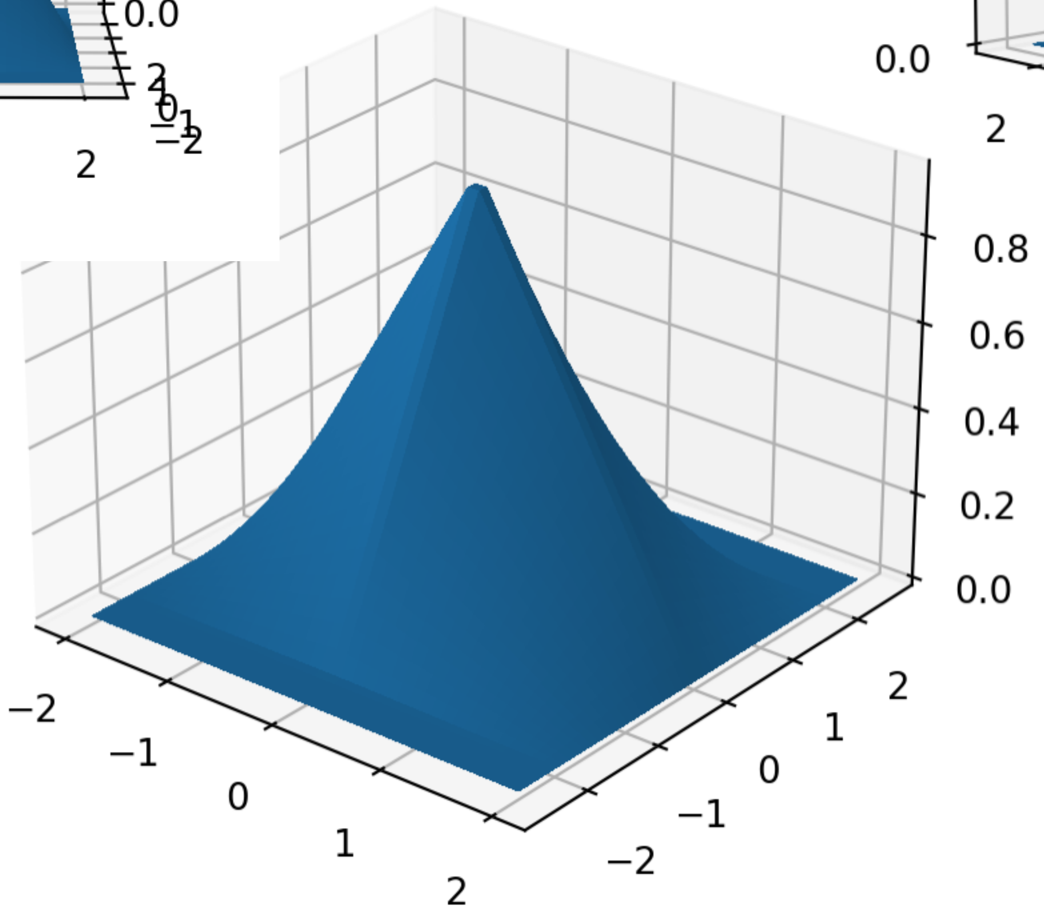
bi-linear

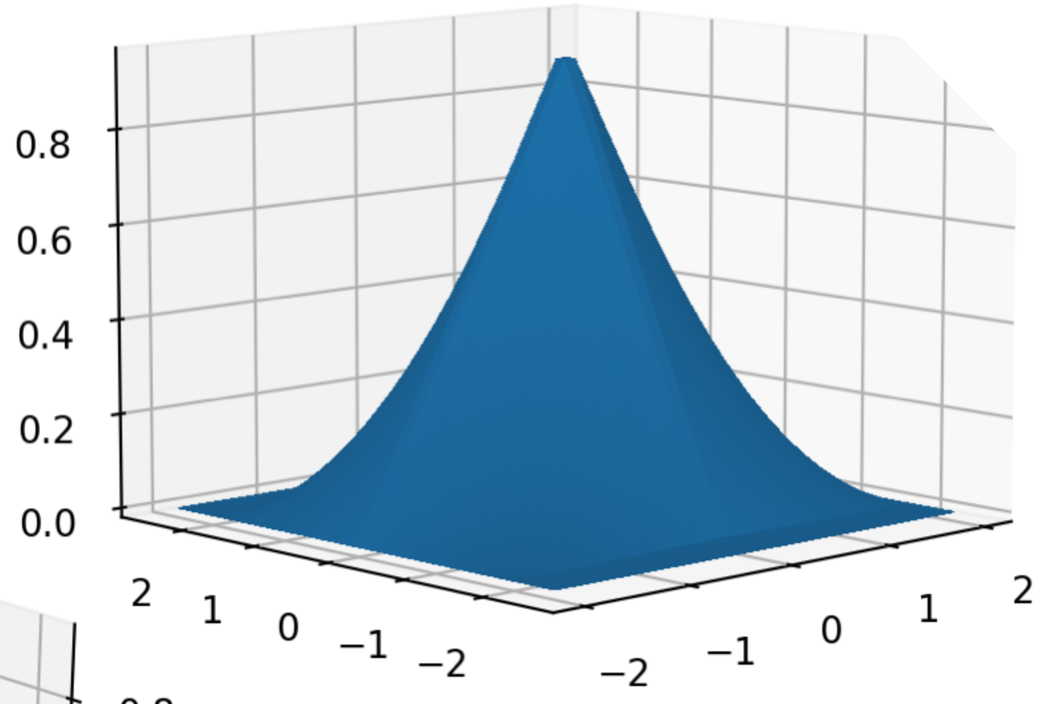
Gaussian/squared-exponential (SE)


Squared Exponential
Bilinear function
0
10
20
30
40
0
10
20
30

Bilinear function







The Maths

a convenient property of the bilinear basis functions is that $h_i(\underline{x}_j) = 0$ for $j \in \{1, 2, 3, \ldots, m\}$ and $i \neq j$. That means that $w_j = f(\underline{x}_j)$, i.e. the Moho depth at $\underline{x}_j$

Also, $\sum_{i=1}^{m} h(\underline{x}_i - \underline{x}_k) = 1$ for all k!

let's assume we wish to use SE or spline functions. Then

$$\sum_{i=1}^{m} h(\underline{x}_i - \underline{x}_k) \neq 1 \quad \text{and}$$

we should write:

$$f(\underline{x}_k) = \sum_{i=1}^{m} w_i \frac{h_i(\underline{x}_k)}{\sum_{i=1}^{m} h_i(\underline{x}_k)}$$

$$f(\underline{x}) = \sum_{i=1}^{m} h(\underline{x}_i - \underline{x}) \cdot w_i$$

the only $f(\cdot)$ values we know are the measurements y

$$f(\underline{x}_k) = \sum_{i=1}^{m} h(\underline{x}_i - \underline{x}_k) \cdot w_i$$

$$\underline{f} = \underline{\underline{G}} \cdot \underline{w} \qquad \text{where } G_{ki} = h(\underline{x}_i - \underline{x}_k)$$

note: f is a linear function of $\underline{w}$

$$\underline{y} - \underline{\underline{G}}\,\underline{w} = \varepsilon$$

minimize the errors (loss) $\underline{\varepsilon} = \underline{y} - \underline{\underline{G}}\,\underline{w}$

$\rightarrow$ find a $\underline{w}$ for which $\underline{\underline{G}} \cdot \underline{w} = \underline{y}$

OK, $\underline{w} = \underline{\underline{G}}^{-1} \underline{y}$? no

minimize length^2 of vector :

$$\min E(\underline{w}) = \| \underline{y} - \underline{\underline{G}}\,\underline{w} \|^2 = \sum_{i=1}^{n} \left(y_i - G_{ij} w_j \right)^2$$

note: $E(\underline{w})$ is a quadratic function of $\underline{w}$

$$\frac{\partial E(\underline{w})}{\partial w_k} = \quad \ldots\ldots \quad = \emptyset$$

$\rightarrow$ for all k

$\rightarrow$ for all k

$$\Longrightarrow \quad \underline{w} = (\underline{\underline{G}}^T \underline{\underline{G}})^{-1} \underline{\underline{G}}^T \underline{y}$$

trick: $y = Gw$

$$G^T y = G^T G w$$

$$(G^T G)^{-1} G^T y = w$$

$$\boxed{\underline{W} = (\underline{\underline{G}}^T \underline{\underline{G}})^{-1} \underline{\underline{G}}^T \underline{y}}$$

What if errors in $\underline{y}$ are not the same?

