



A Spatial-Autocorrelation Workflow for Data-Driven Confidence Mapping of Geological Surfaces

Patrizio Petricca*, Chiara D'Ambrogi

Dipartimento per il Servizio Geologico d'Italia, ISPRA, Rome

Introduction

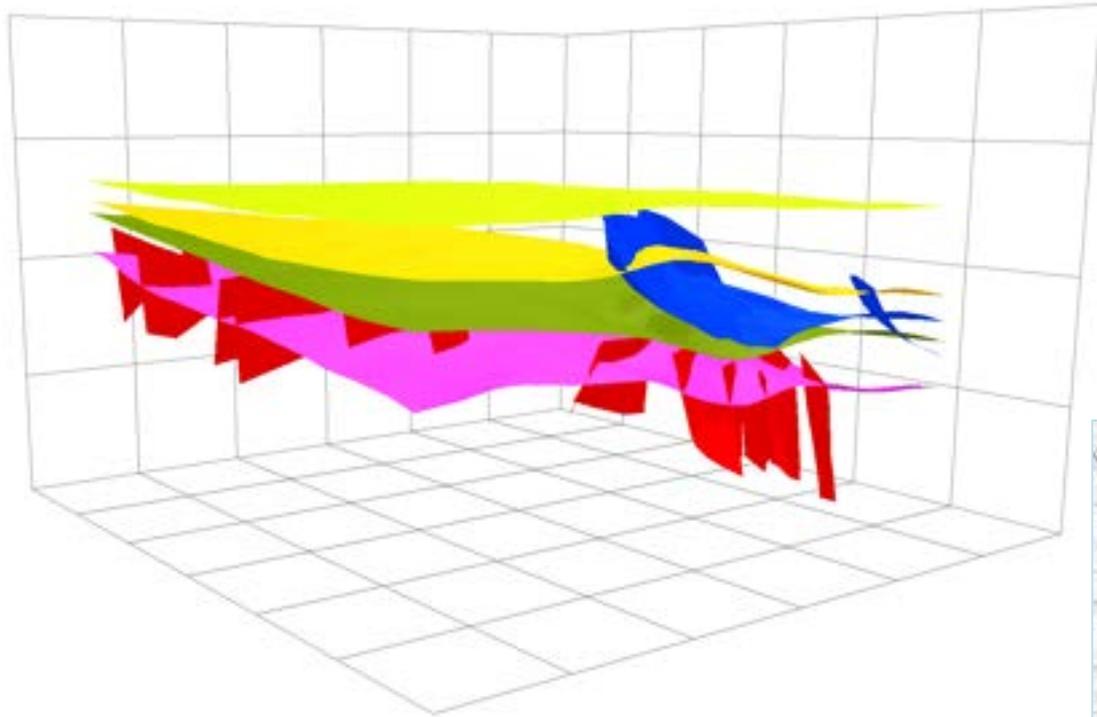
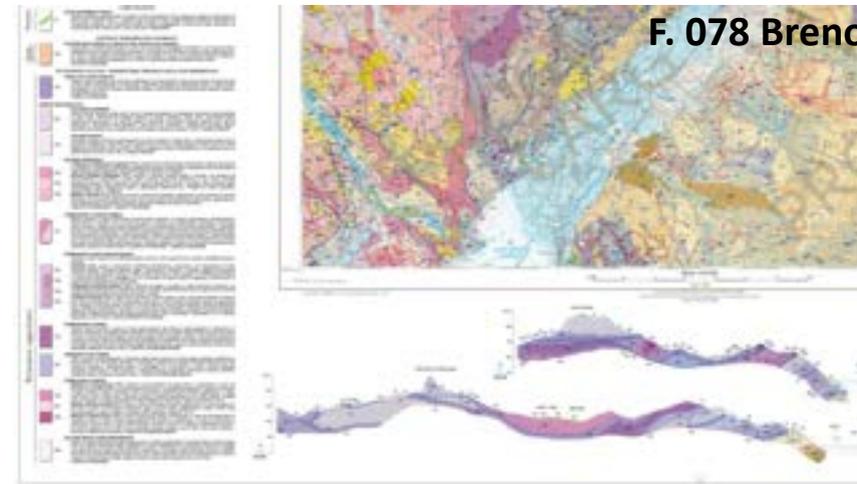


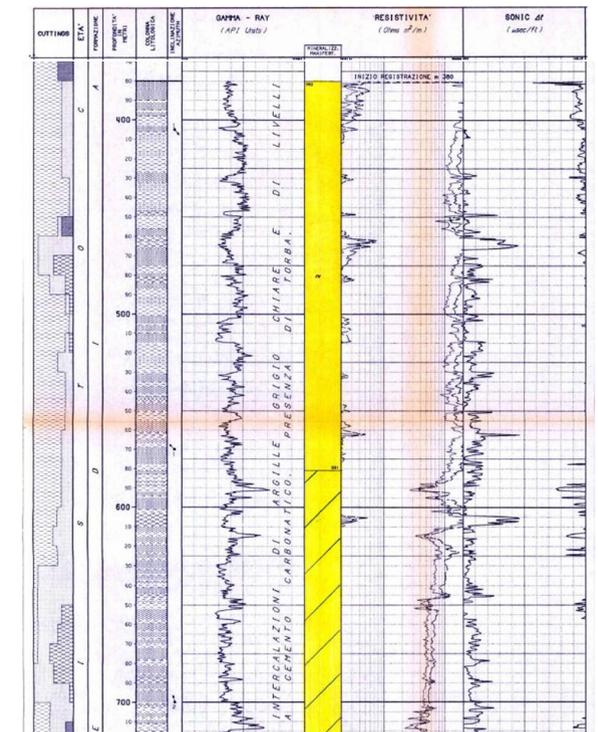
Image from **GEO-IT3D**
Geological 3D models in Italy

3D web-viewer

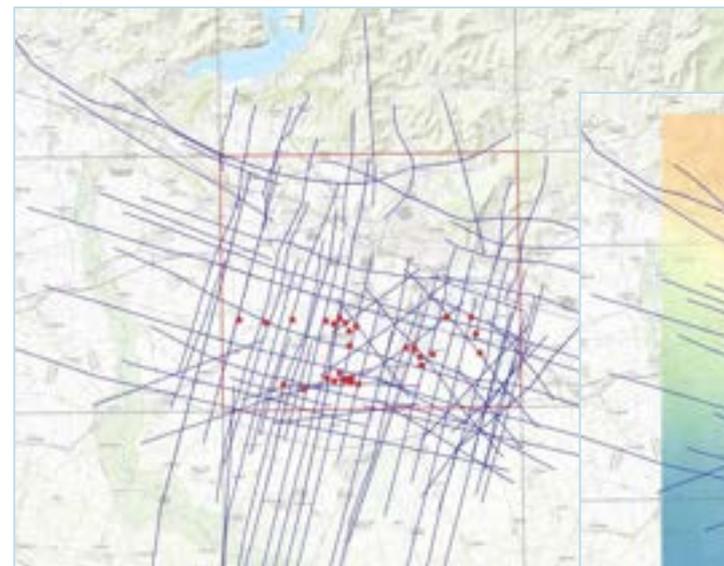
<https://geo-it3d.isprambiente.it/map>



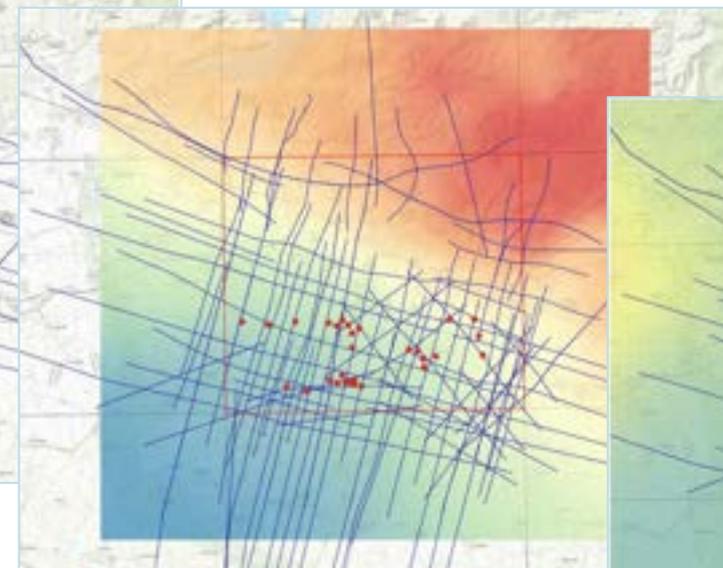
Geological survey



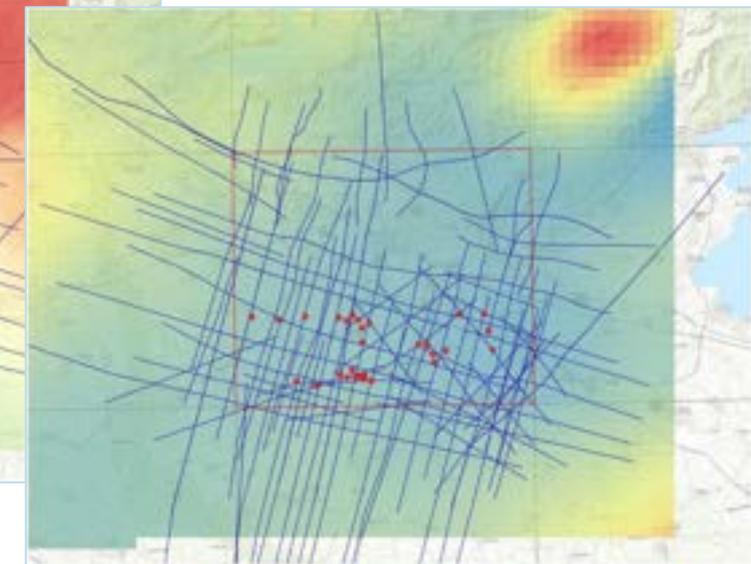
Composite log analysis



Seismic lines and boreholes



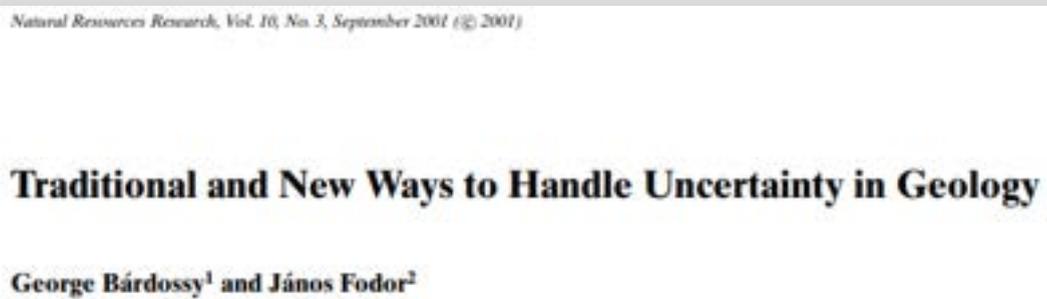
Gravimetric surveys



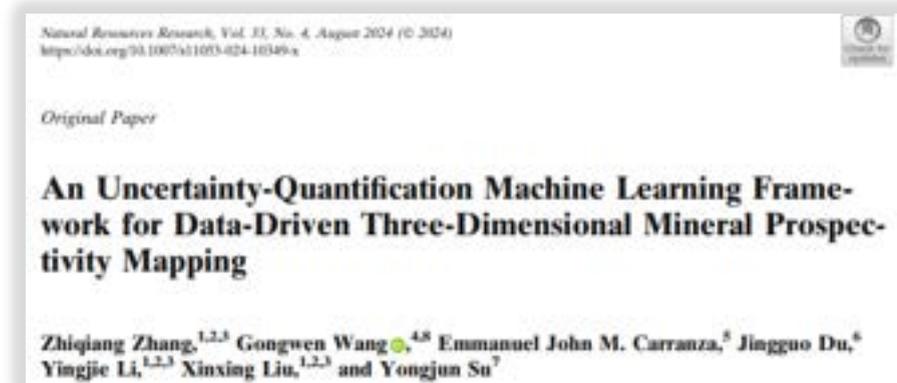
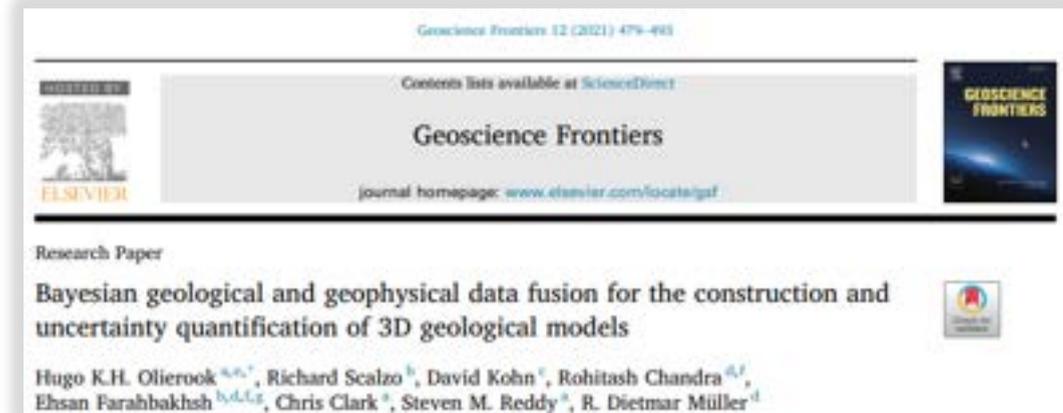
Aeromagnetic surveys

State of The Art

«**Quality flagging** is the basic approach used to quantify uncertainty within a spatial dataset and is done by assessing metadata fields; This approach may range from simply indicating data presence or absence to more complex methods yielding a comprehensive array of quantitative error ranges»



«**Probabilistic and Machine Learning models** are usually difficult to set up, computationally demanding as well as difficult to interpret for decision makers»

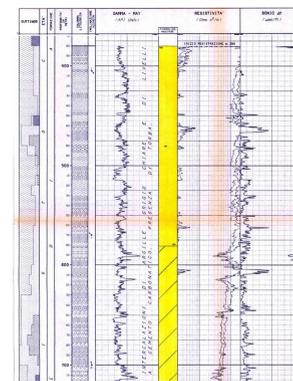
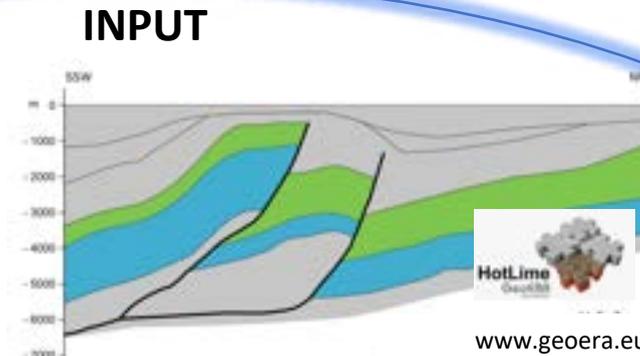
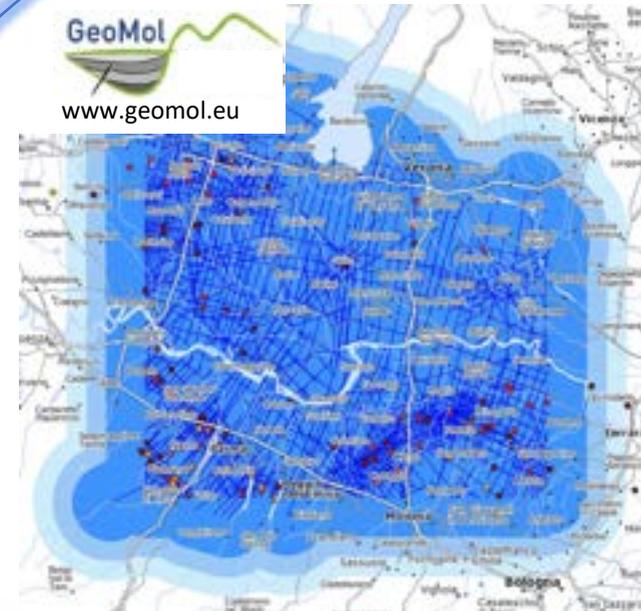


Limitations to a Generalized Approach for uncertainty

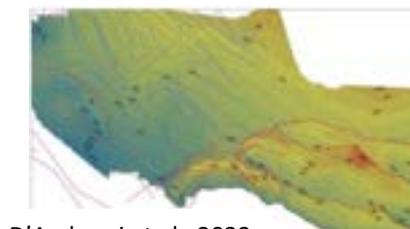
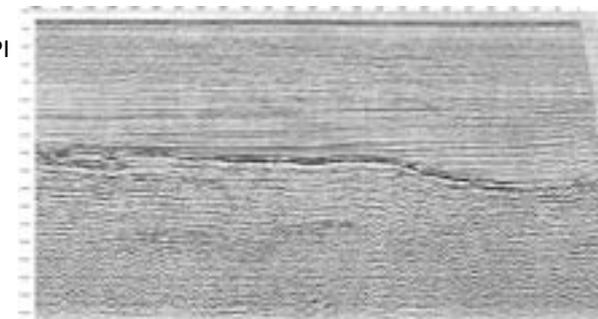
- i) Areas with few constraining data**
- ii) Data with no attached uncertainty**
- iii) Inhomogeneous data availability and distribution**
- iv) No information about data acquisition techniques and instrumental accuracies**
- v) Errors on data models (e.g. Digital Terrain and Time/depths conversion models)**
- vi) Action Errors, Selection Errors, Information Retrieval Errors, Checking Errors (=Human errors)**

Method

What kind?
Types of Spatial Data
- punctual
- linear
- polygonal



ViDEPI



D'Ambrogi et al., 2022

How many?
Count of Spatial Data
- n° punctual
- n° linear
- n° polygonal

How are they distributed?

Spatial autocorrelation measures the correlation between the values of a geographical variable as a function of the distance from the sampling points. **Key concept:** geographically close points tend to have similar values (positive autocorrelation).