$\downarrow$

standard deviation σ_i

$$\underline{\underline{C}}_d = \begin{pmatrix} \sigma_1^2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 & & & \phi & \cdot \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & \phi & & & \cdot & \cdot \\ 0 & \cdot & \cdot & & \cdot & \sigma_n^2 \end{pmatrix}$$

$$\underbrace{\underline{\underline{C}}_d^{-1/2}\,\underline{y}}_{\underline{y}'} = \underbrace{\underline{\underline{C}}_d^{-1/2}\,\underline{\underline{G}}}_{G'}\,\underline{w}$$

$$\underline{w} = \left(\underline{\underline{G}}'^T\,\underline{\underline{G}}'\right)^{-1}\underline{\underline{G}}'^T\,\underline{y}'$$

What if I have prior expectations for $\underline{w}$, e.g. when not enough data is available?

$$\underline{\underline{C}}_m = \begin{pmatrix} A^2 & 0 & \cdots & & 0 \\ 0 & A^2 & & & \vdots \\ \vdots & & \ddots & \emptyset & \vdots \\ \vdots & \emptyset & & \ddots & \vdots \\ 0 & \cdots & & \cdots & A^2 \end{pmatrix}$$

$$\begin{pmatrix} \underline{y} \\ \cdots \\ \underline{0} \end{pmatrix} = \underbrace{\begin{pmatrix} G \\ \cdots \\ C_m^{-1/2} \end{pmatrix}}_{\underline{\underline{G}}'} \begin{pmatrix} \underline{w} \end{pmatrix}$$

$$\downarrow$$

$$\underline{y}'$$

$$\Rightarrow \underline{w} = \left(\underline{\underline{G}}'^{T} \underline{\underline{G}}'\right)^{-1} \underline{\underline{G}}'^{T} \underline{y}'$$

combined:

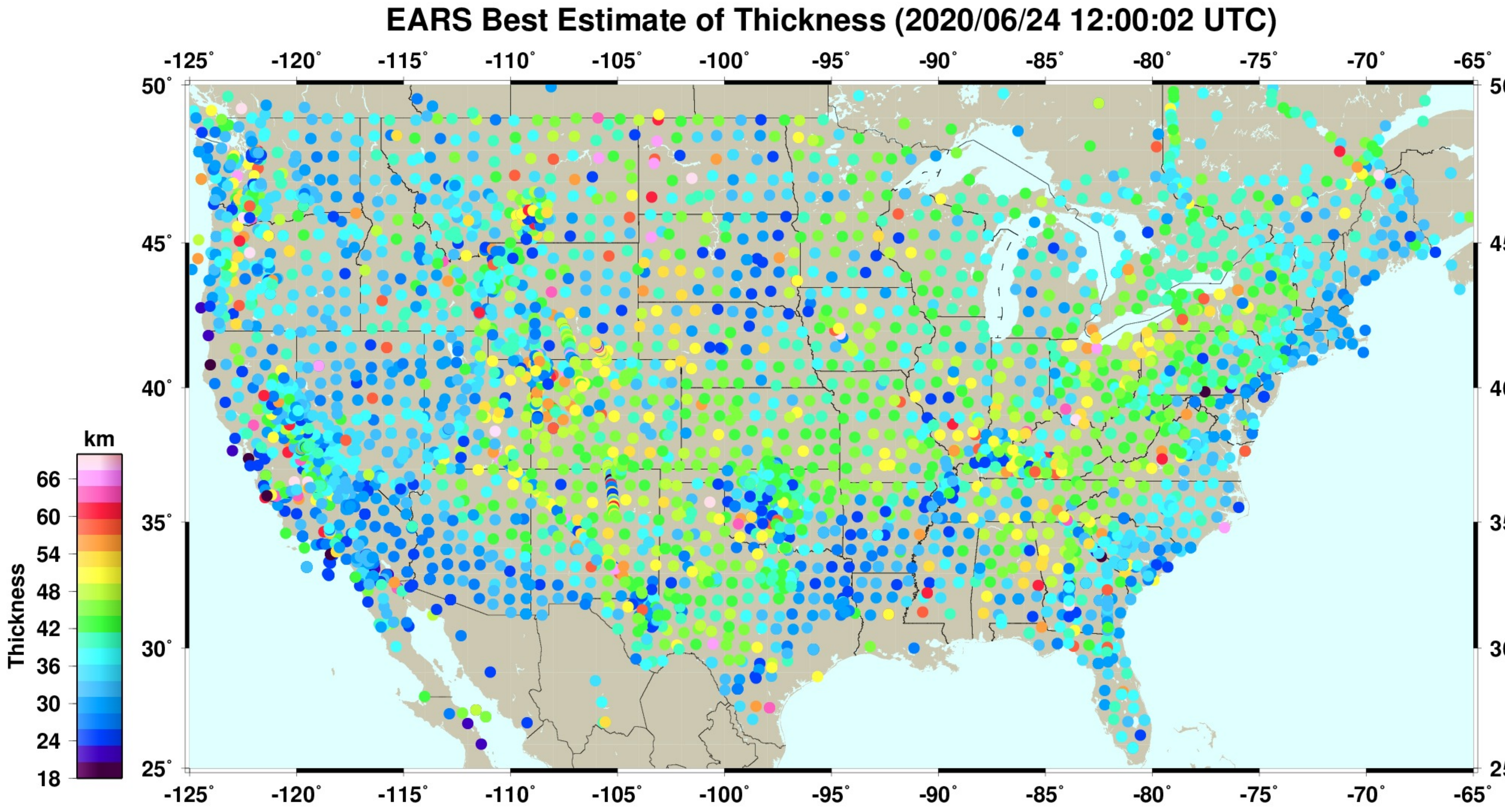
$$\underline{W} = \left(\underline{\underline{G}}^T \underline{\underline{C}}_d^{-1} \underline{\underline{G}} + \underline{\underline{C}}_m^{-1}\right)^{-1} \underline{\underline{G}}^T \underline{\underline{C}}_d^{-1} \underline{y}$$

Weighting by the inverse square root of the Data Covariance matrix helps to deal with data errors that are not drawn from the same distribution (we have more precise or accurate knowledge of some Moho depths than of others).

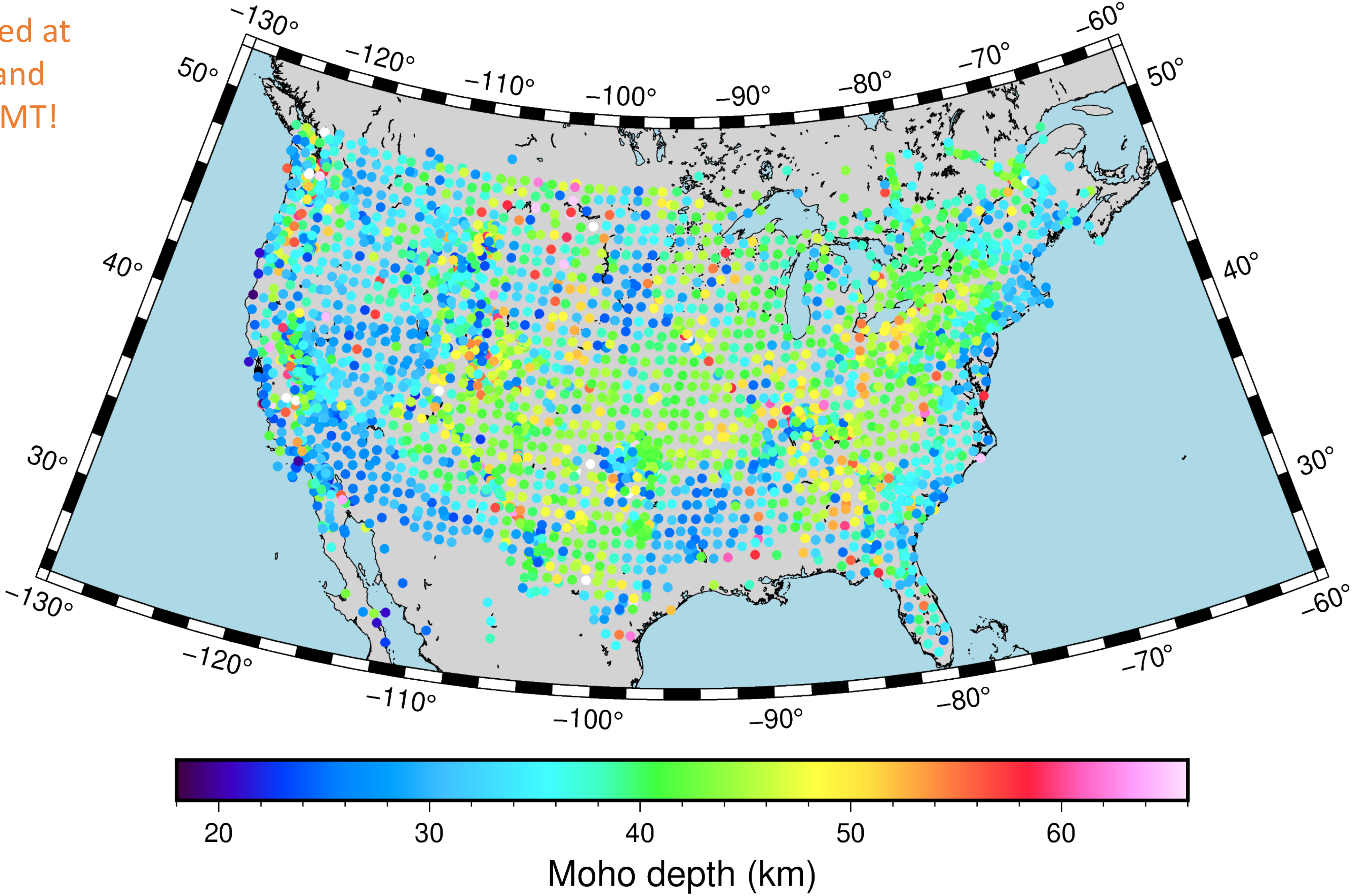
Weighting by the inverse square root of the Model Covariance matrix helps to deal with spatial gaps in data (blank regions). Assume the simplest Moho that is consistent with prior knowledge. Damping, flattening, smoothing, …