Workflow

Pre-processing

Select the surface

outline the study area

transform it into a regular grid

tag the observations

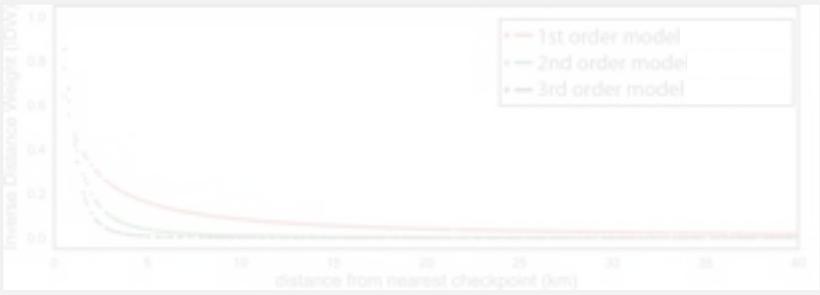
checkpoints order
 1st wells/outcrops ●
 2nd seismic lines/geologic sections
 3rd isobaths

r^p = distance between the point of observation and the checkpoint of order p

$$IDW = \frac{1}{1 + r^p}$$

$$IDW_i = \frac{ID - ID_{min}}{ID_{MAX} - ID_{min}}$$

$$IDW_{avg} = \frac{1}{n} \sum_{i=1}^p IDW_i$$

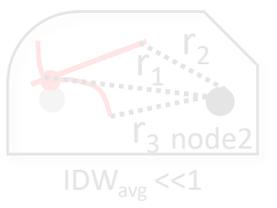
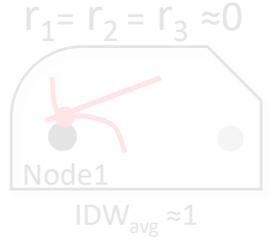


evaluate the horizontal confidence

evaluate the vertical confidence

interpolator (linear/NN)

$$\Delta_{norm} = 1 - \frac{\Delta - \Delta_{min}}{\Delta_{MAX} - \Delta_{min}}$$



collect matrices of confidence values

plot

Workflow

Select the surface

outline the study area

transform it into a regular grid

tag the observations

checkpoints order
 1st wells/outcrops ●
 2nd seismic lines/geologic sections ~
 3rd isobaths ~

Pre-processing

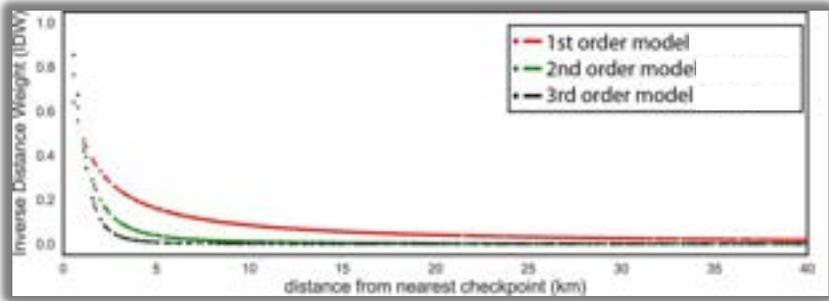
model

r^p = distance between the point of observation and the checkpoint of order p

$$IDW = \frac{1}{1 + r^p}$$

$$IDW_i = \frac{ID - ID_{min}}{ID_{MAX} - ID_{min}}$$

$$IDW_{avg} = \frac{1}{n} \sum_{i=1}^p IDW_i$$

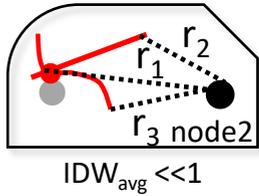
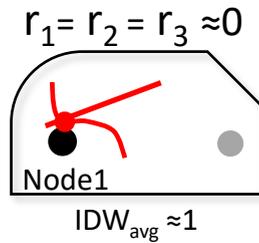


evaluate the horizontal confidence

evaluate the vertical confidence

interpolator (linear/NN)

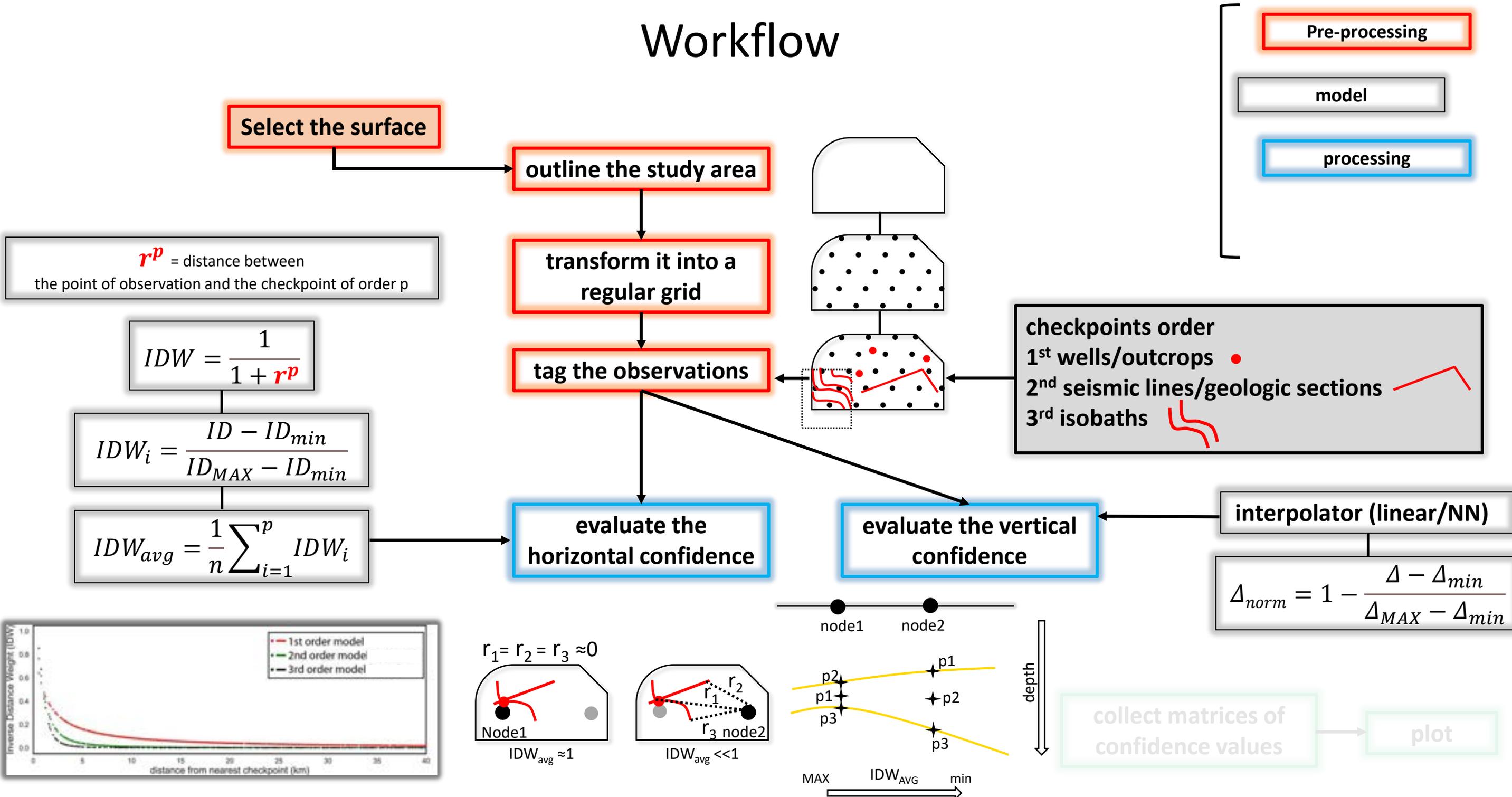
$$\Delta_{norm} = 1 - \frac{\Delta - \Delta_{min}}{\Delta_{MAX} - \Delta_{min}}$$



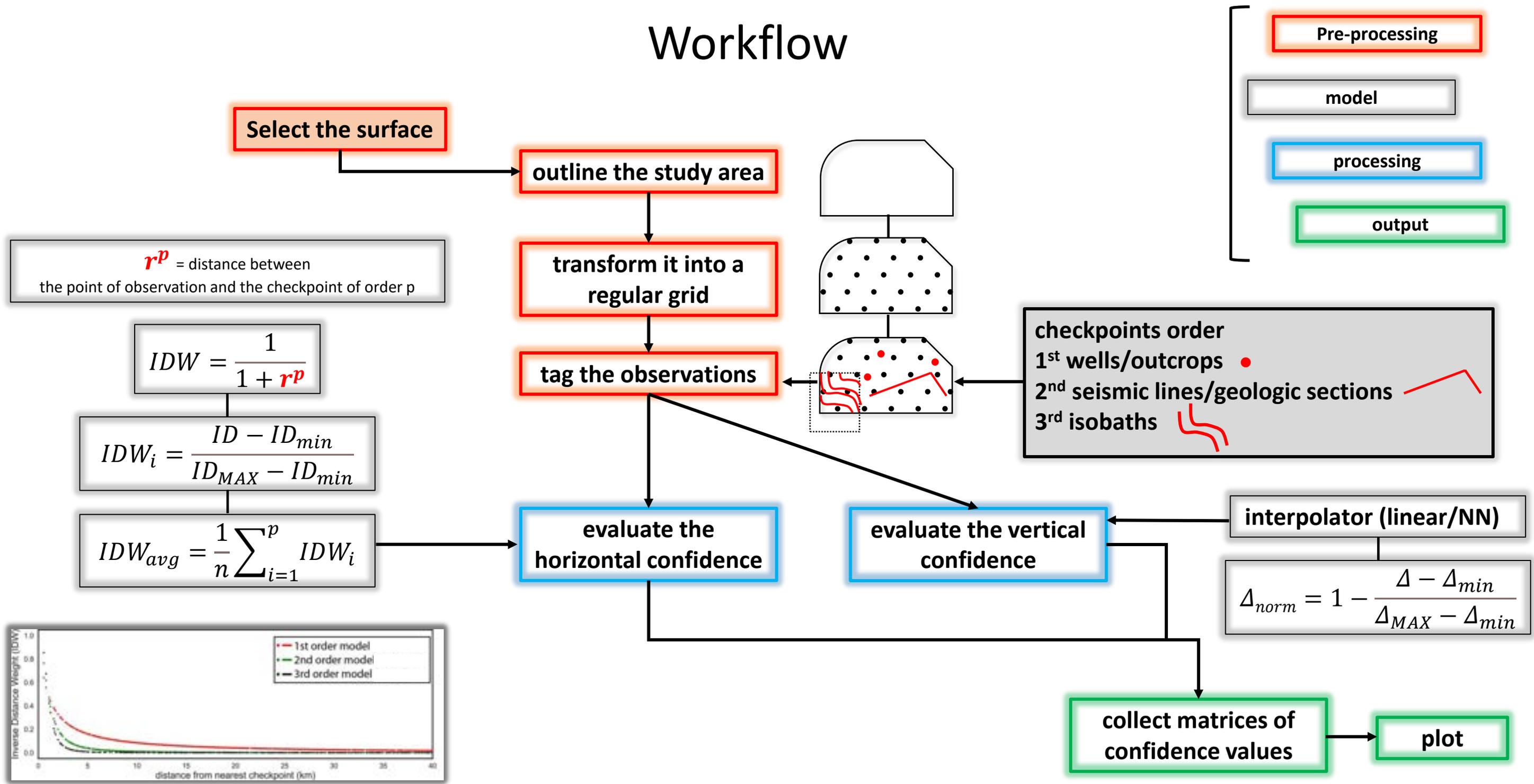
collect matrices of confidence values

plot

Workflow

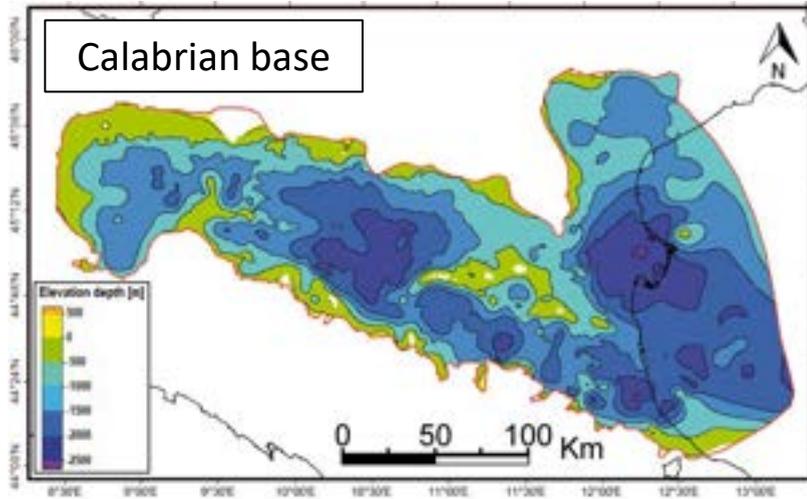


Workflow

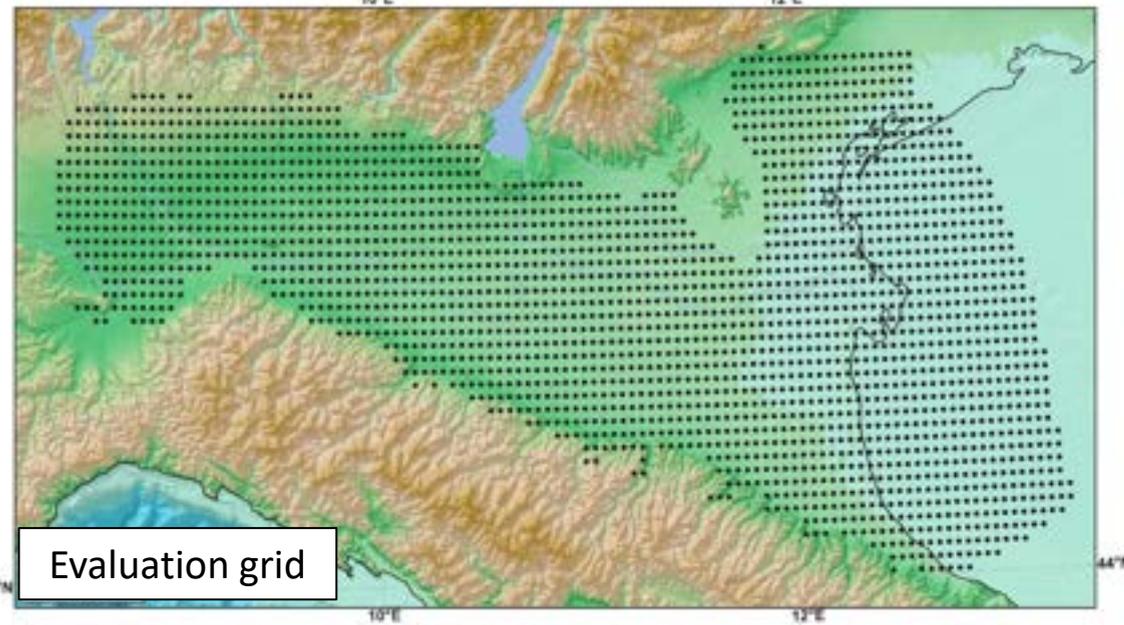


Example – evaluation grid

Pre-processing

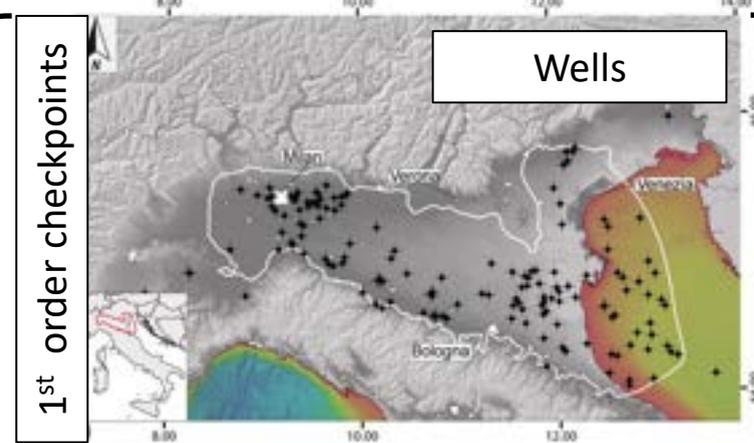


Livani et al., 2023

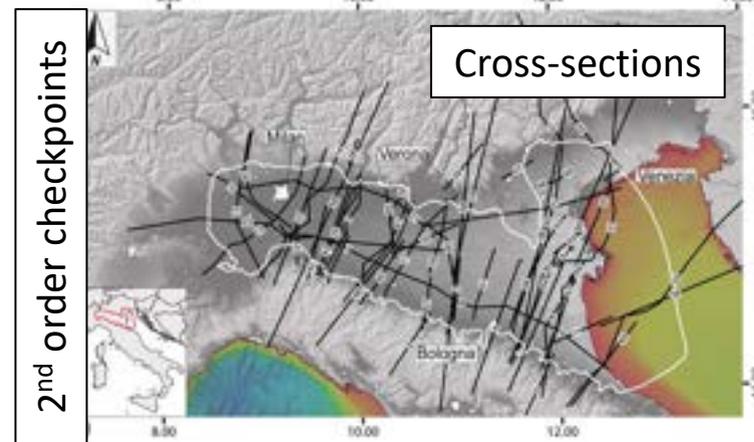


Evaluation grid

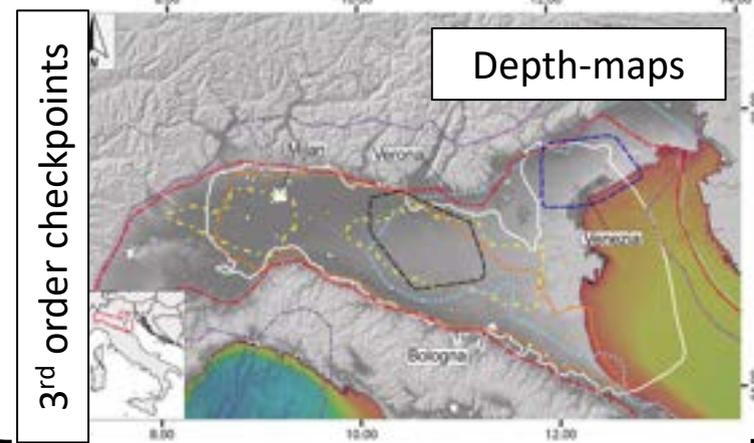
n. Checkpoint (p1)	n. Checkpoint (p2)	n. Checkpoint (p3)	Checkpoint density
139	3065	2299	2.36 per node



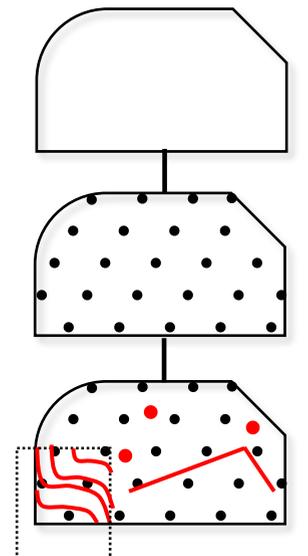
1st order checkpoints



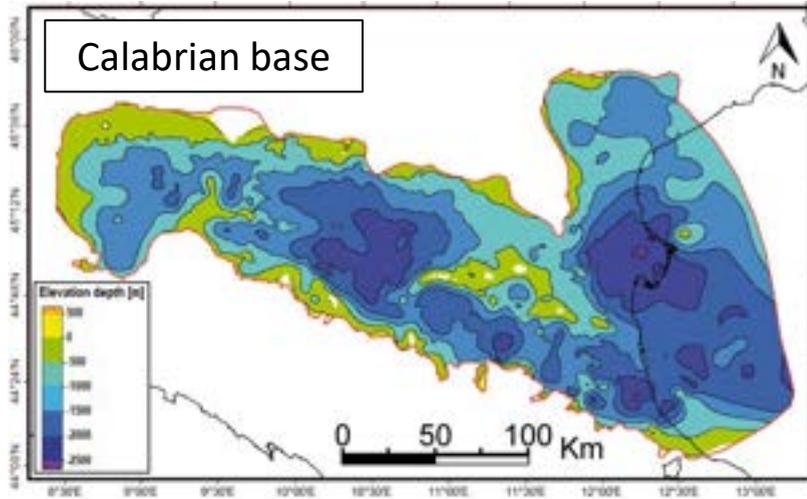
2nd order checkpoints



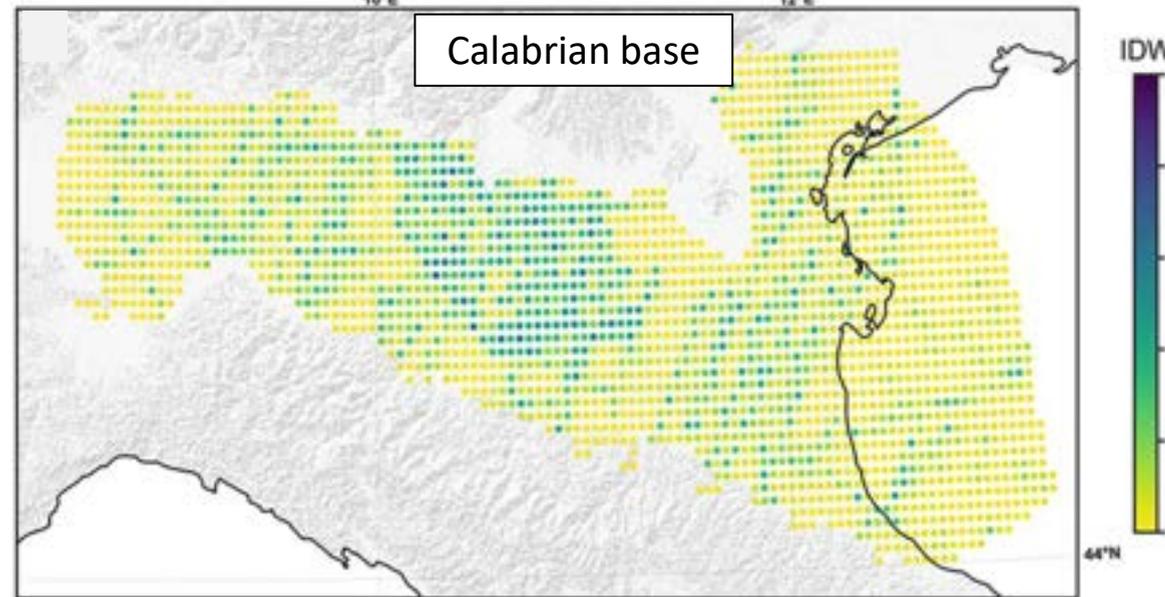
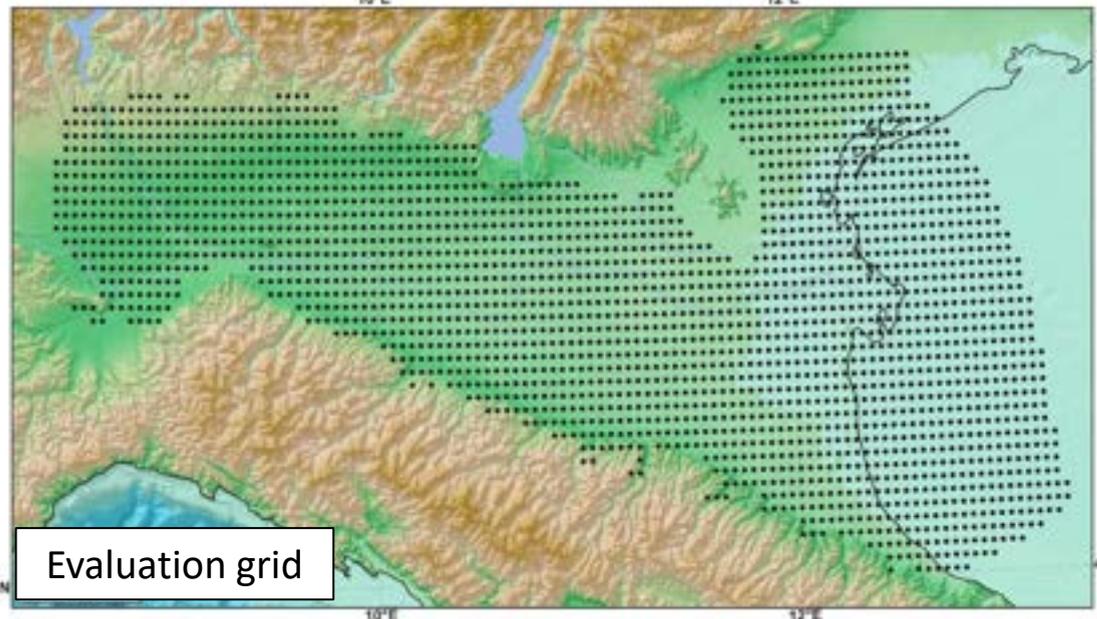
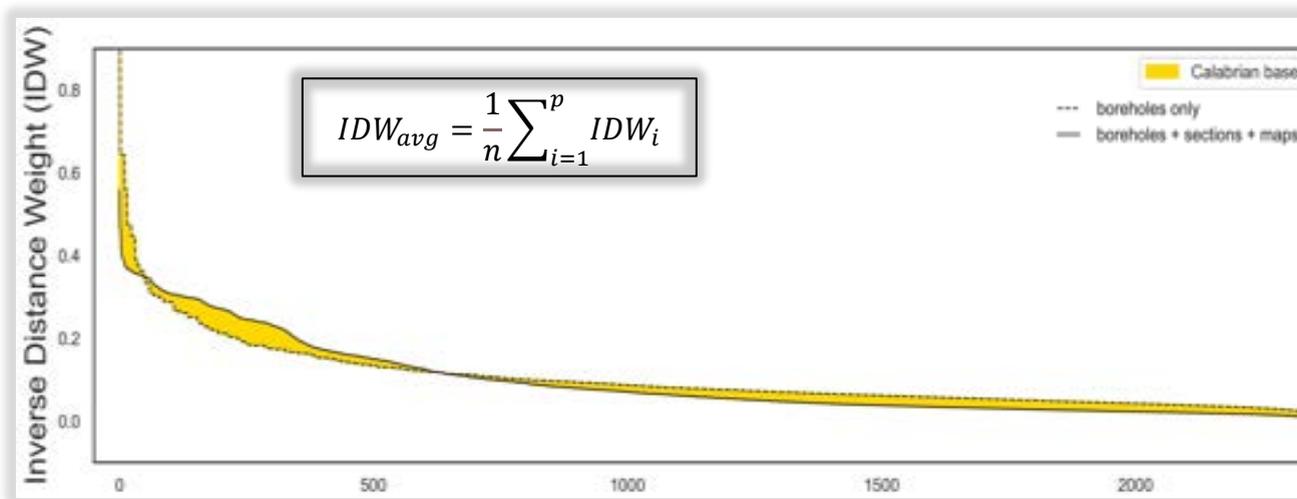
3rd order checkpoints



Example – horizontal accuracy



Livani et al., 2023

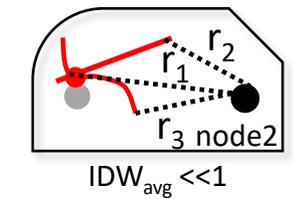
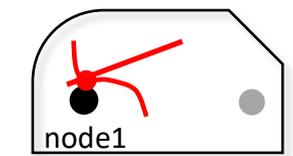


Pre-processing

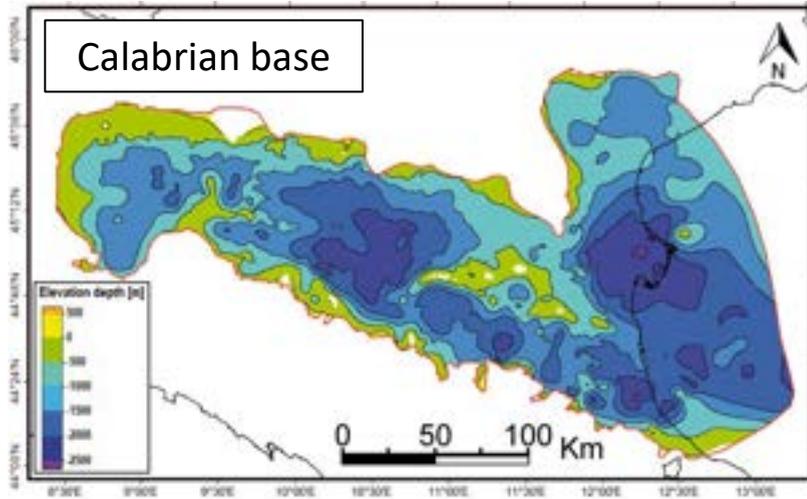
model

processing

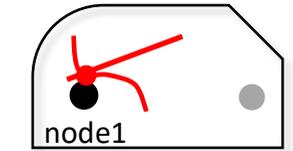
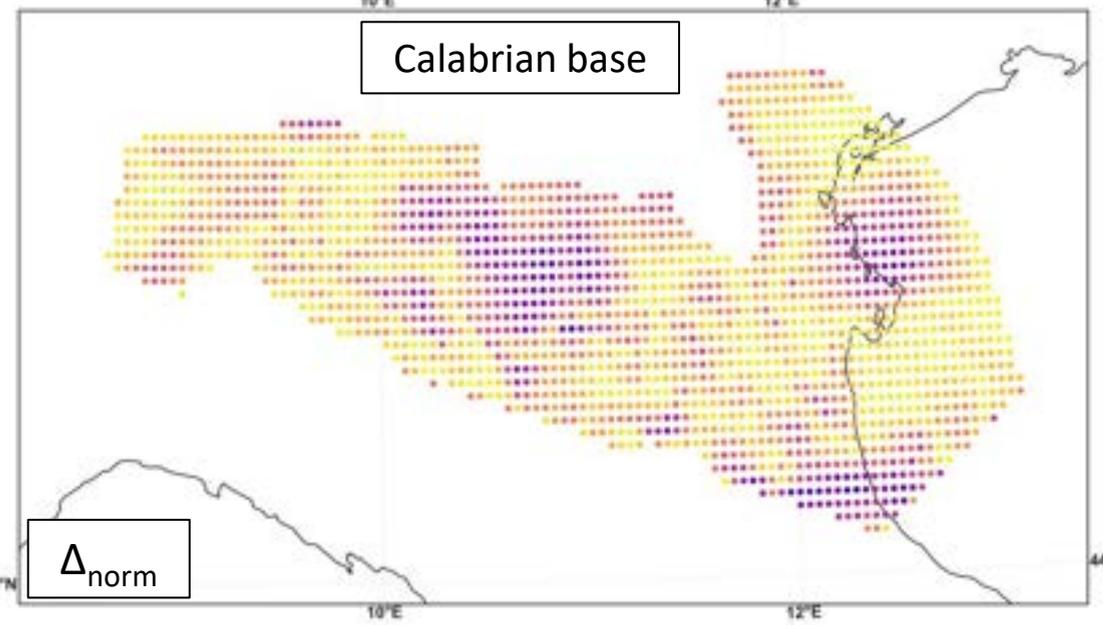
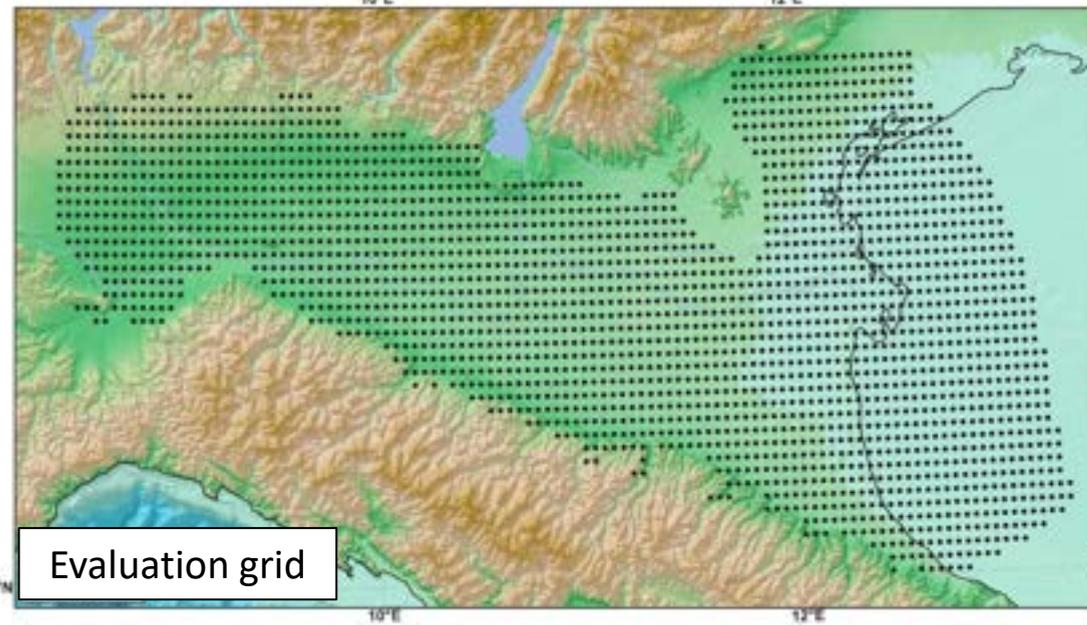
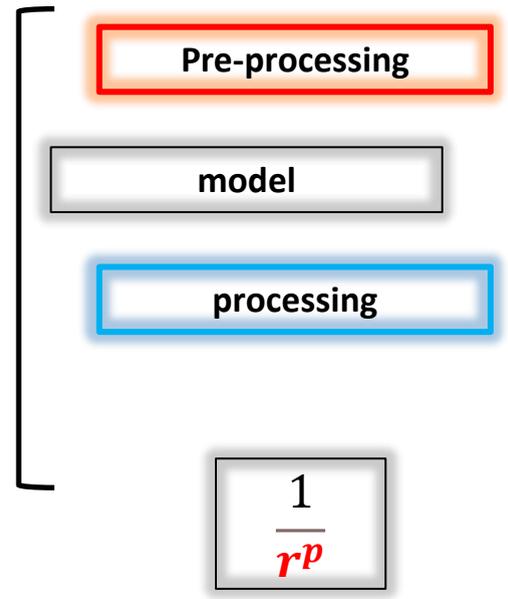
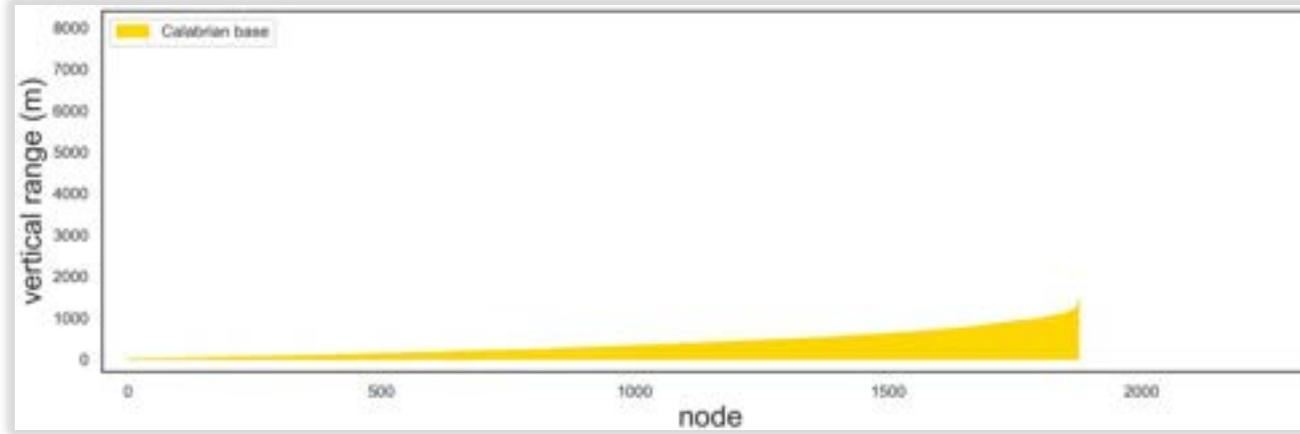
$$\frac{1}{r^p}$$