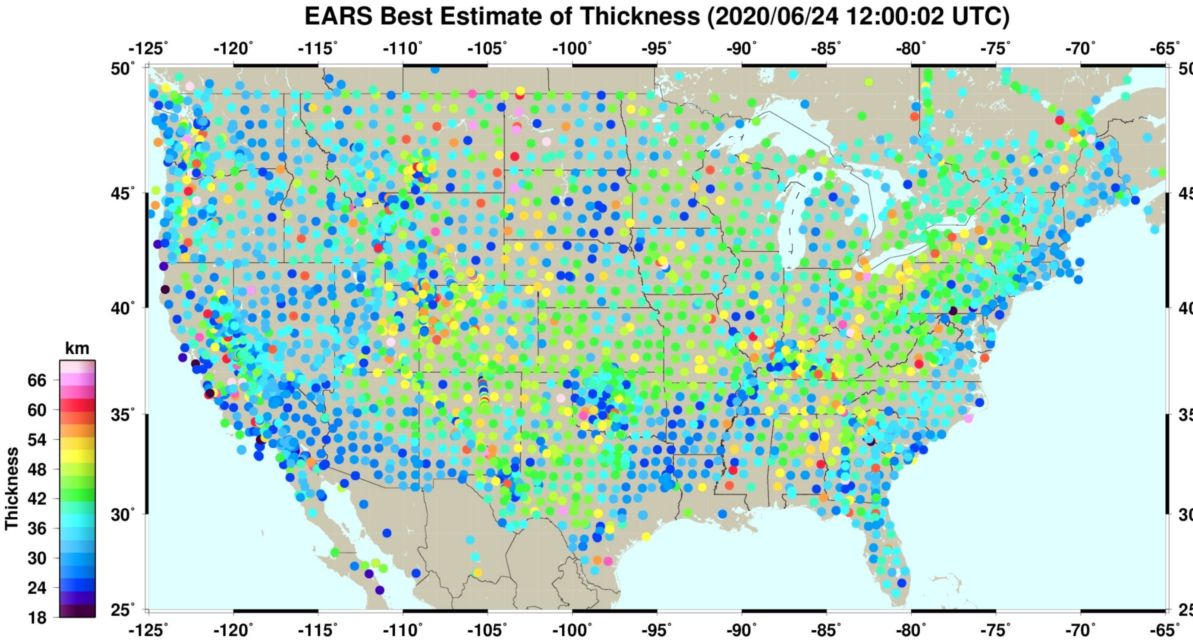
The Data

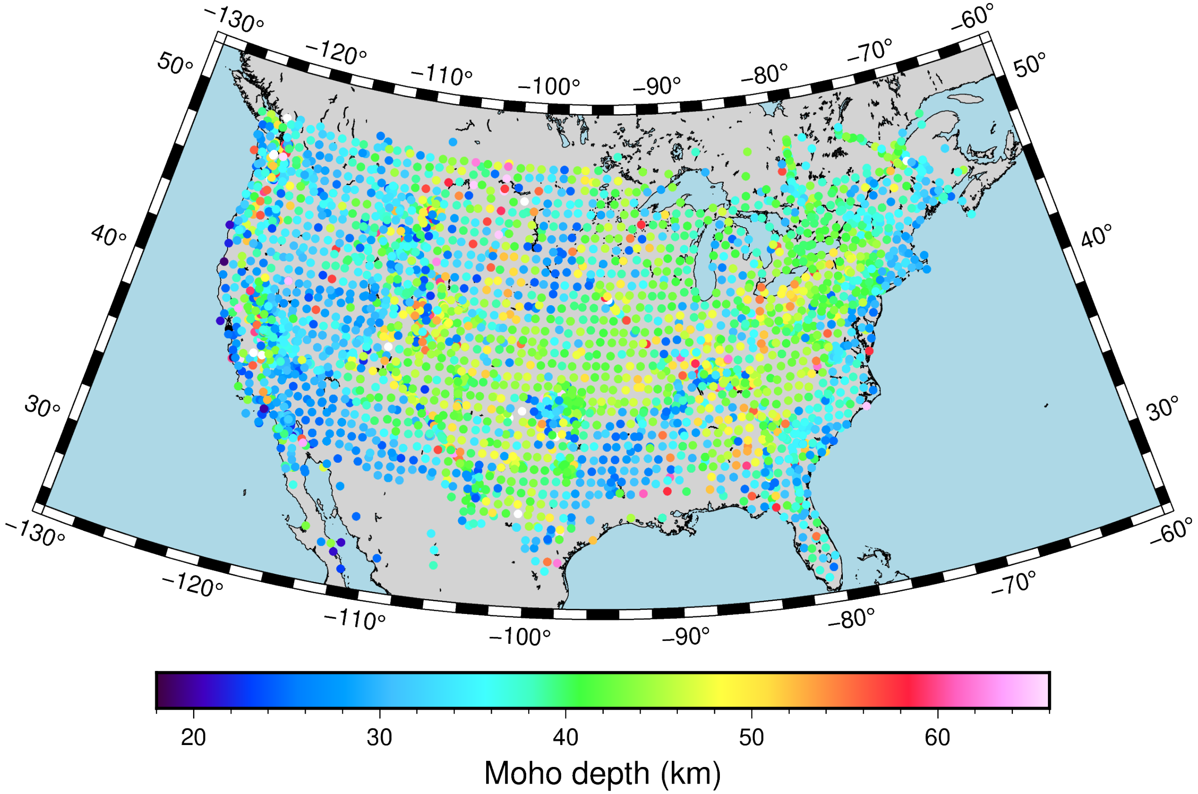
Provided by EARS data service



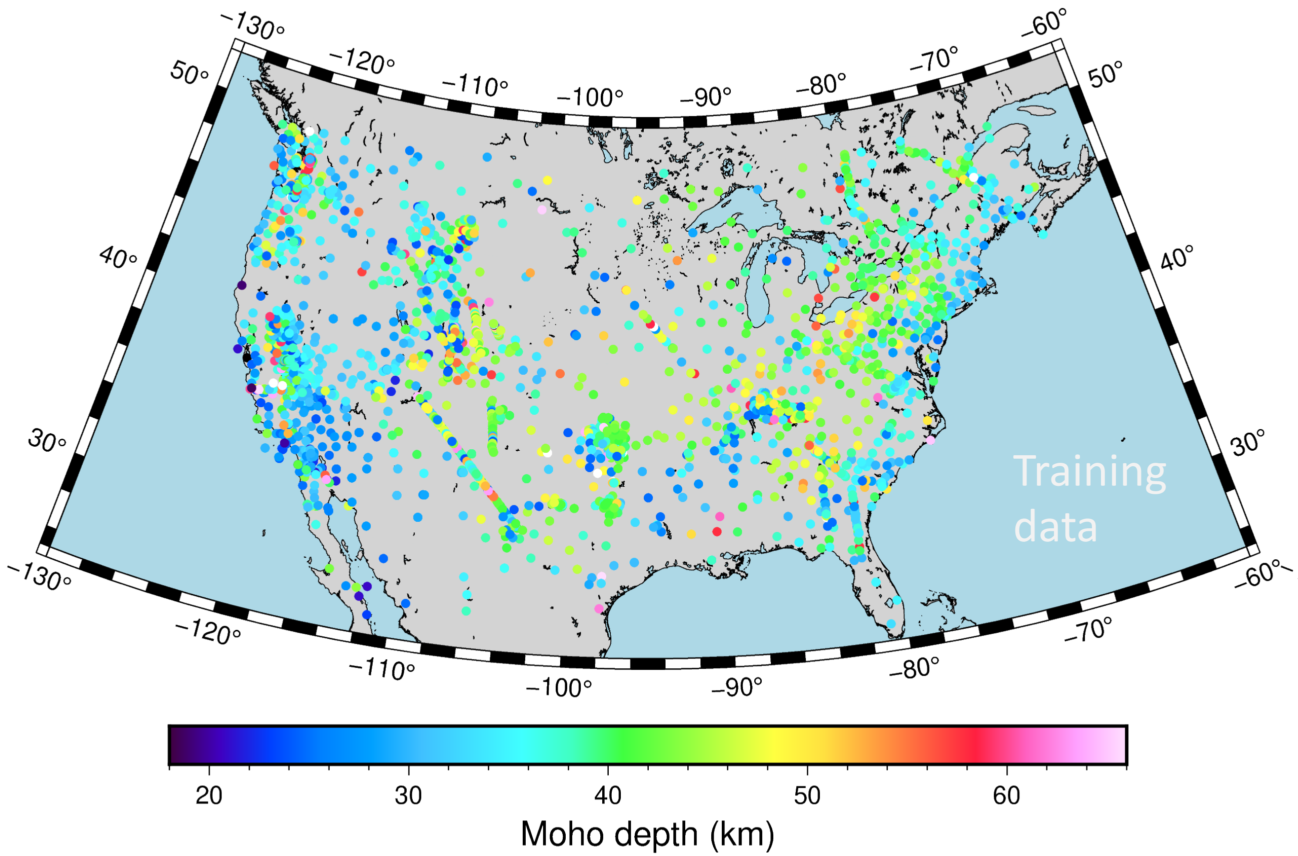
EARS downloaded at an earlier date and Plotted with pGMT!