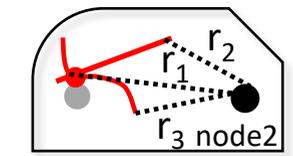
Example – vertical range of variation



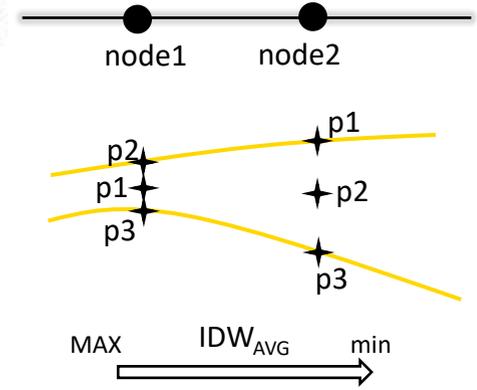
Livani et al., 2023



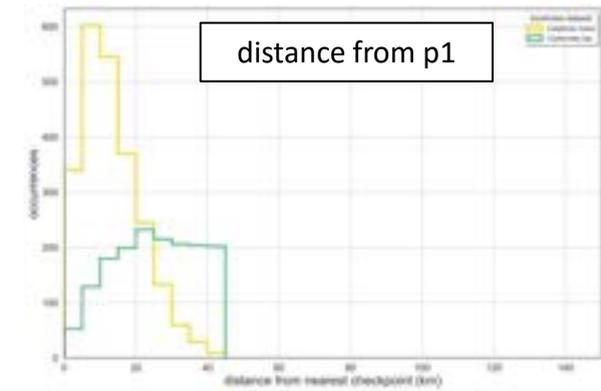
IDW_{avg} ≈ 1



IDW_{avg} << 1

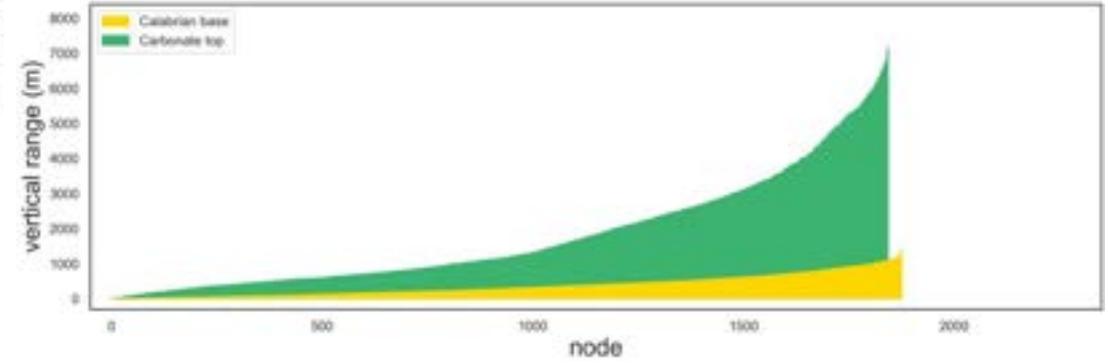
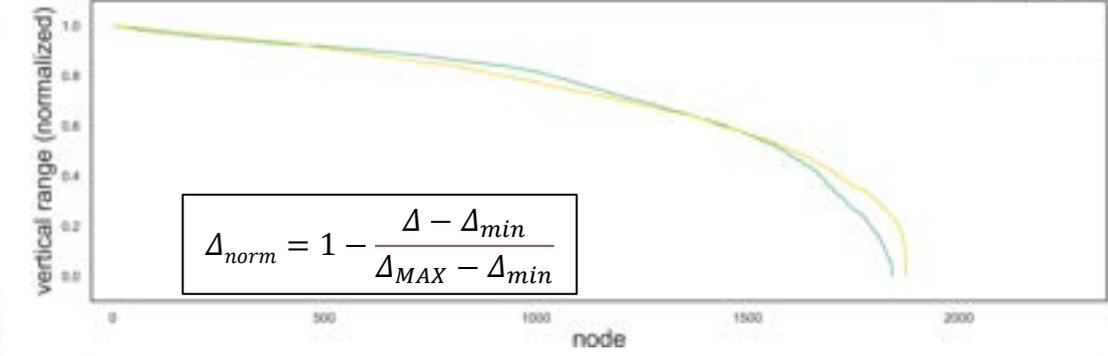
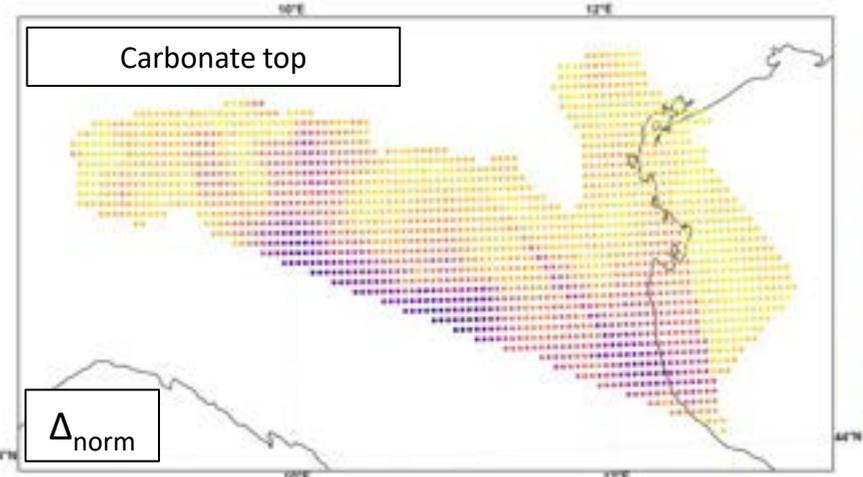
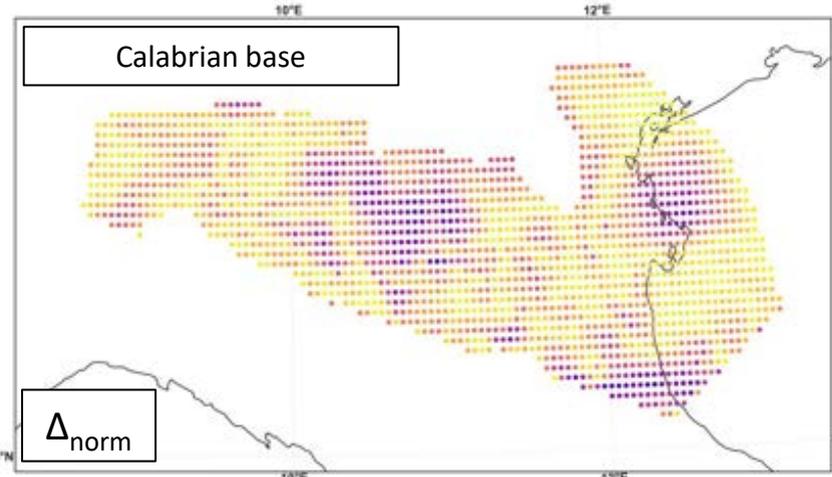
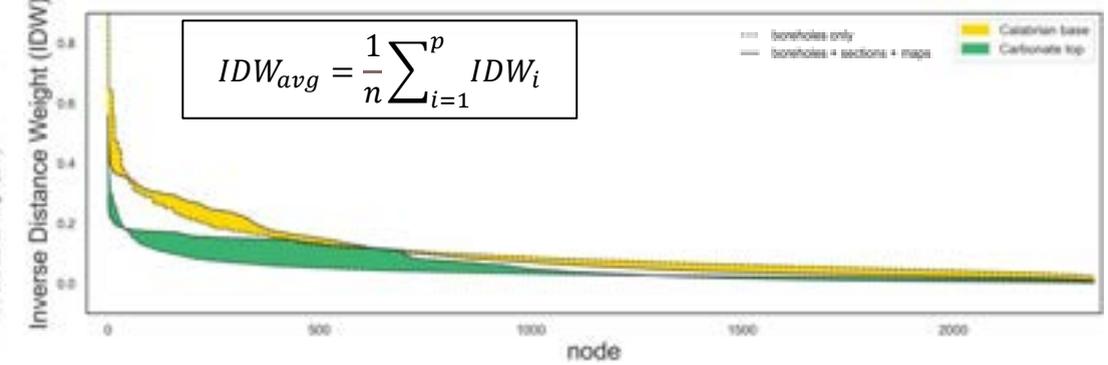
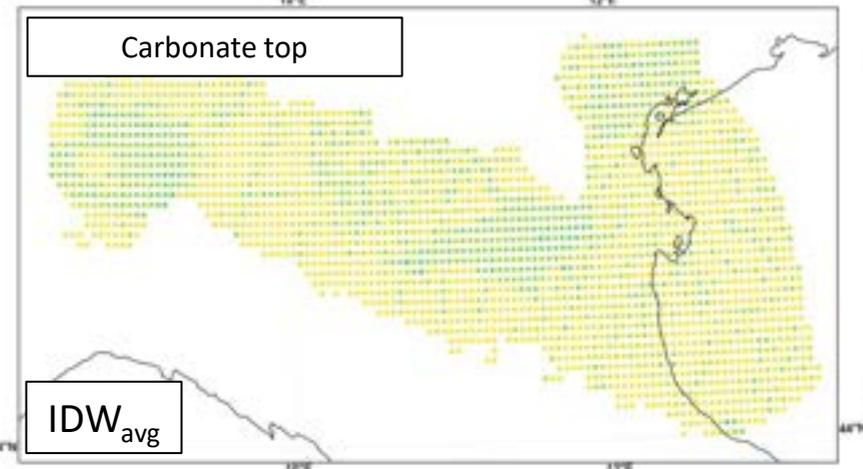
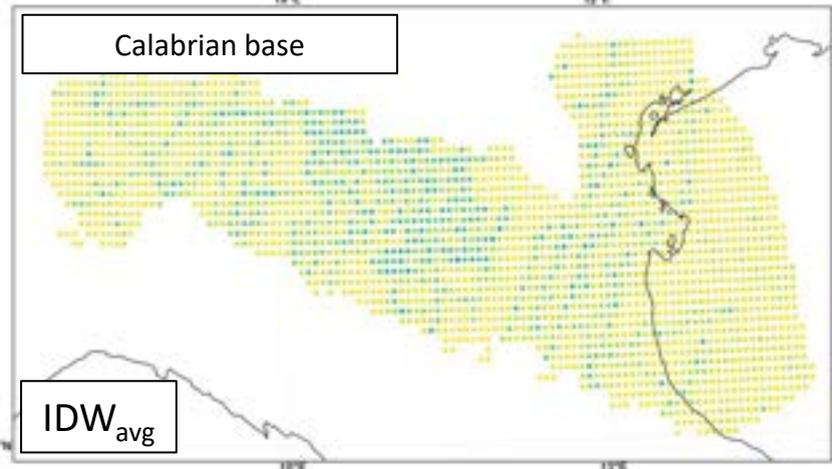


Example – comparing surfaces



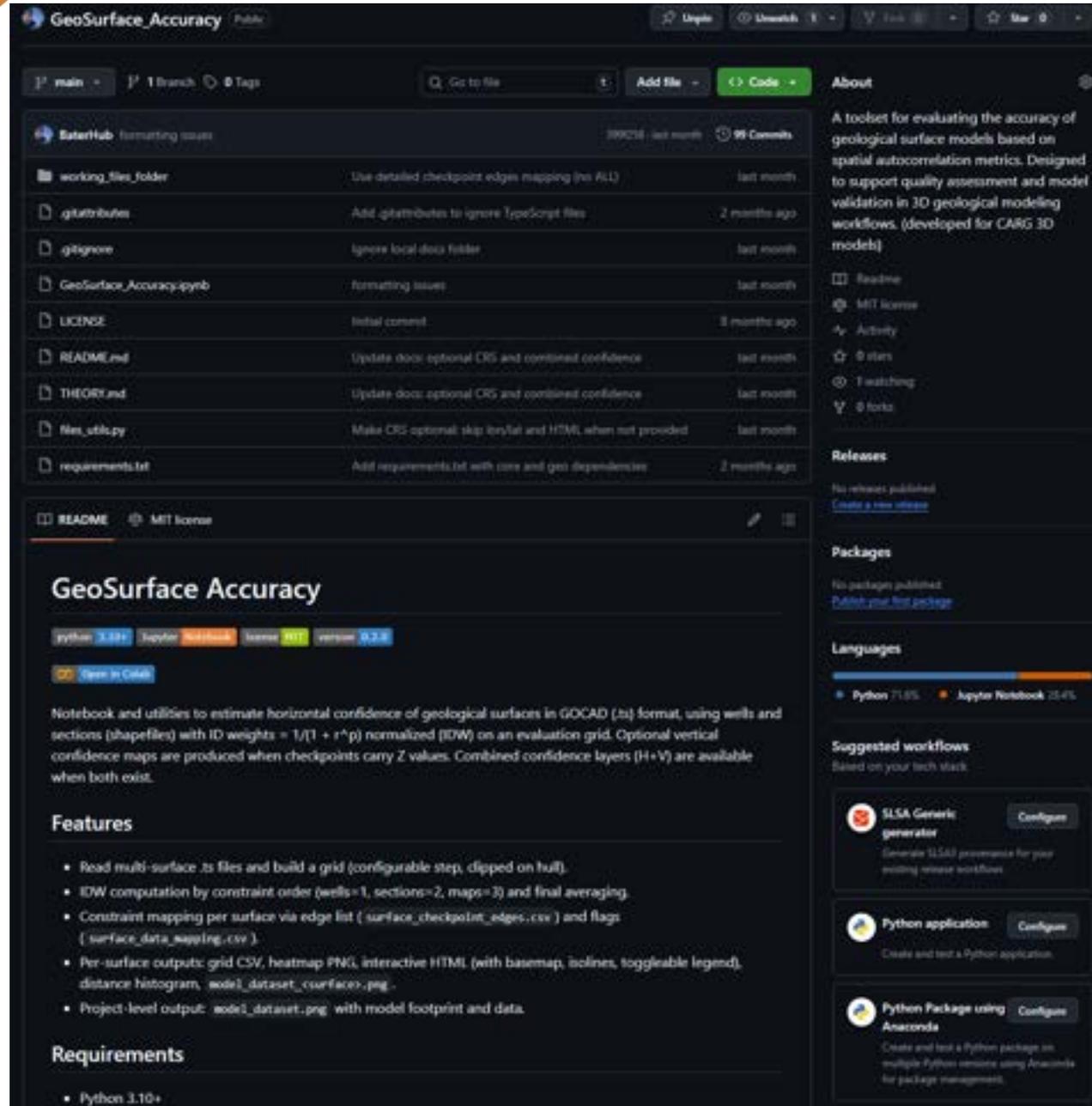
p1 checkpoints =139

p1 checkpoints =17

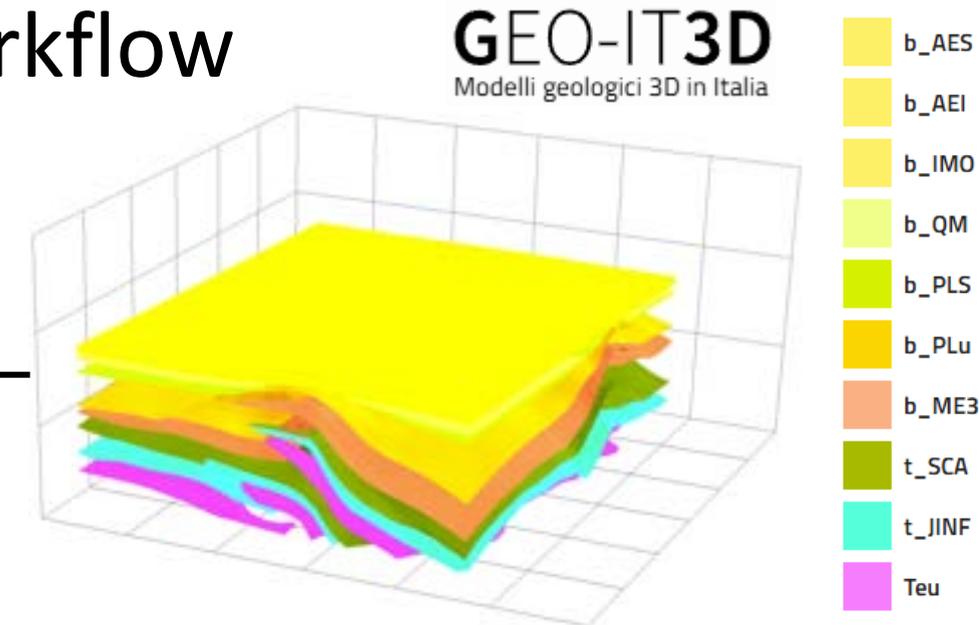


Livani et al., 2023

From concept to reproducible workflow



https://github.com/BaterHub/GeoSurface_Accuracy



<https://geo-it3d.isprambiente.it/>

INPUT FILES format:

- GOCAD .ts (structured surfaces)
- ESRI Shapefile .shp
- Tabular data .csv, .txt

OUTPUT FILES:

- confidence maps .png
- numerical outcomes .csv
- statistics .csv

-> Open and reproducible Python notebook
 -> Based on interoperable formats
 -> Horizontal component fully implemented
 -> Vertical component under development

Conclusions

- i) **Transparent and reproducible confidence mapping**
- ii) **Spatially explicit, data-driven approach**
- iii) **Designed for heterogeneous Geological Survey datasets**

Open points

- i) **Sensitivity analysis (grid resolution/search radius)**
- ii) **Systematic analysis of several case studies**
- iii) **Comparison with alternative spatial models**
- iv) **Integration of horizontal and vertical confidence**

Thanks for listening

https://github.com/BaterHub/GeoSurface_Accuracy

Advancing probabilistic uncertainty- incorporated 3D voxel modelling – the application of uncertainty

Thomas Højland Lorentzen

Anne-Sophie Høyer

Rasmus Bødker Madsen (Presenter)

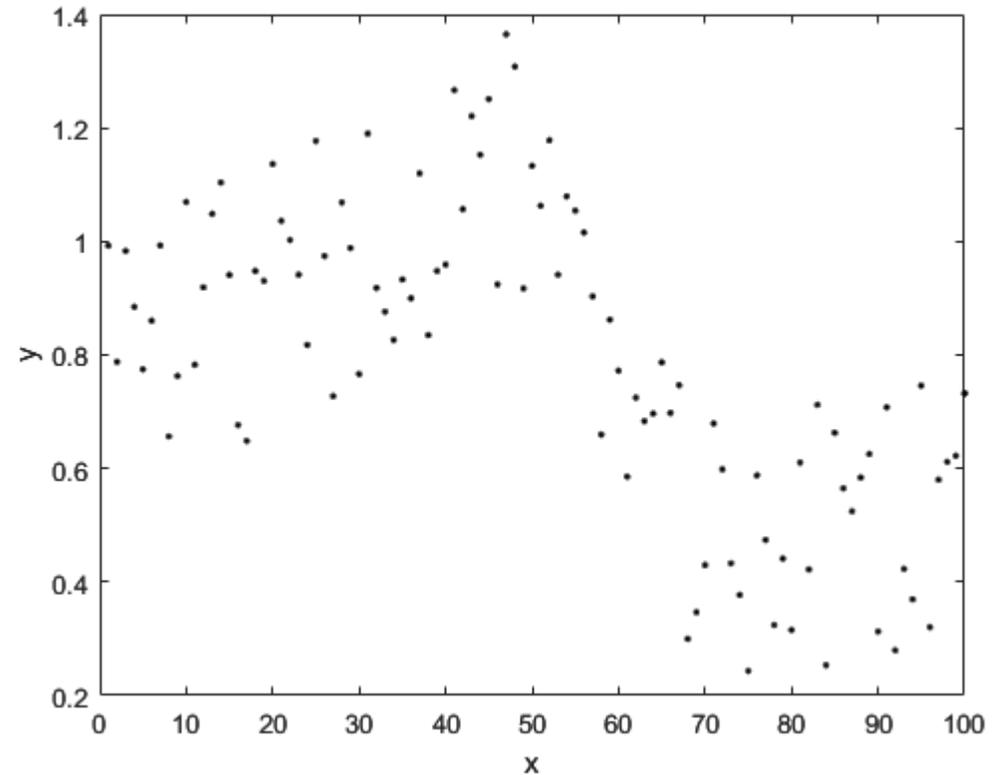


Data Don't Tell the Whole Story

Local measurements only

No context between observations

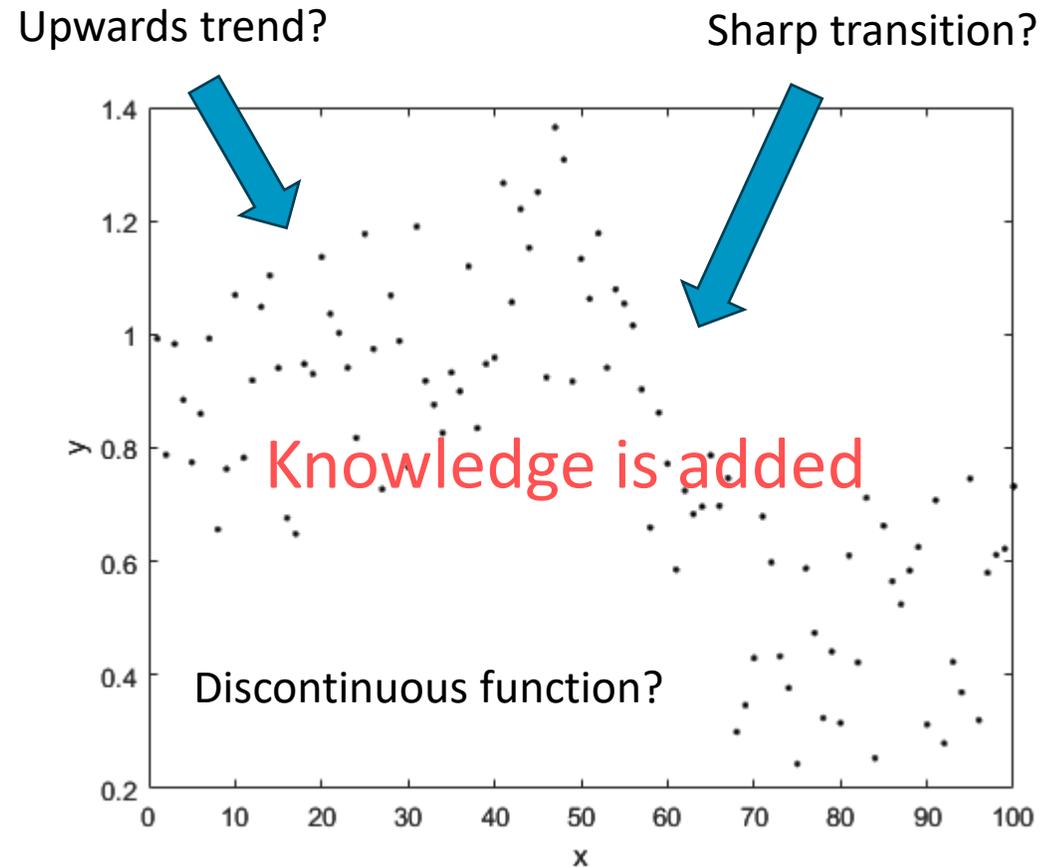
No structure



Models Give Context to Data

Connects data
Interpret structure
Generalize beyond
measurements

Model = data + interpretation +
assumption



Models Give Context to Data

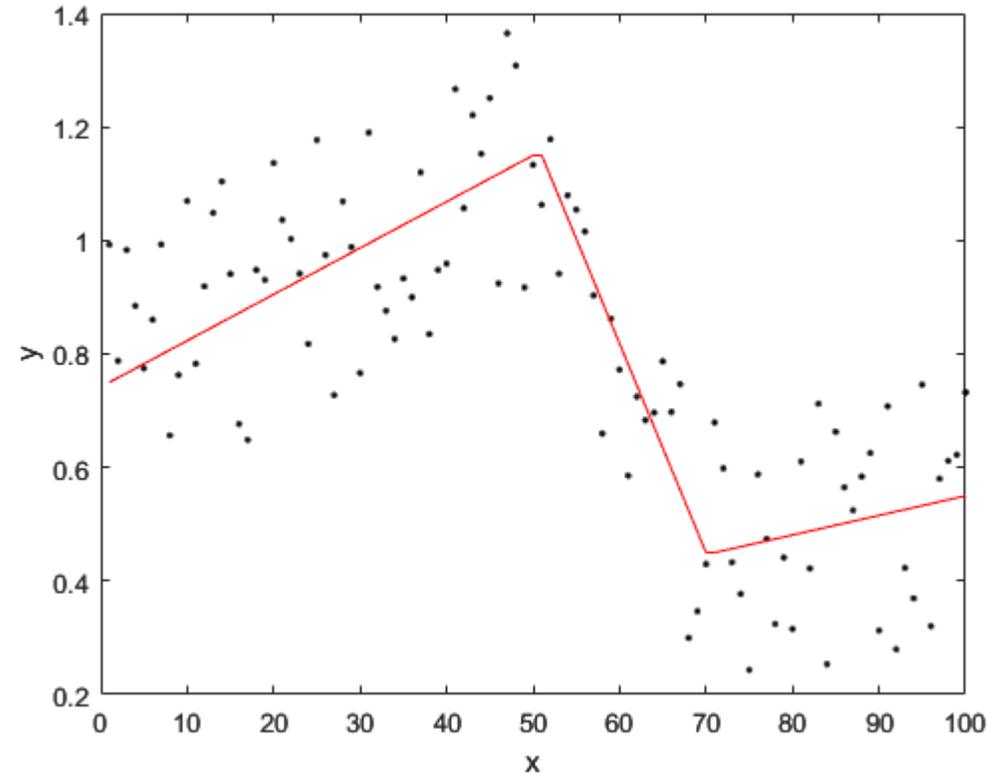
Connects data

Interpret structure

Generalize beyond measurements

Model = data + interpretation + assumption

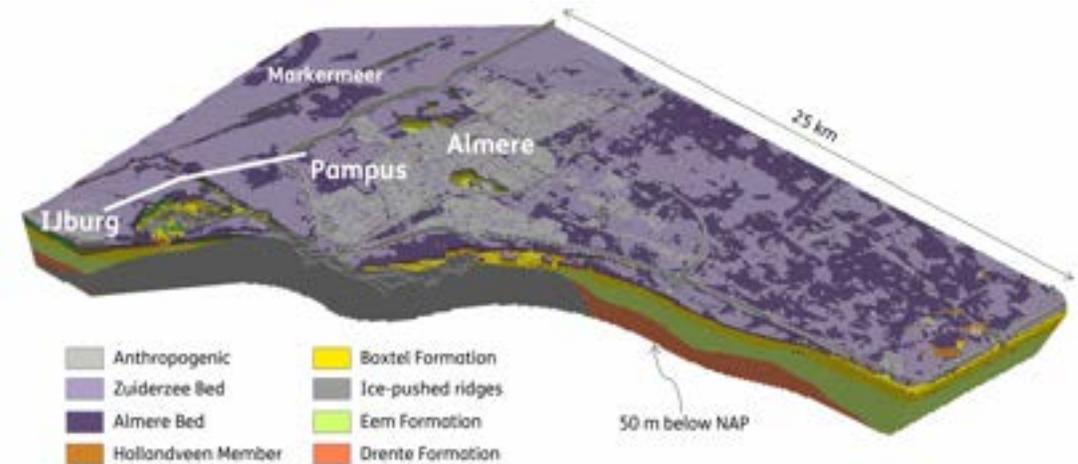
“Model representation” to convey information



Model representations are key for communication! – Especially in geoscience

Example: GEOTOP

- Netherlands
- National-scale voxel model
- Multiple applications
- Working tool
- Communication with stakeholders and end-users
 - Knowledge ingrained (“self explanatory”)



<https://www.dinoloket.nl/en/nieuws/geotop-extended-to-almere>

Motivation: National voxel model of Denmark

Today: Hydrostratigraphic model

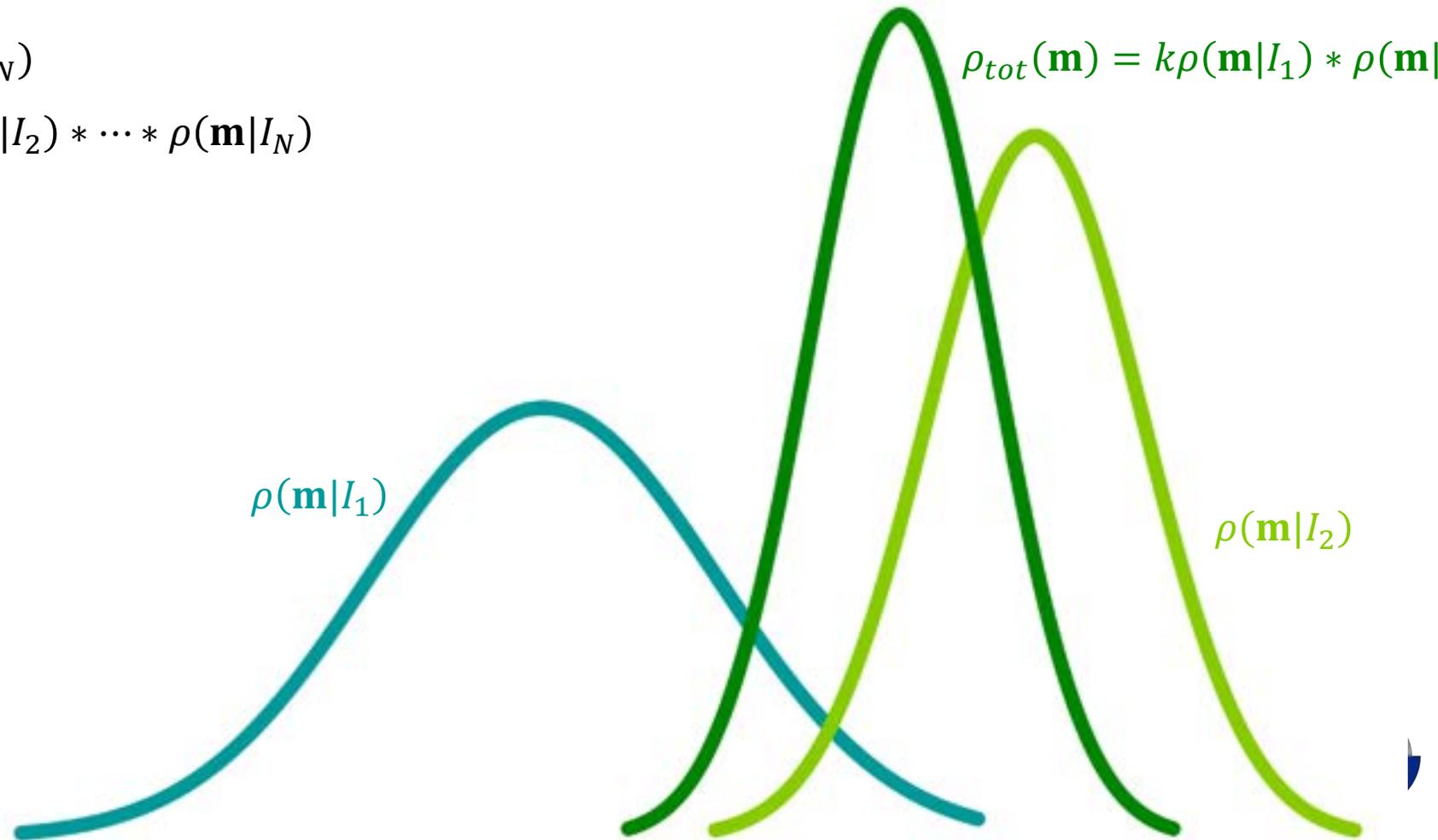
- Variable quality
 - Lack of near surface resolution
 - Variable modeling strategies
 - Variable data availability
- Not fit for multipurpose use
- Unused subsurface data

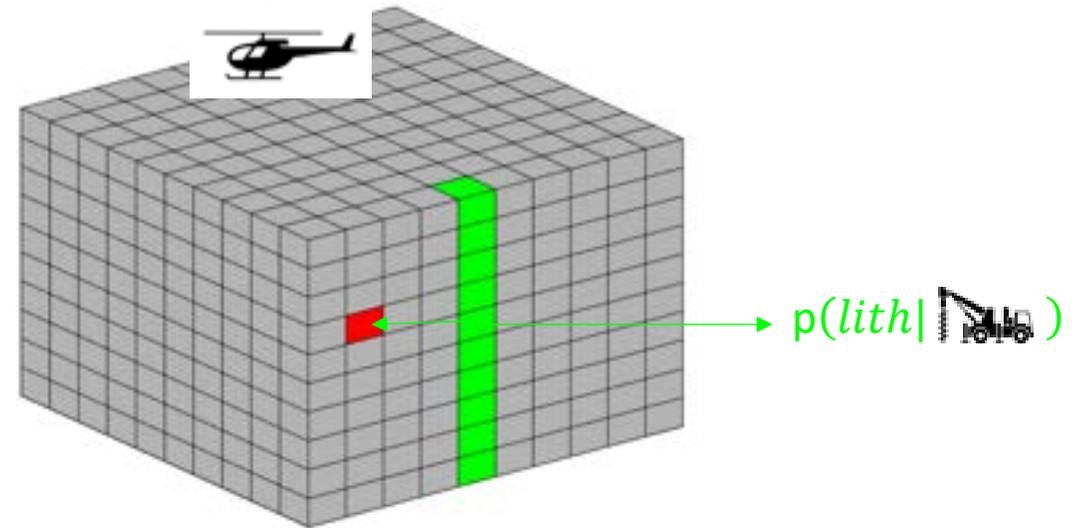
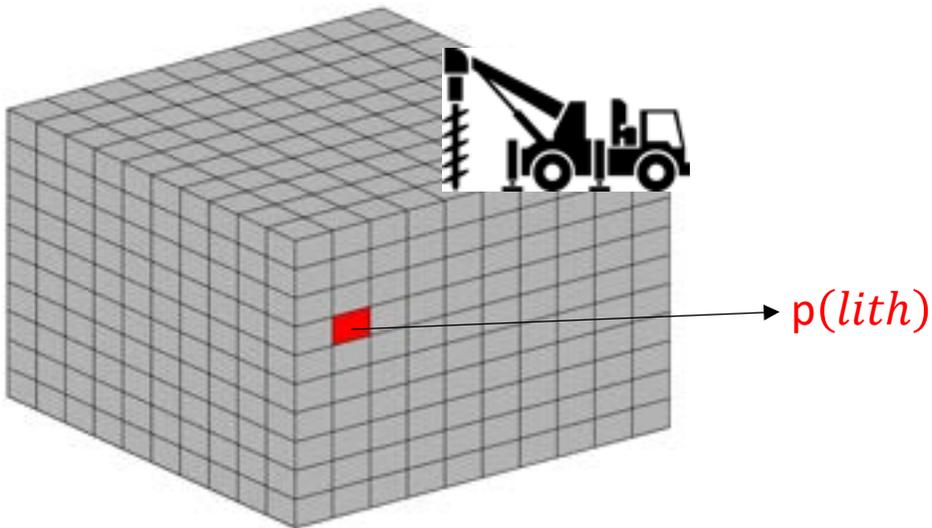


The answer: Probabilistic data integration

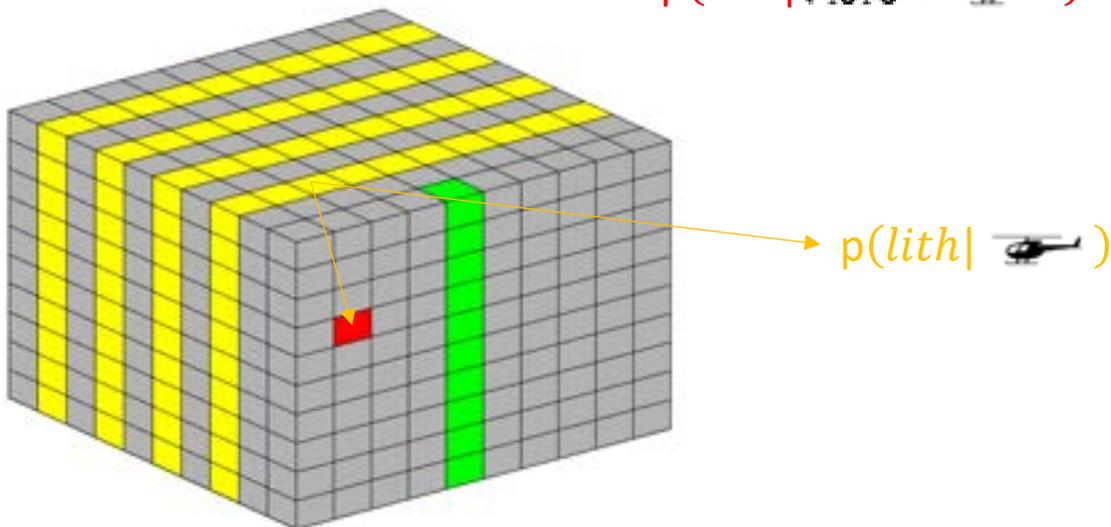
$$\begin{aligned}\rho_{tot}(\mathbf{m}) &= k\rho(\mathbf{m}|I_1, I_2, \dots, I_N) \\ &= k\rho(\mathbf{m}|I_1) * \rho(\mathbf{m}|I_2) * \dots * \rho(\mathbf{m}|I_N)\end{aligned}$$

$$\rho_{tot}(\mathbf{m}) = k\rho(\mathbf{m}|I_1) * \rho(\mathbf{m}|I_2)$$





$$p(lith | \text{crane icon} \ \& \ \text{helicopter icon}) = k \ p(lith | \text{crane icon}) * p(lith | \text{helicopter icon}) * p(lith | I_3) \dots p(lith | I_n)$$



Challenge!