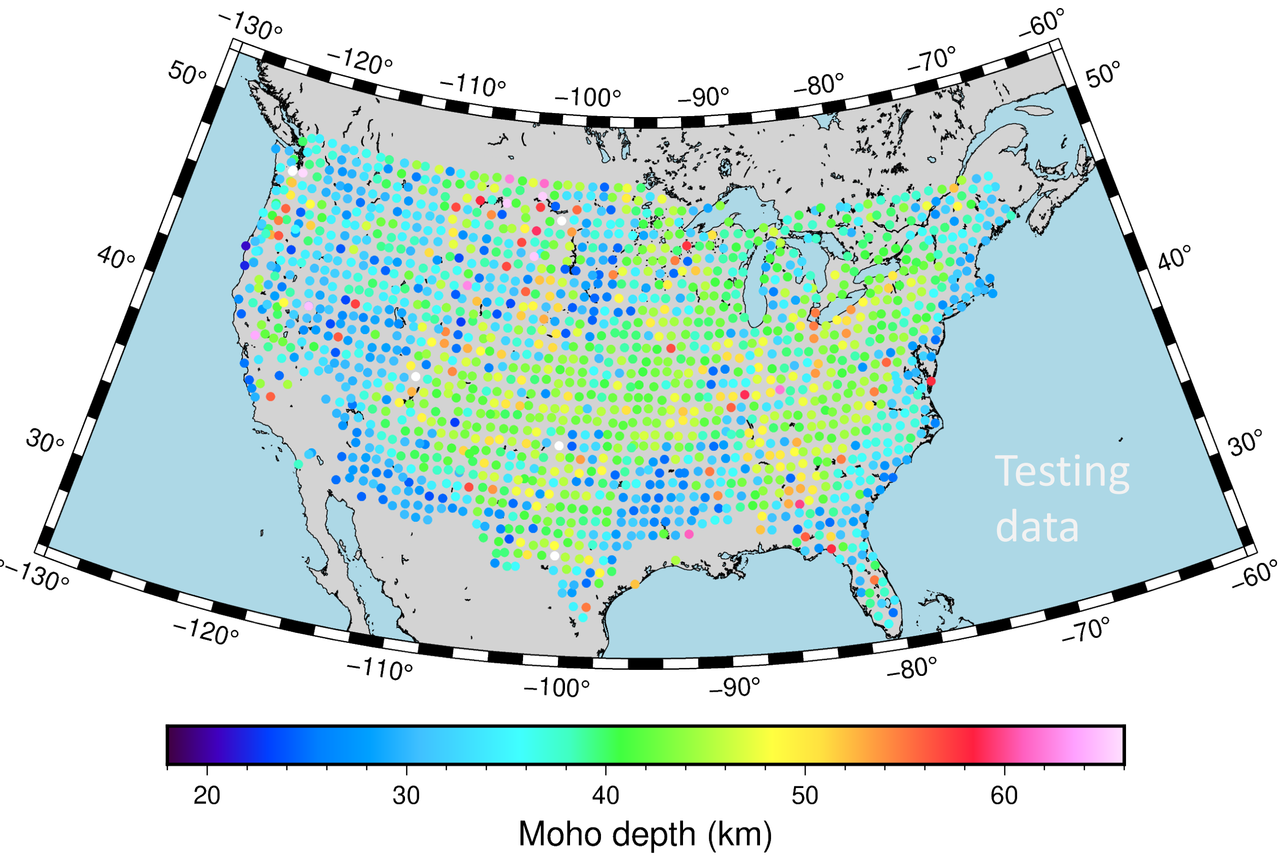




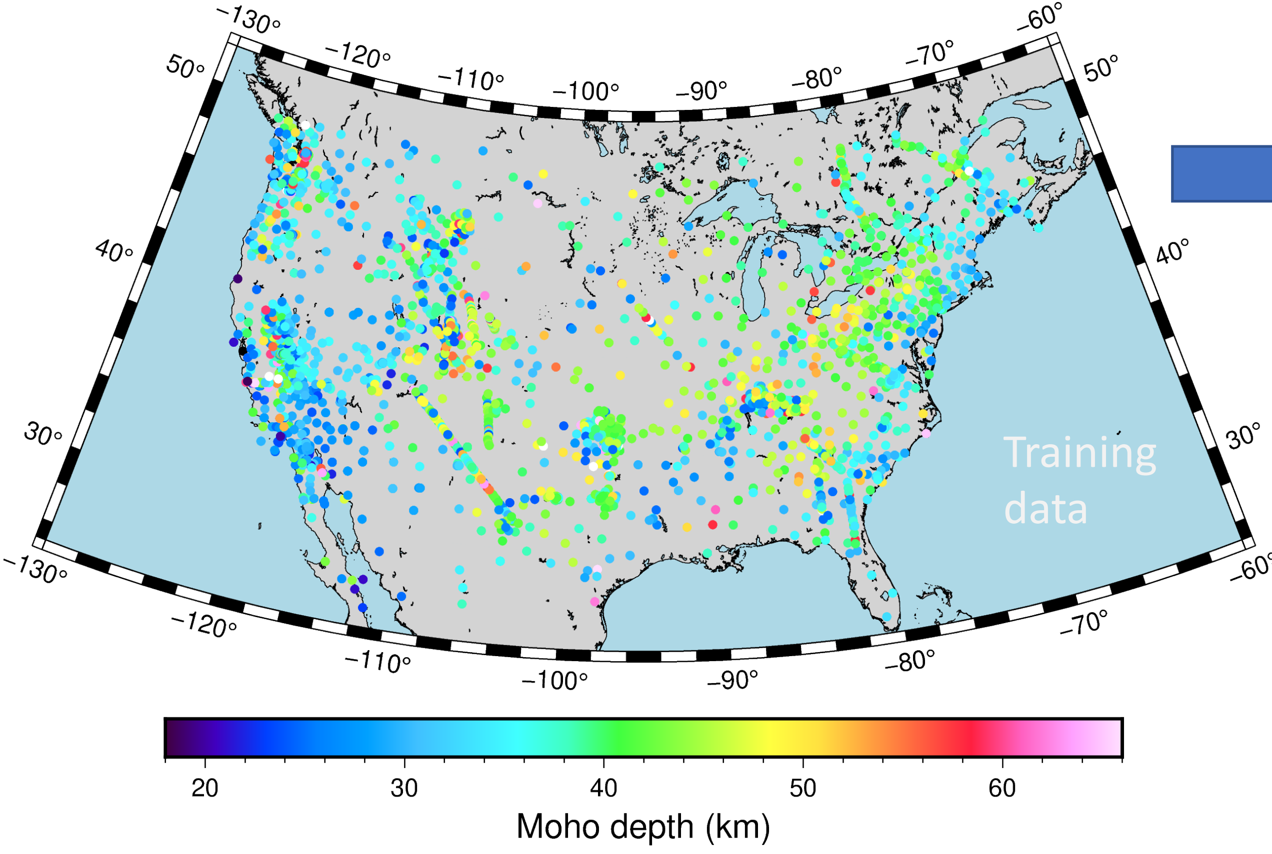


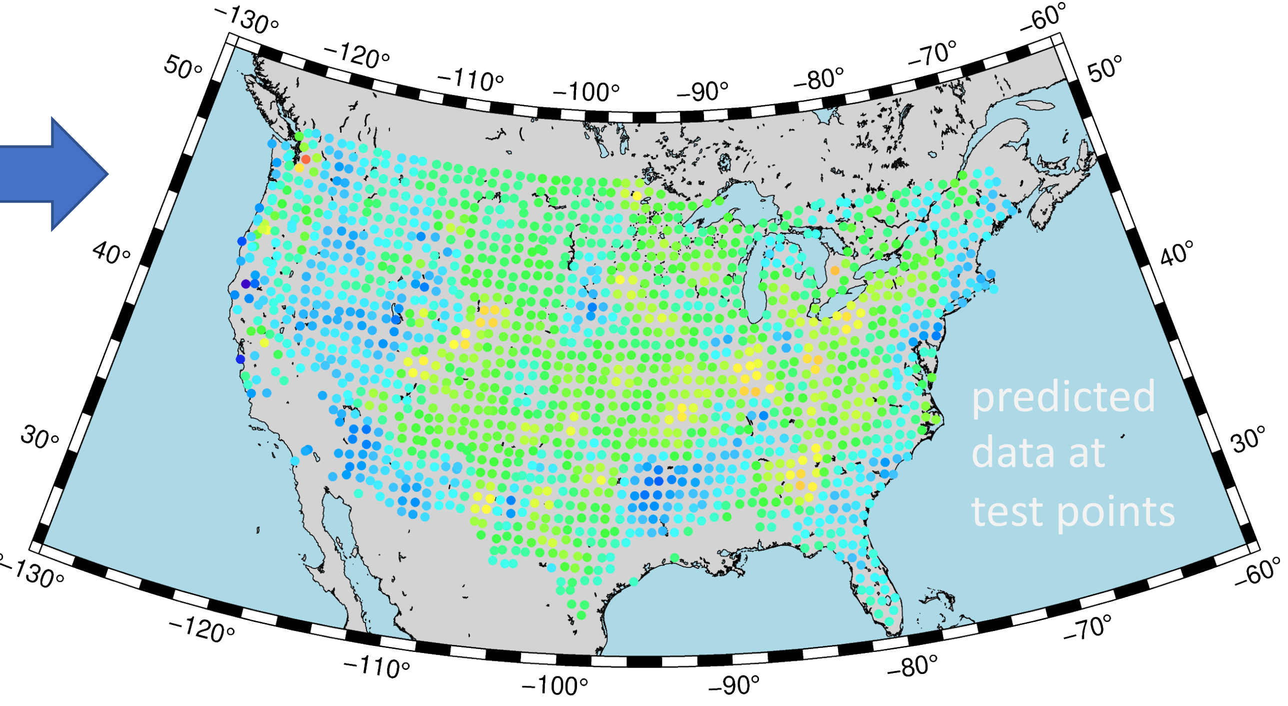
Next: Split EARS data into estimates for USArray TA stations and non-TA stations

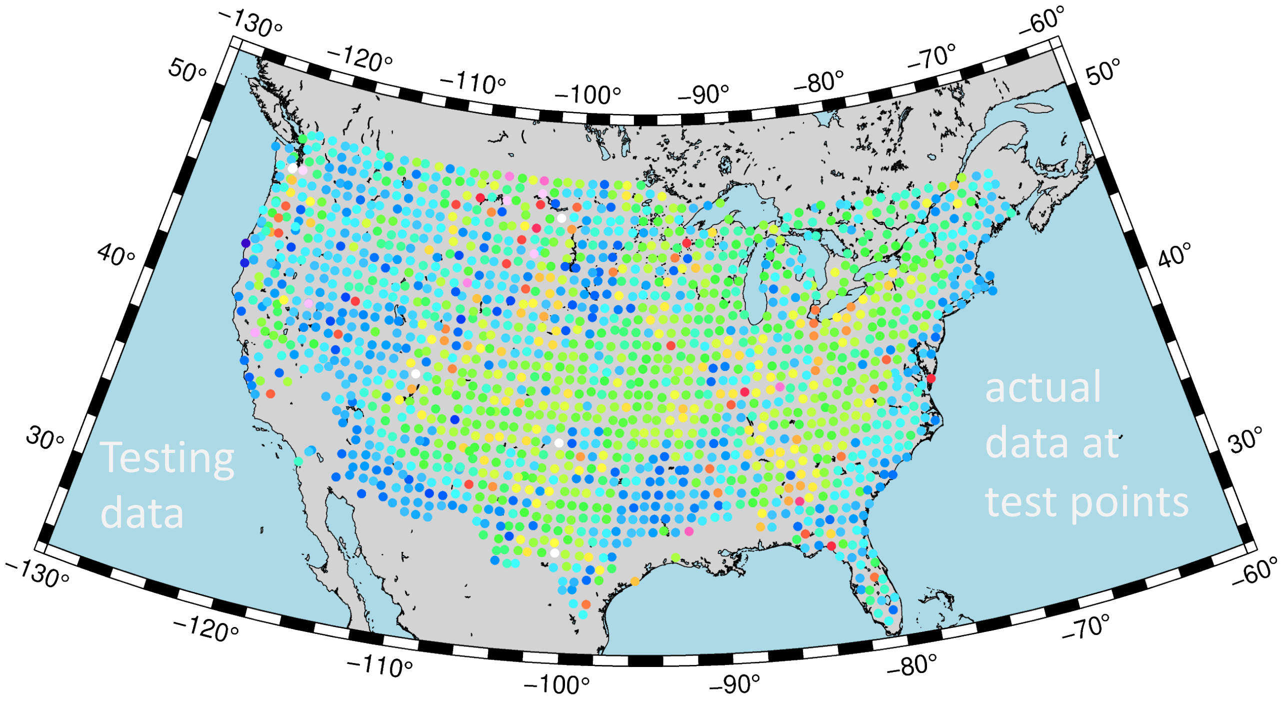




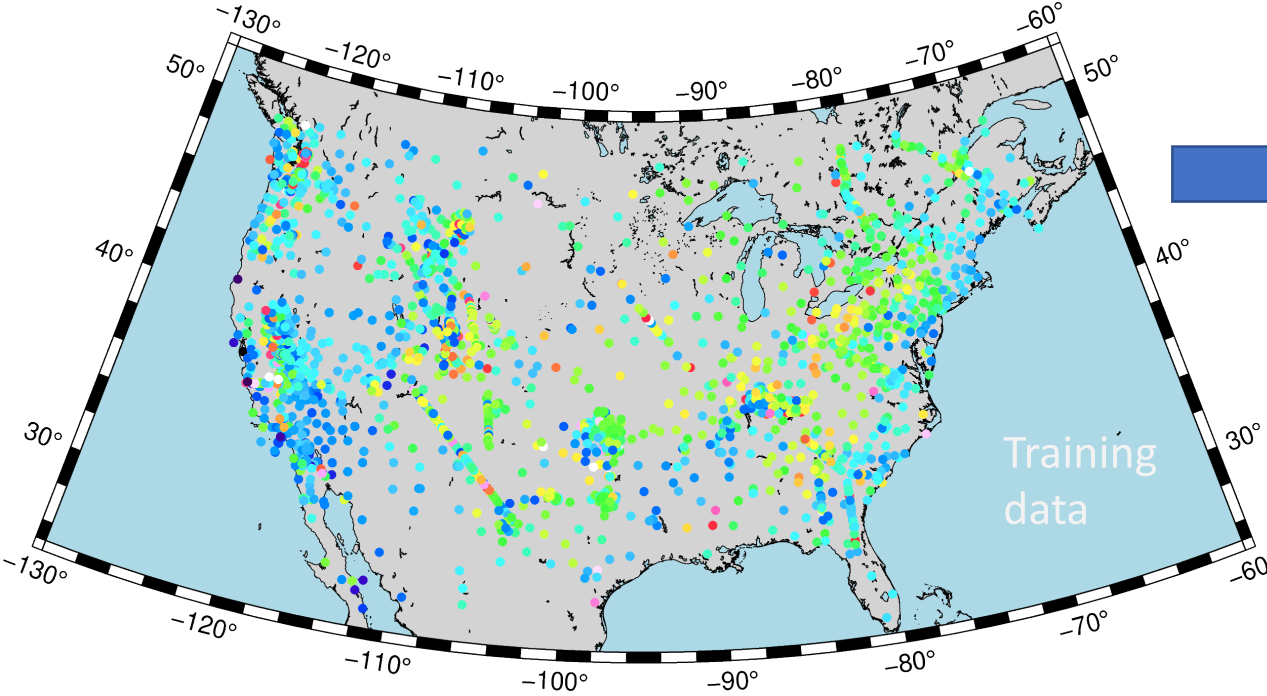
Next: Use Moho estimates from non-TA stations to predict Moho depth at TA stations

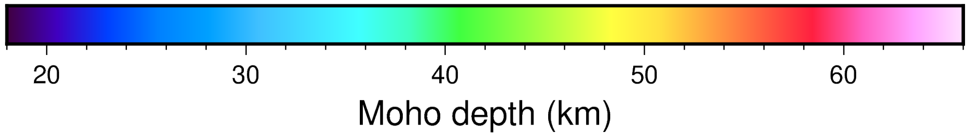

Training
data
20
30
40
50
60
Moho depth (km)