1. How to **describe/translate** all subsurface Information (with uncertainty) as probability
2. How to make geologically informed **extrapolation** of the information

Two new concepts

Independent Information

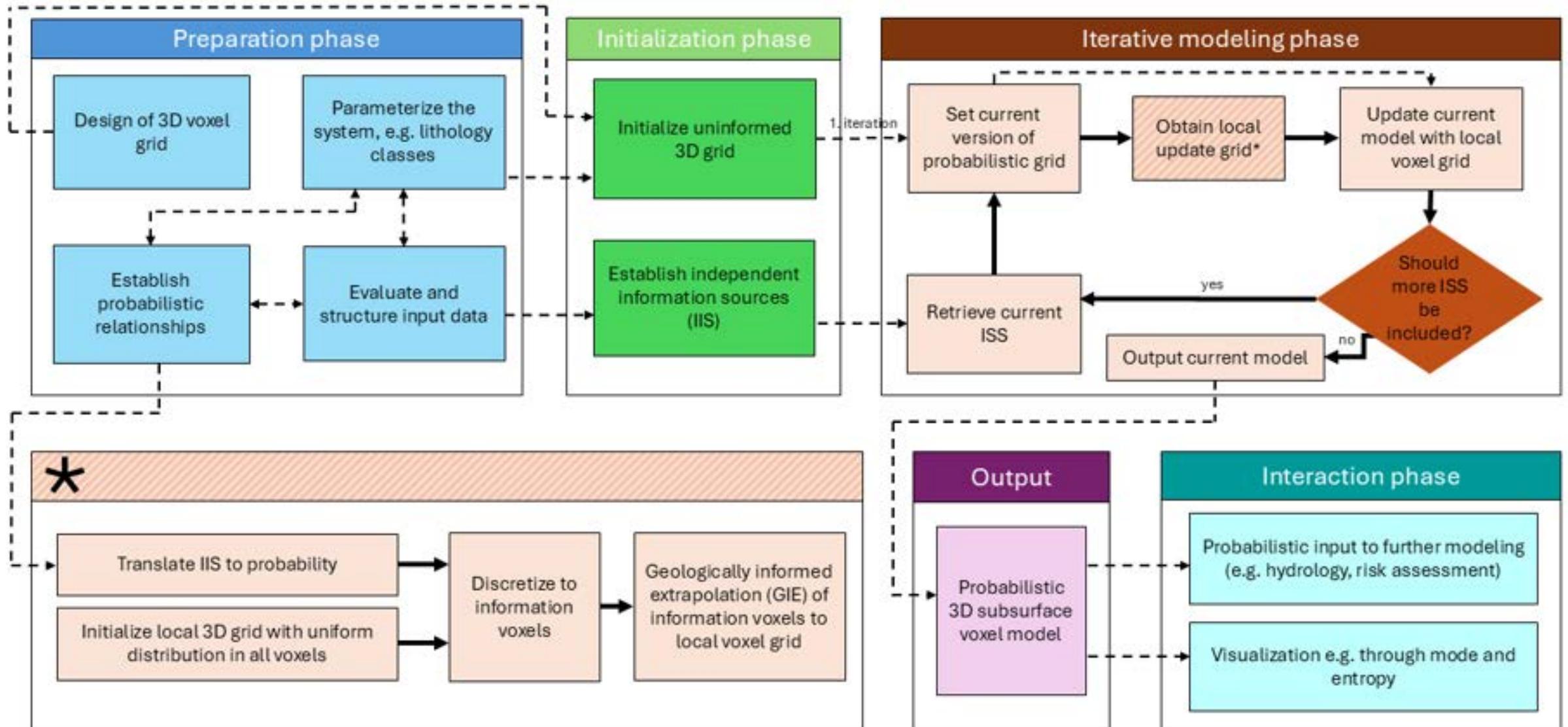
Sources (IIS)

- Self-contained units of subsurface information (e.g. boreholes, geophysics, interpretations)
- Translated into voxel-scale probability distributions with explicit uncertainty
- Assumed conditionally independent, enabling systematic combination and local updates

Geologically Informed Extrapolation (GIE)

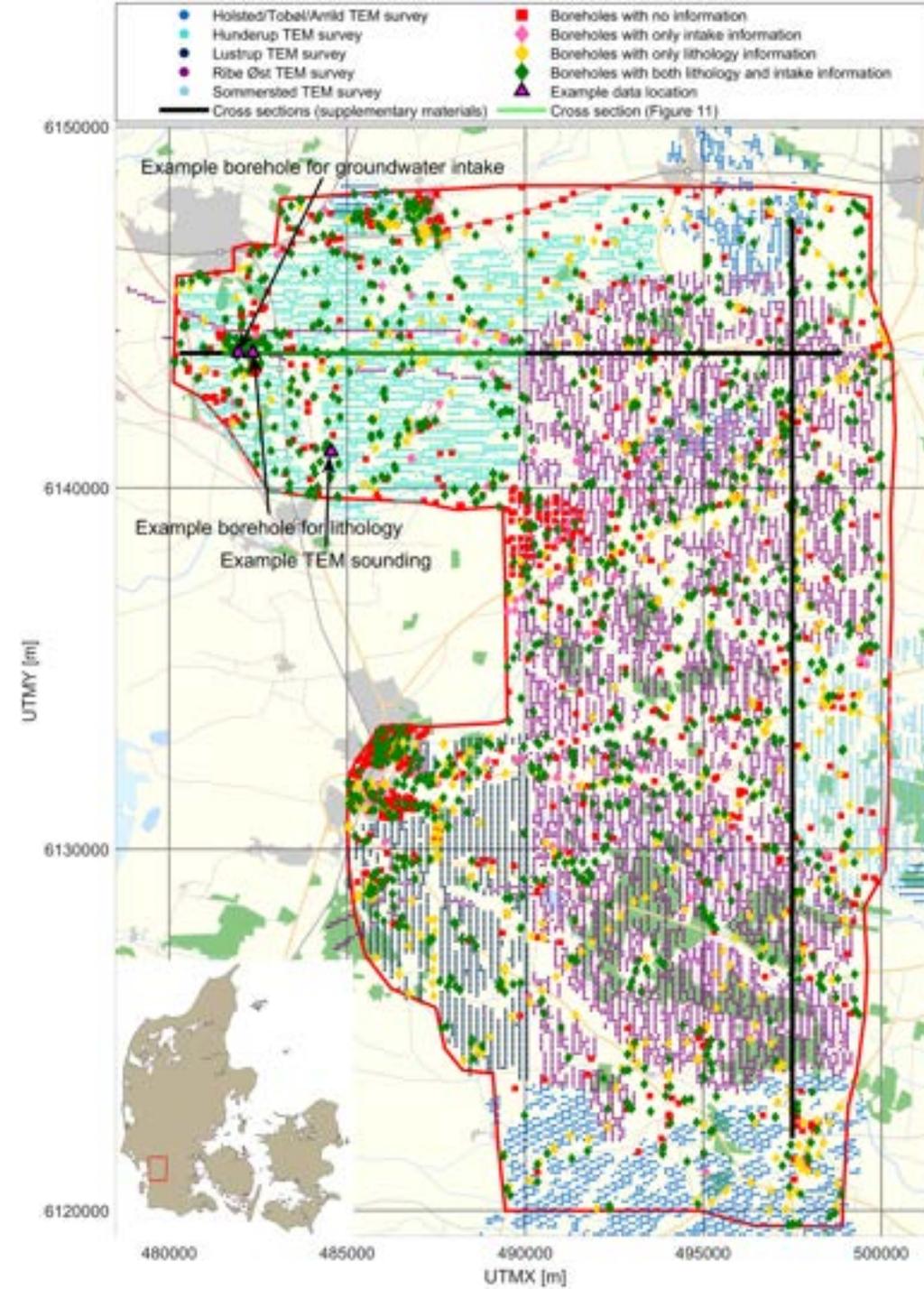
- Mechanism for spatially propagating IIS information into the 3D voxel grid
- Based on geological reasoning rather than statistical interpolation
- Defines controlled spatial influence in a transparent and efficient way

Workflow

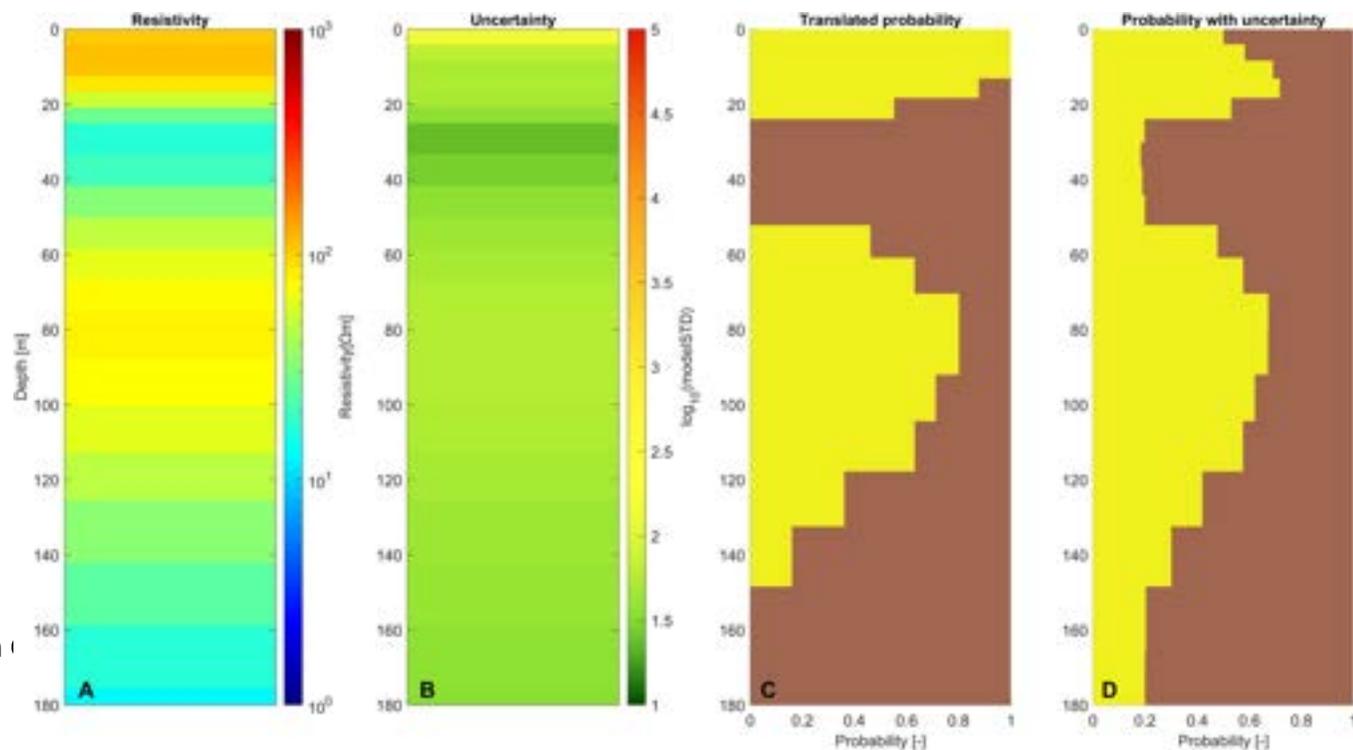
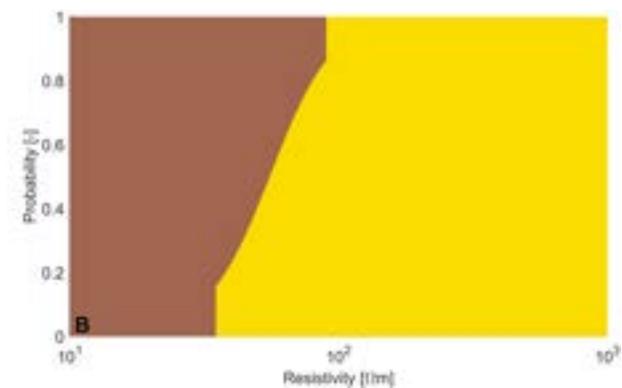
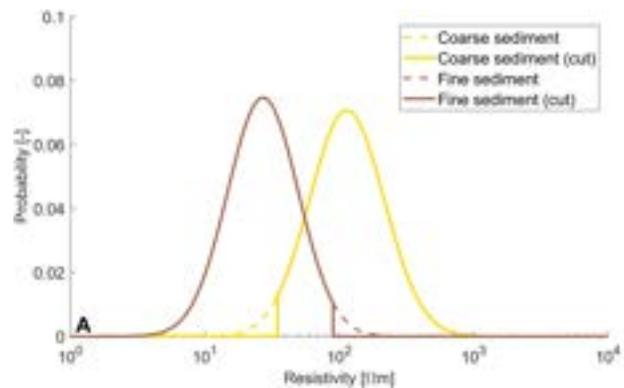
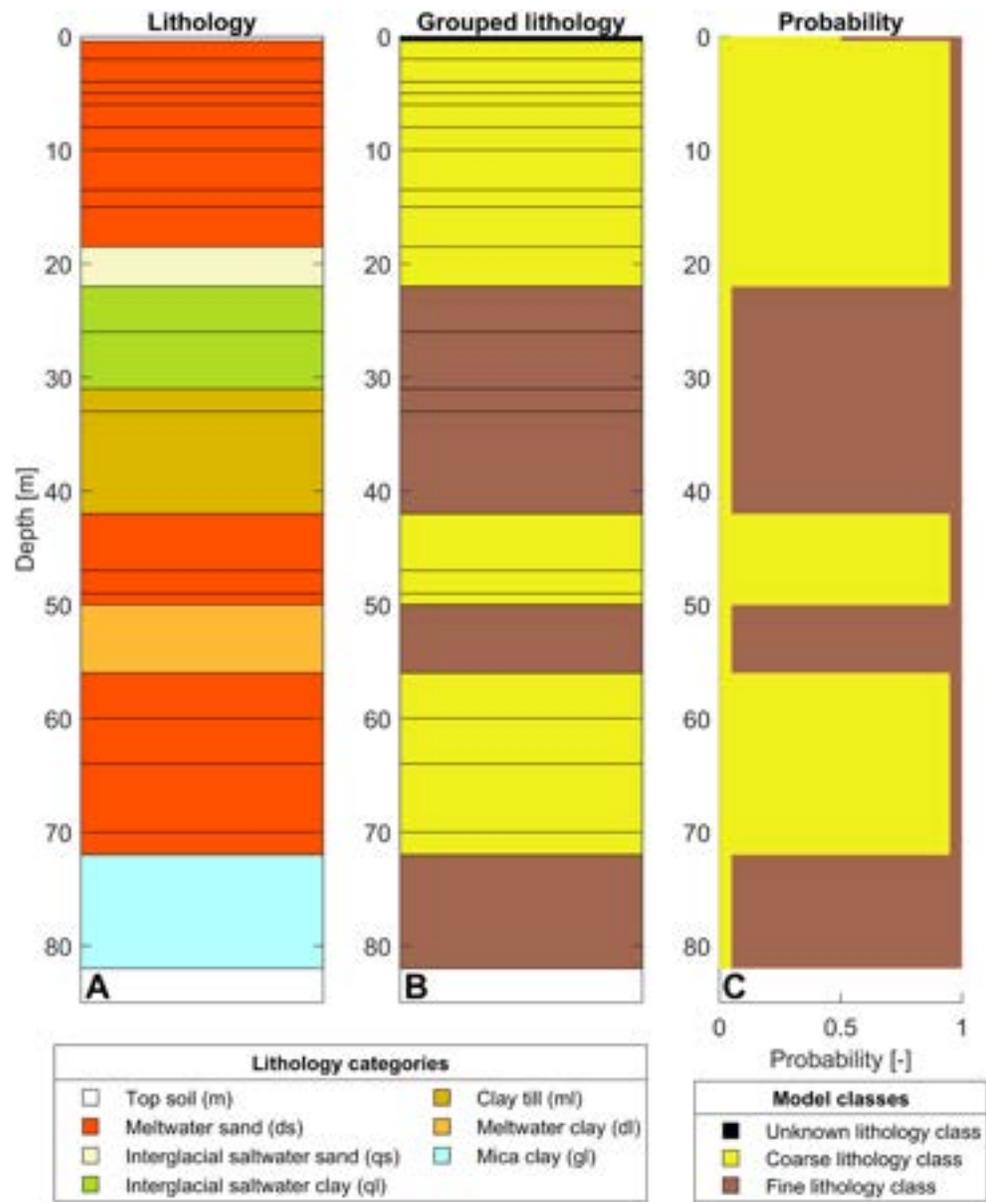


Example: Hydrogeological modeling Ribe, Denmark

- The heterogenous Saalian unit was modeled using the proposed voxel modeling framework
- Data sources
 - Boreholes
 - SkyTEM surveys
 - Water extraction

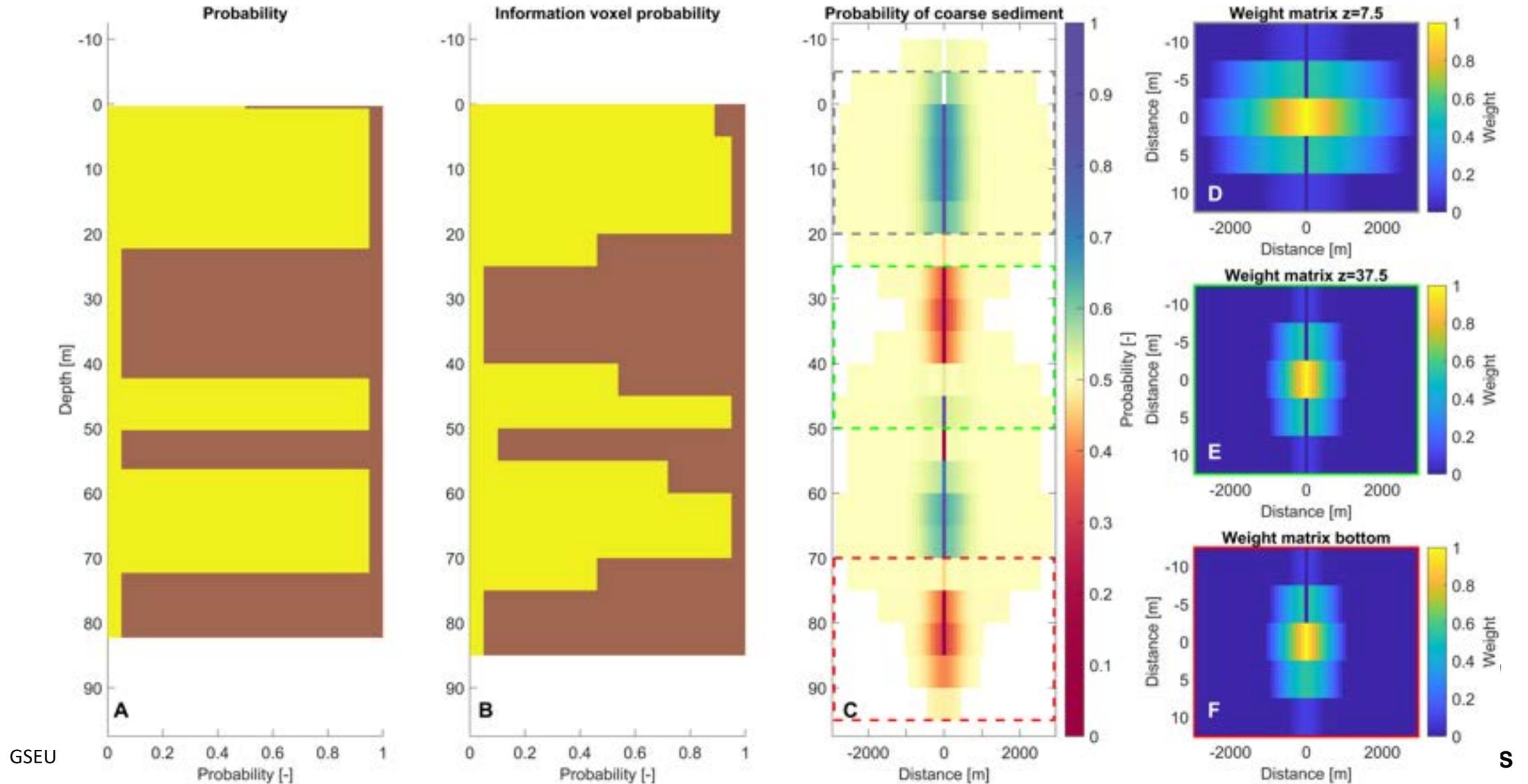


Independent Information Sources (IIS)

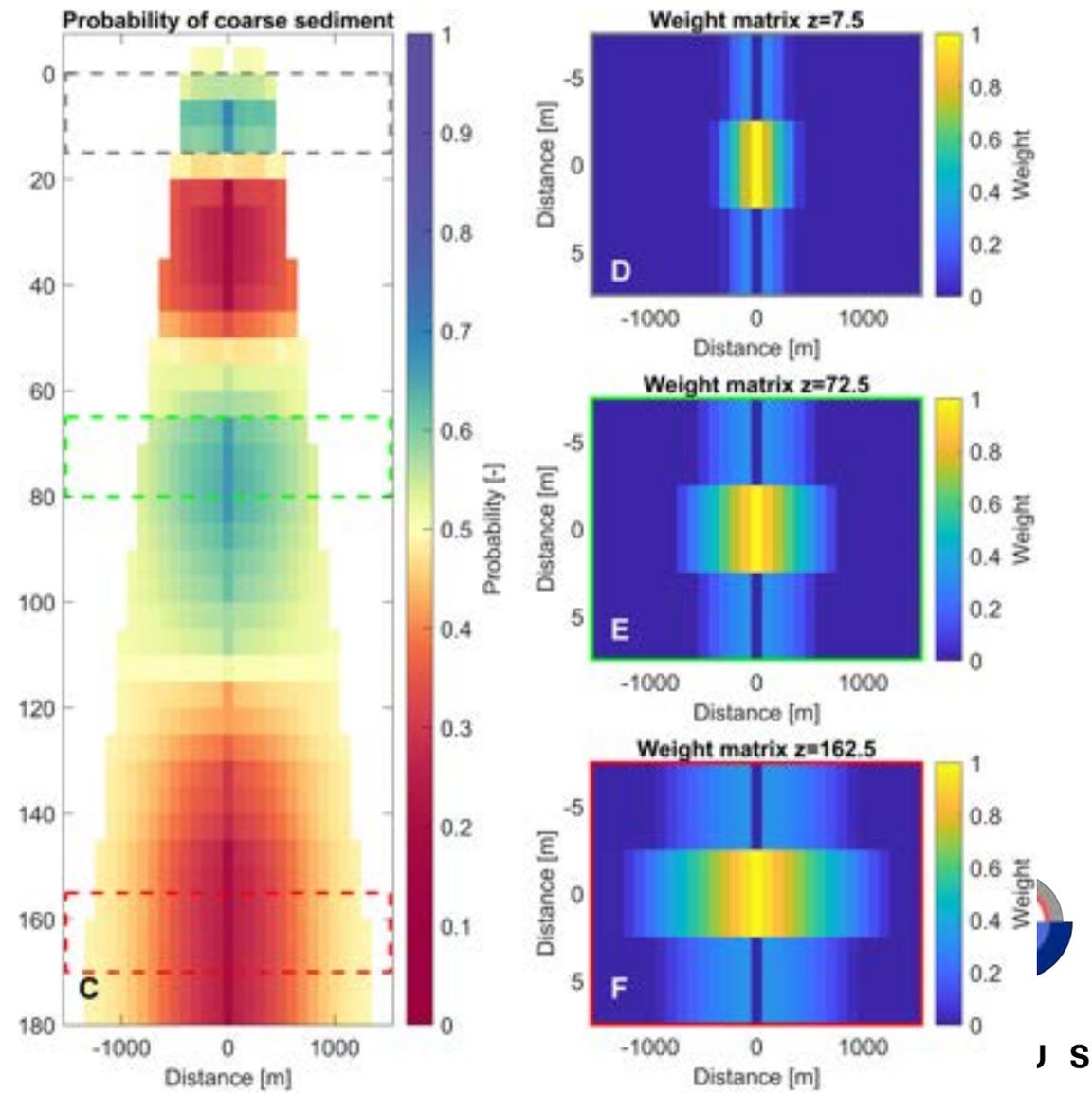
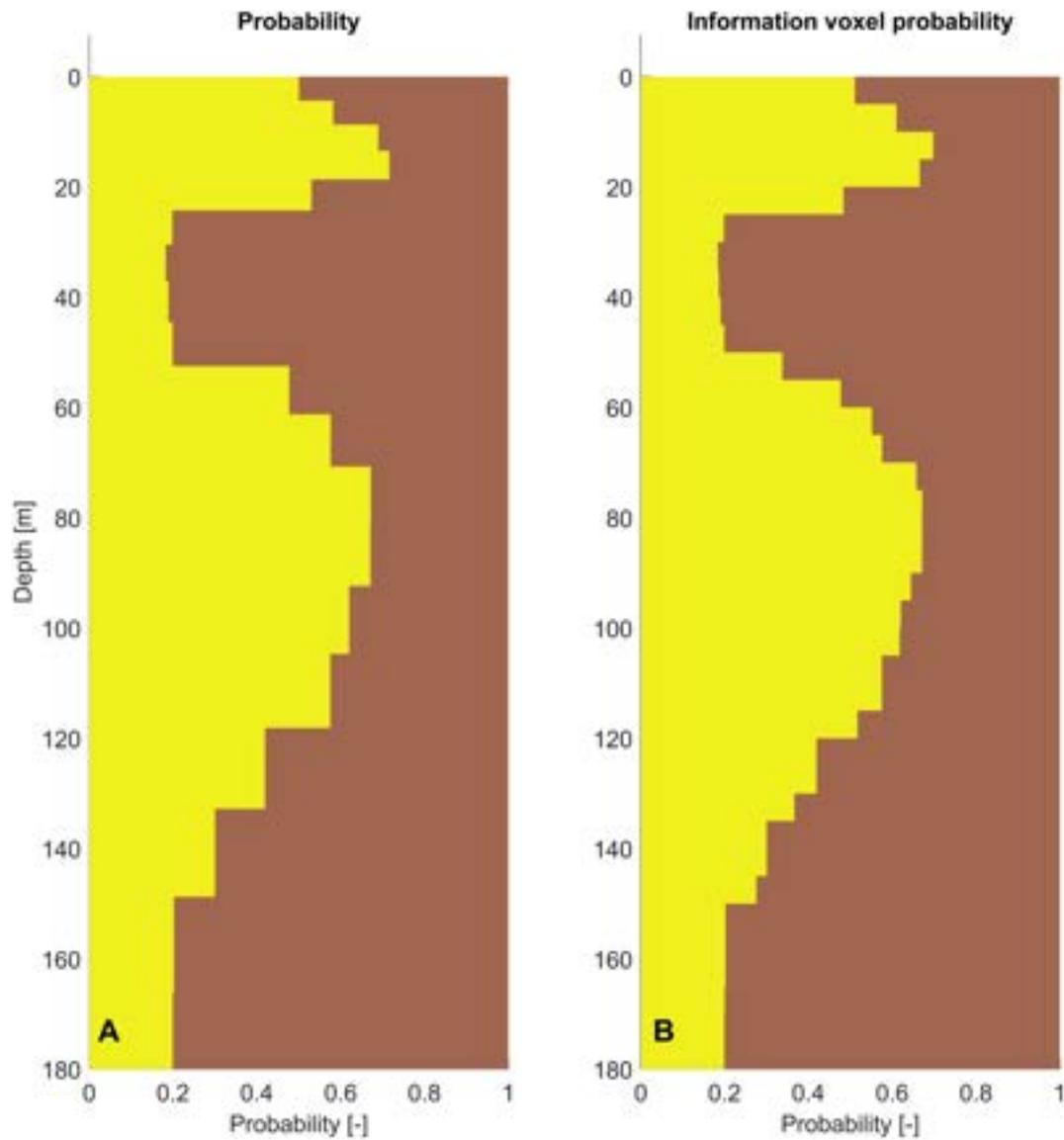


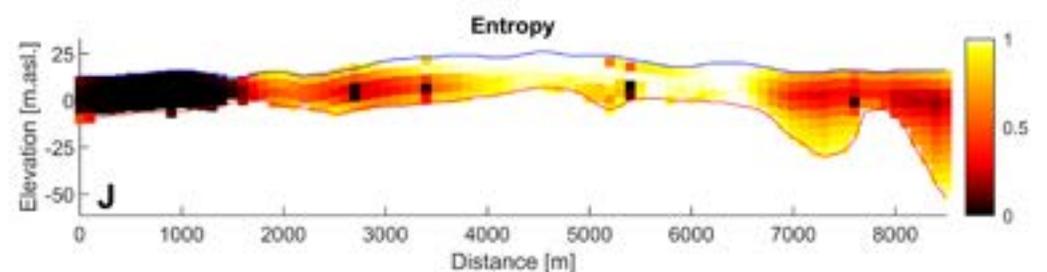
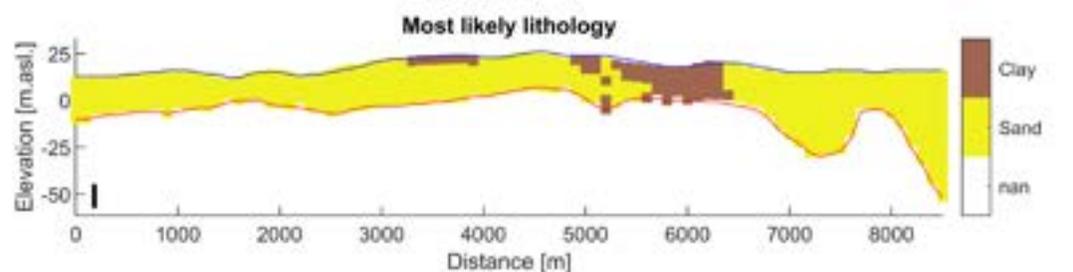
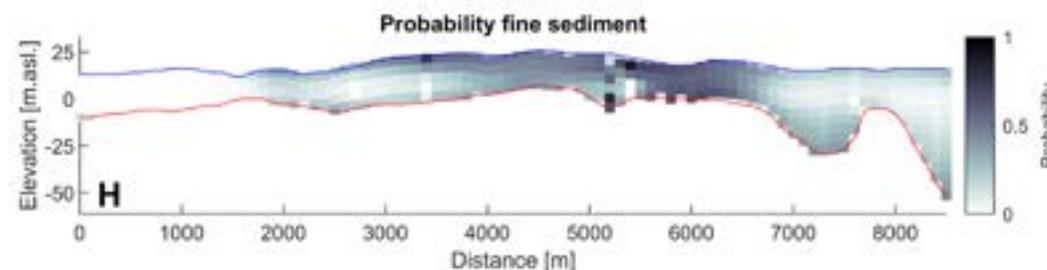
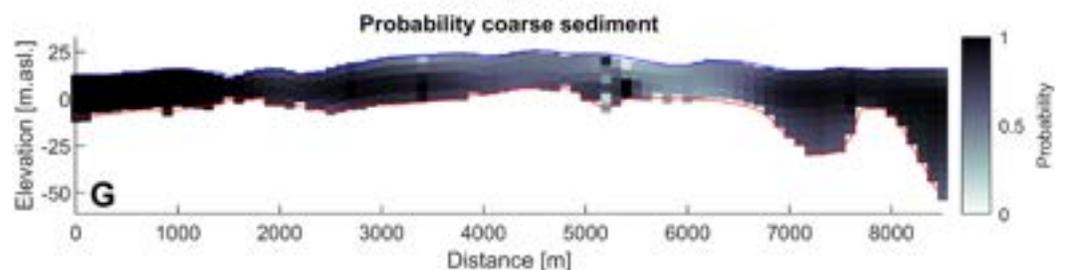
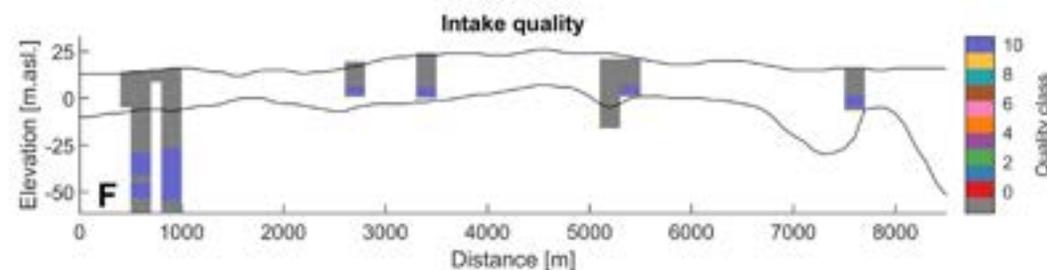
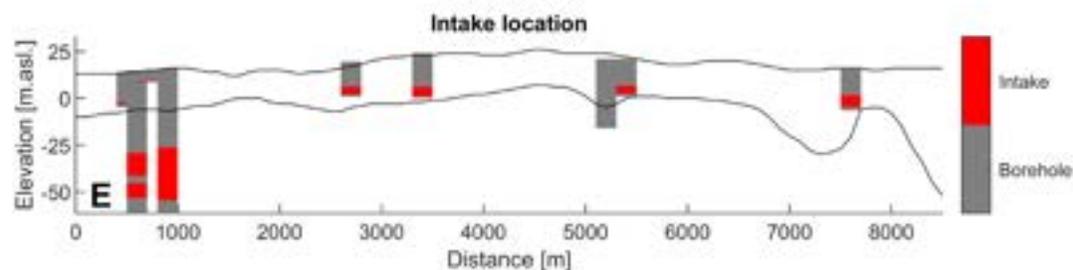
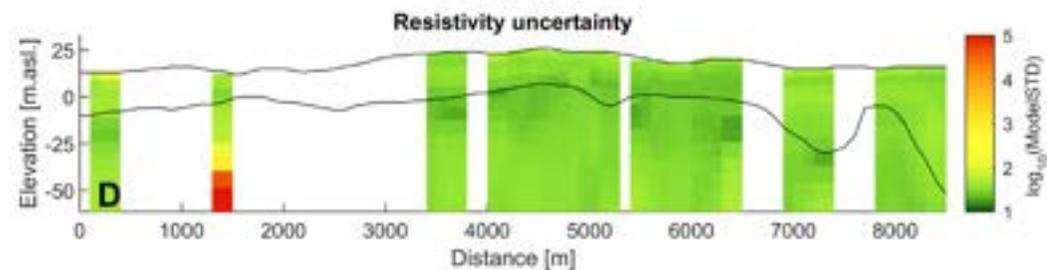
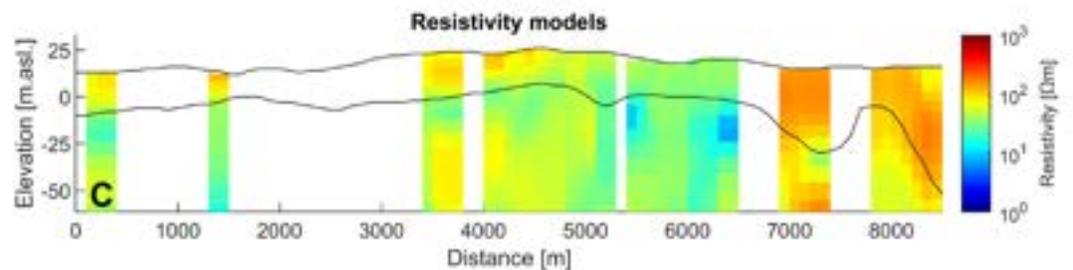
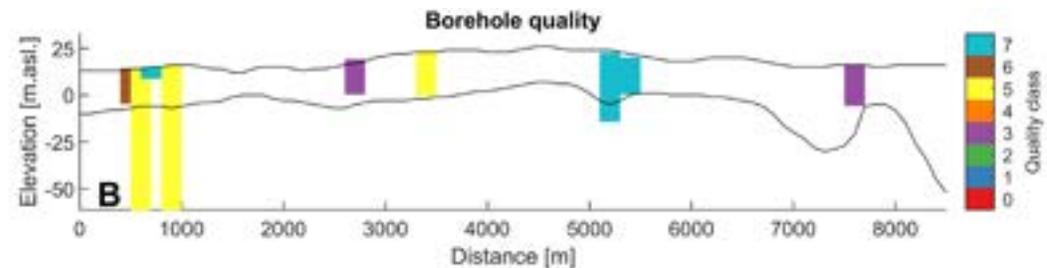
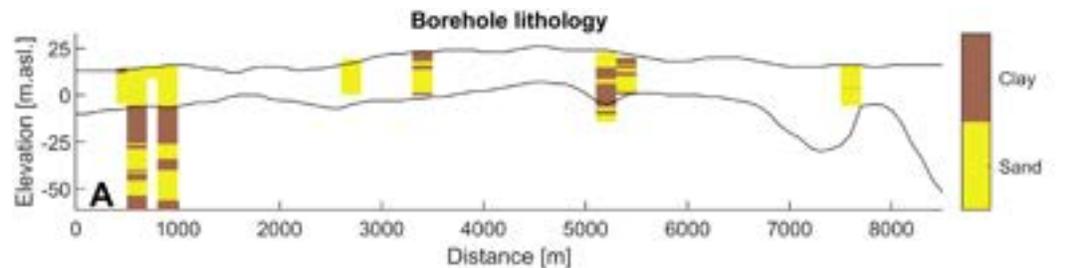
ainties in

Geologically Informed Extrapolation (GIE) - Borehole



Geologically Informed Extrapolation (GIE) - SkyTEM





U S

Describe information as probability distributions of the model parameters

- We have considered:
 - Boreholes
 - Lithology
 - Water production
 - TEM (SkyTEM, tTEM)
 - Surface geology maps
- Any subsurface information can in theory be used if a relation to chosen model parameters can be established

Geologically informed extrapolation of the information

- We have considered:
 - Simple geological knowledge
 - Shape and weight are completely free to design
 - Geophysical method based extrapolation
 - TEM sensitivity functions
- The modeling framework is adaptable to any spatial information available

Paper submitted to Engineering Geology

Title: Advancing probabilistic uncertainty-incorporated 3D voxel modelling – the application of uncertainty

Thank you

How are uncertainties taken into account in the SPDE process?