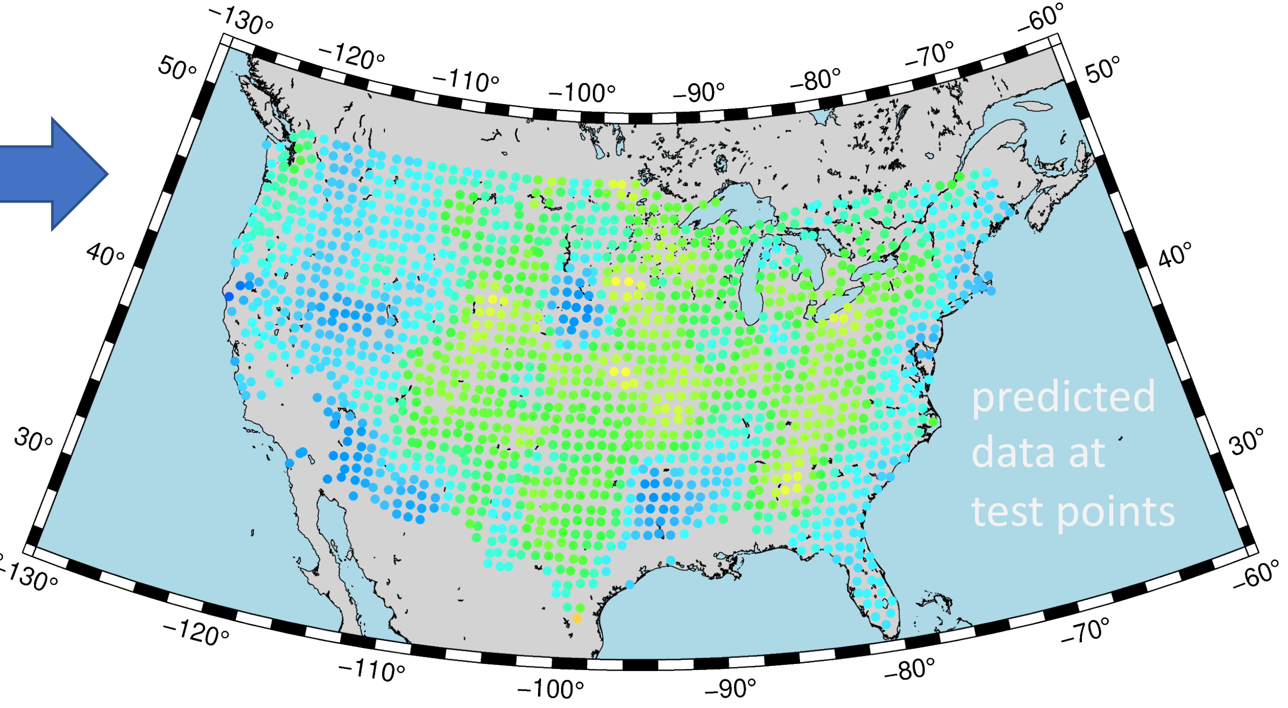

predicted
data at
test points

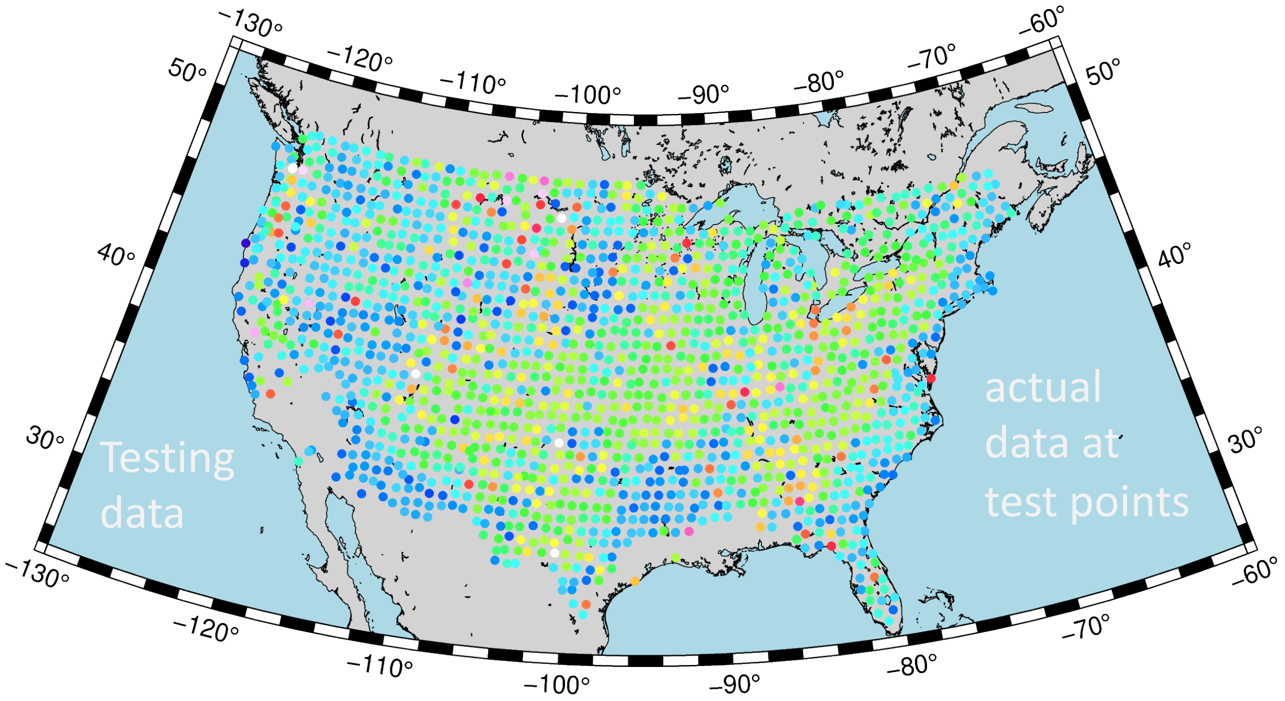

Testing
data
actual
data at
test points

Kriging

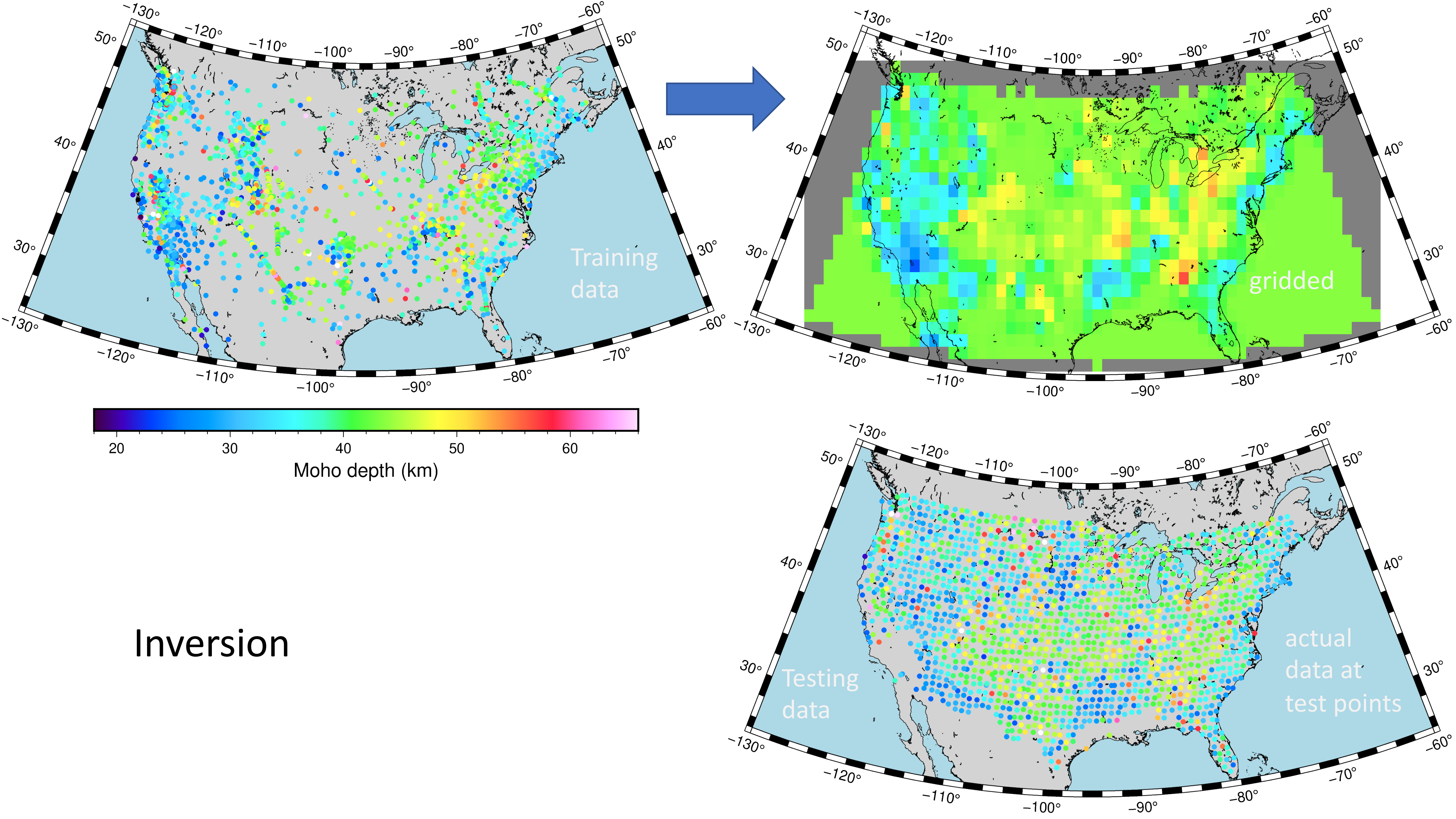


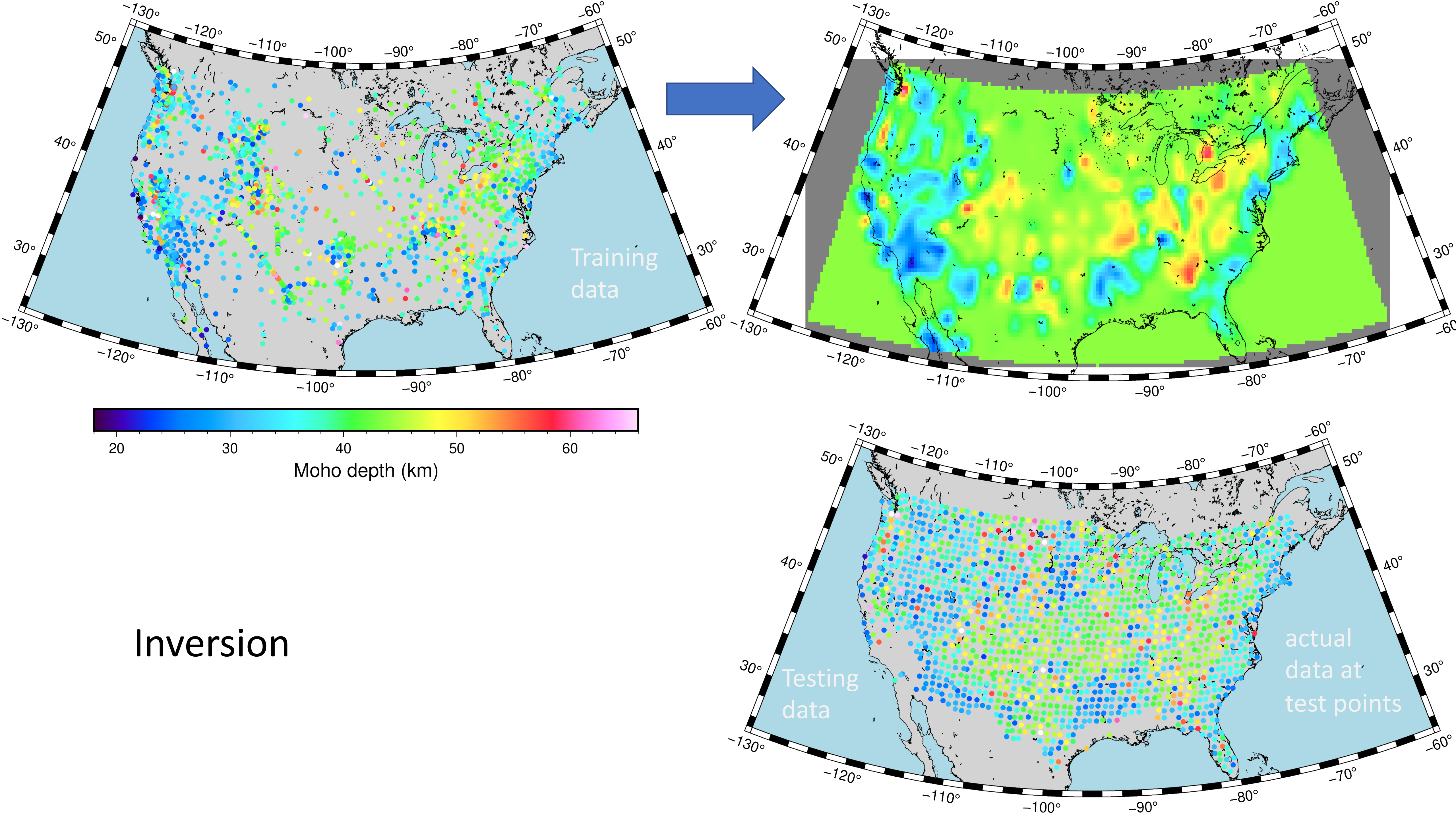






Weighted Nearest-Neighbor averaging


Training data
gridded
Testing data
actual data at test points
20
30
40
50
60
Moho depth (km)
Inversion


Training
data
20
30
40
50
60
Moho depth (km)
Inversion
Testing
data
actual
data at
test points

The Lab

1. Download EARS data

2. Split EARS data into estimates for USArray stations and non-TA stations

3. Explore different ways of using these data from non-TA stations to predict Moho depths at TA stations:
    a. Without gridding (weighted nearest neighbors, kriging)
    b. With gridding (using 1x1 grid and bilinear basis functions)

    Detail on point b (gridding):
    i. Define bilinear basis functions – example on 1x1 grid.
    ii. Use non-TA data on Moho depth and their error estimates, and prior expectations to infer optimal weights **w** (which themselves are Moho depths at the support points of the regular grid!) → this is called *inversion*
    iii. Use **w** to predict the Moho for TA stations
    iv. Calculate fit.
    v. More details in Lab Notebook

QUESTIONS?