Charlie GARAYT^(1,2), Nicolas Desassis⁽¹⁾, Nicolas Clausolles⁽²⁾,
Simon Lopez⁽²⁾

(1) Mines Paris, PSL

(2) BRGM

02/03/2026

What is SPDE? (Lindgren and Rue, 2011)

SPDE: stochastic partial differential equation

$$(\kappa^2 - \Delta)^{\alpha/2} Z = \overbrace{W}^{\text{Gaussian white noise}}$$

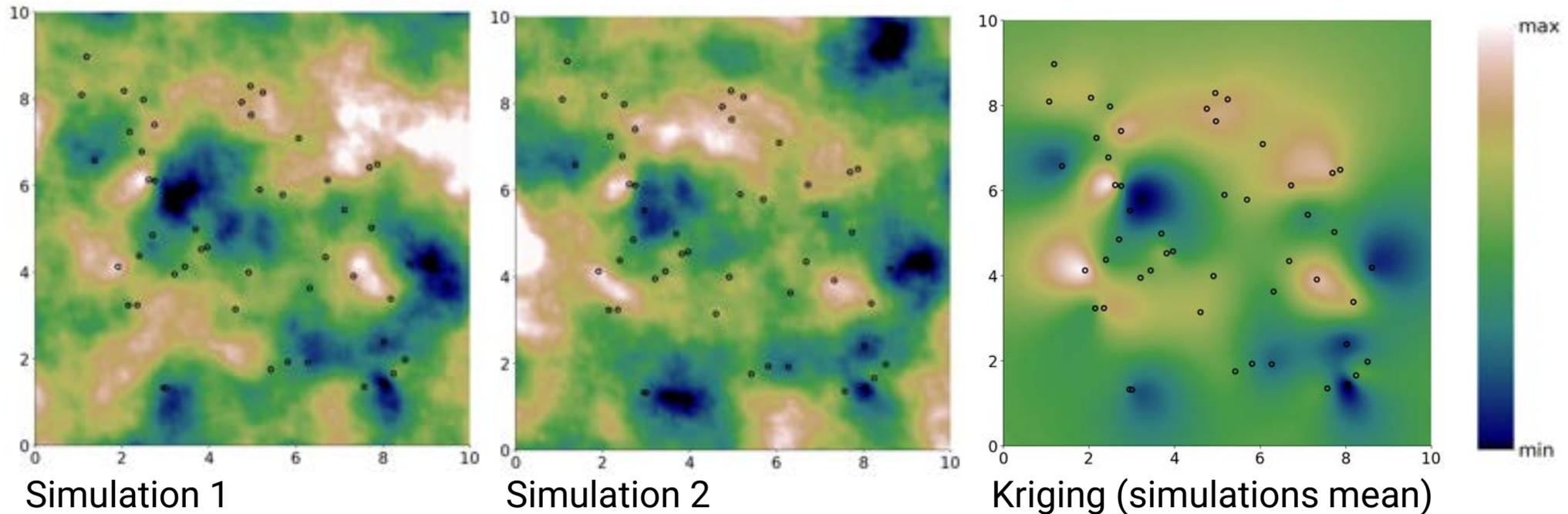
What is SPDE? (Lindgren and Rue, 2011)

SPDE: stochastic partial differential equation

What we are looking for $\underbrace{\hspace{1.5cm}}$ Gaussian white noise $\underbrace{\hspace{1.5cm}}$

$$(\kappa^2 - \Delta)^{\alpha/2} \mathbf{Z} = \mathbf{W}$$

\mathbf{Z} : a Gaussian random field (GRF)



What is SPDE? (Lindgren and Rue, 2011)

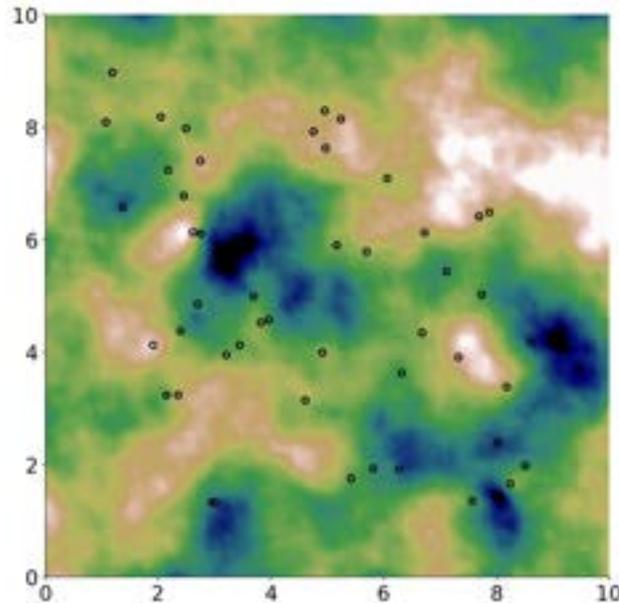
SPDE: stochastic partial differential equation

What we are
looking for Gaussian white noise

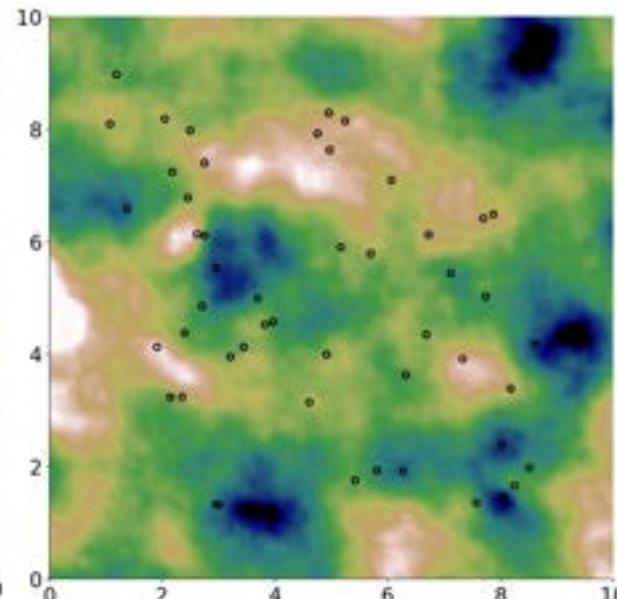
$$(\kappa^2 - \Delta)^{\alpha/2} Z = W$$

Two parameters > 0

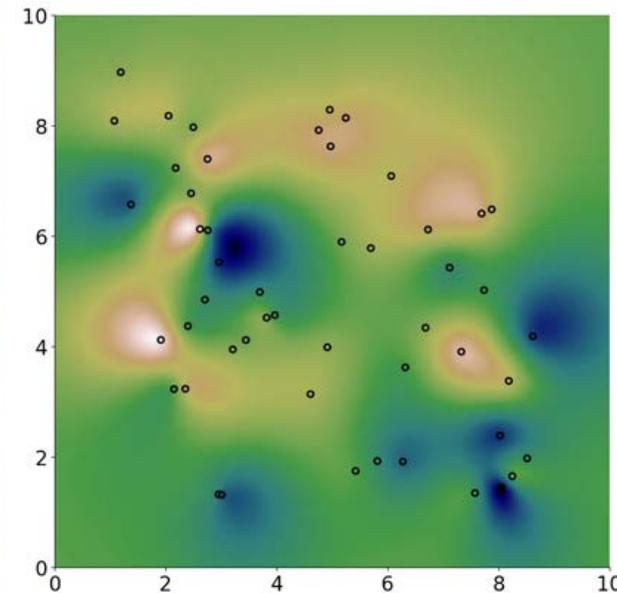
Z : a Gaussian random field (GRF)



Simulation 1



Simulation 2



Kriging (simulations mean)

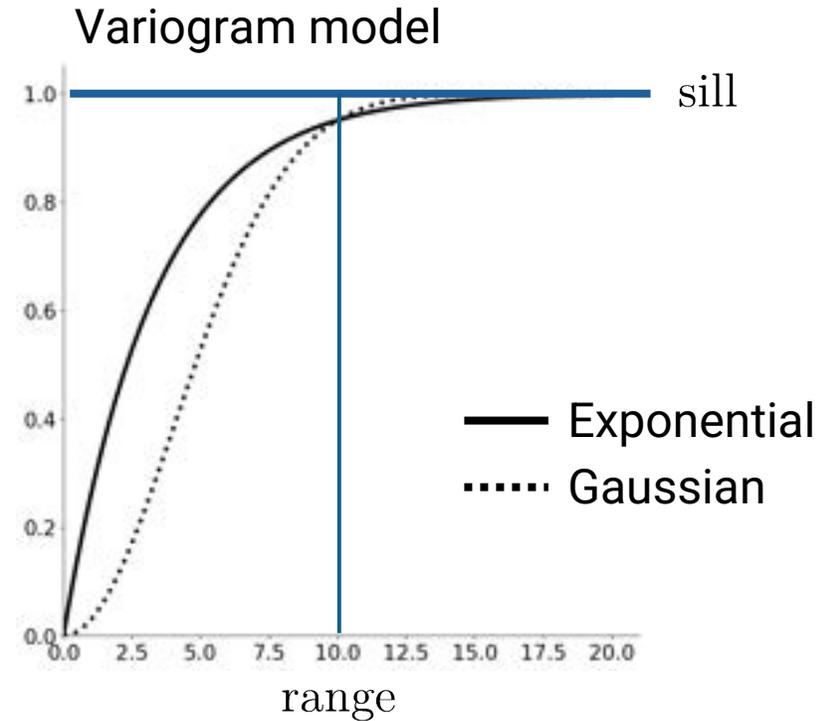


Computation of a Gaussian random field

“Classic” geostatistics, covariance approach

1. Definition of a covariance/variogram model from data
2. Computation of the covariance matrix
3. Inversion of the covariance matrix

Computation time function of n_{data} ($O(n^3)$)



Computation of a Gaussian random field

“Classic” geostatistics, covariance approach

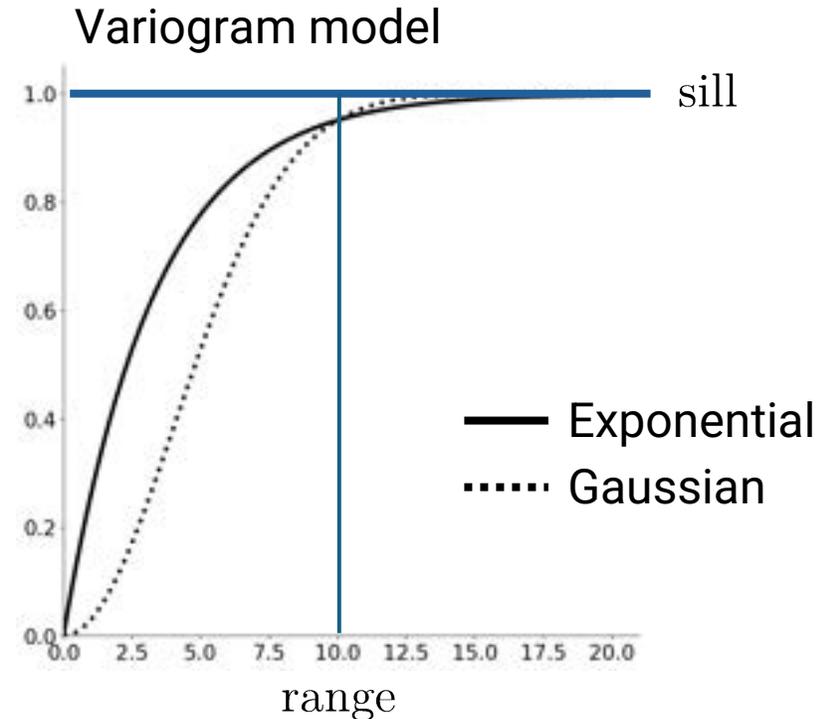
1. Definition of a covariance/variogram model from data
2. Computation of the covariance matrix
3. Inversion of the covariance matrix

Computation time function of n_{data} ($O(n^3)$)

SDPE approach

1. Definition of parameters κ and α (see next slide)
2. Definition of an underlying mesh for finite element method (FEM)
3. Computation of the precision matrix (inverse of the covariance matrix) by solving the SPDE with FEM

Computation time not function of n_{data}

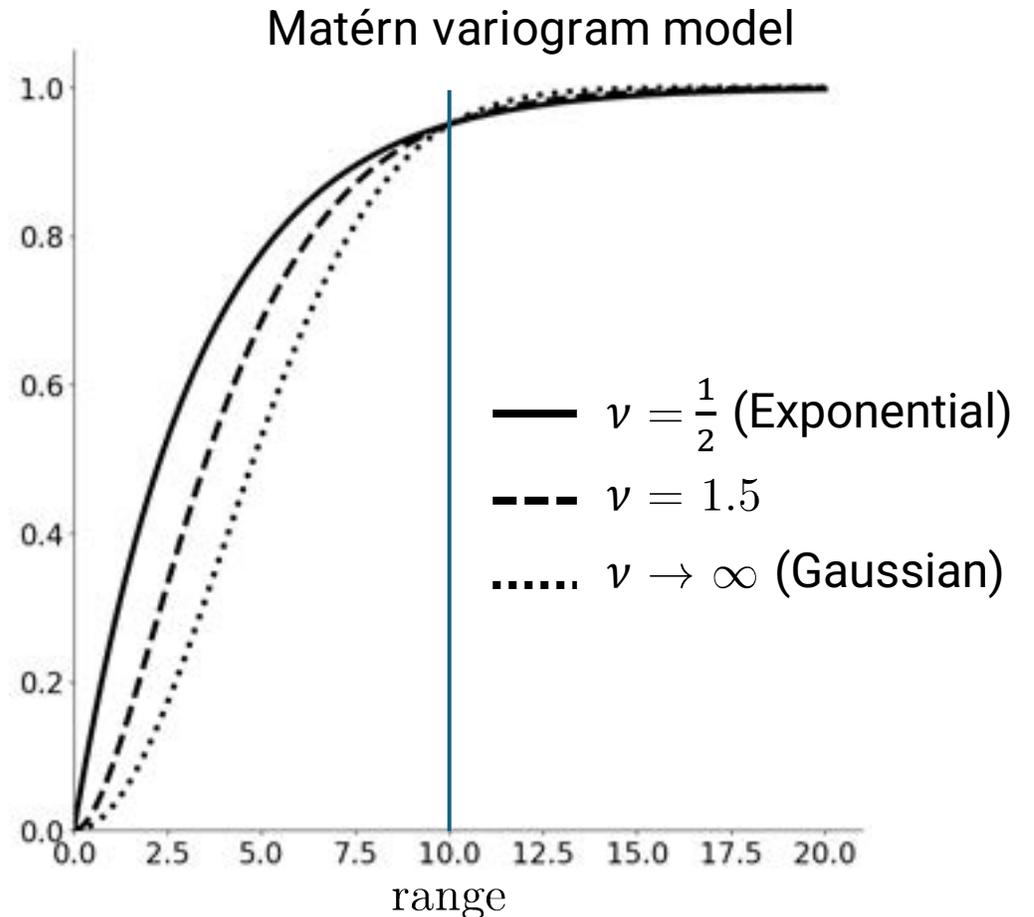
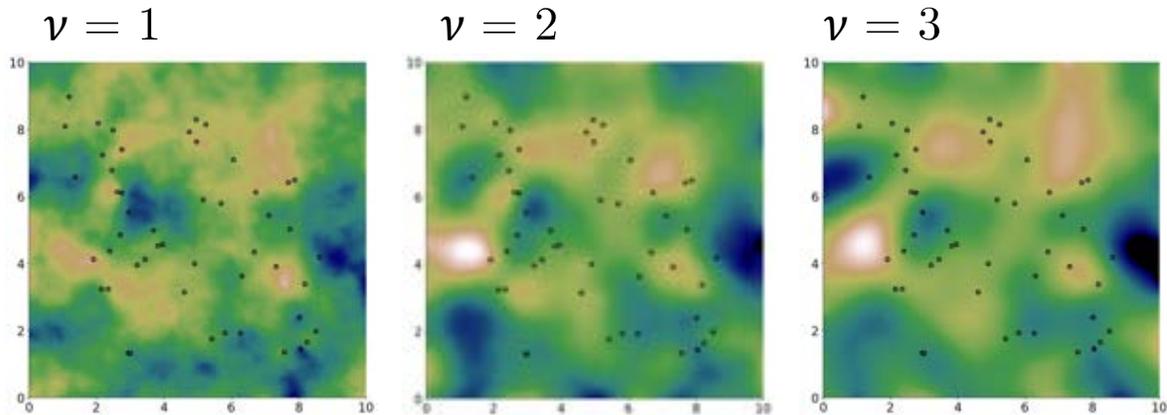


Link between κ and α and variogram models

Solution Z of the SDPE $(\kappa^2 - \Delta)^{\alpha/2} Z = W$ is a GRF parametrized by a Matérn covariance
(Whittle, 1954)

Link between κ and α and variogram models

Solution Z of the SDPE $(\kappa^2 - \Delta)^{\alpha/2} Z = W$ is a GRF parametrized by a Matérn covariance
→ Equivalence between the SPDE approach and the covariance approach (Whittle, 1954)



Link between κ and α and variogram models

Solution Z of the SDPE $(\kappa^2 - \Delta)^{\alpha/2} Z = W$ is a GRF parametrized by a Matérn covariance (Whittle, 1954)

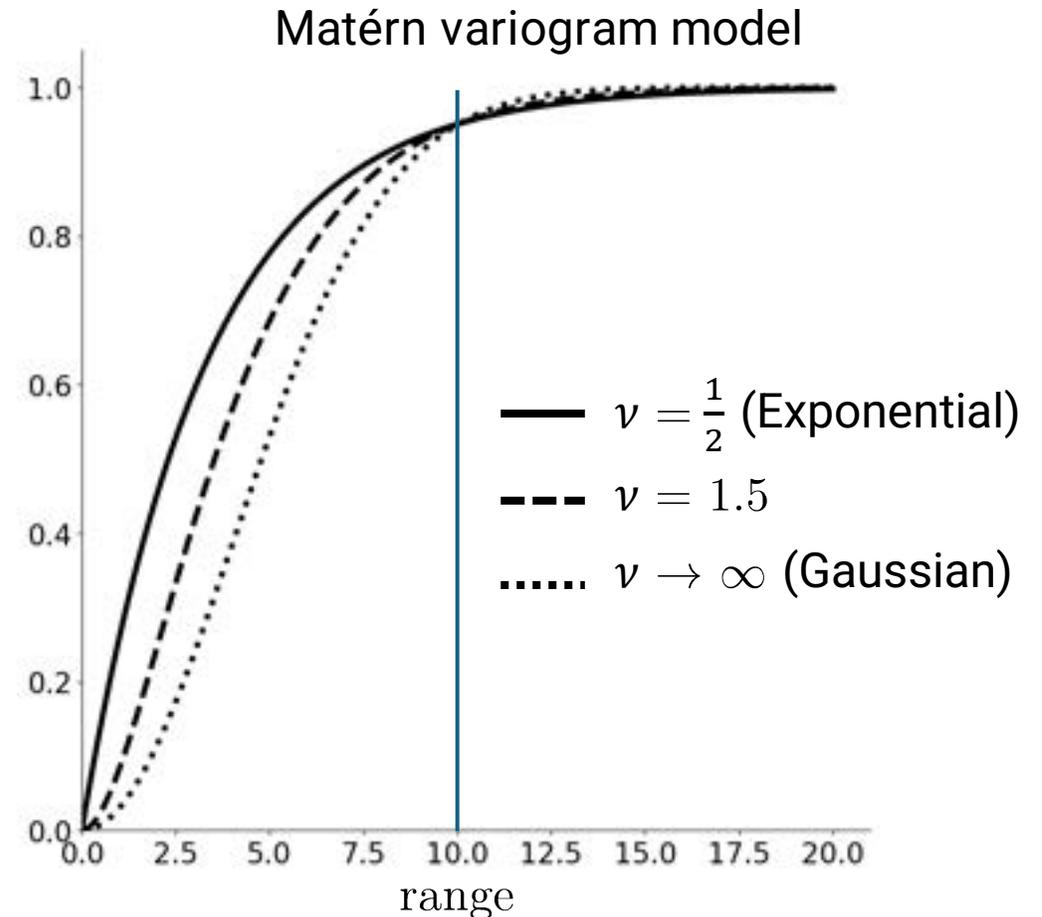
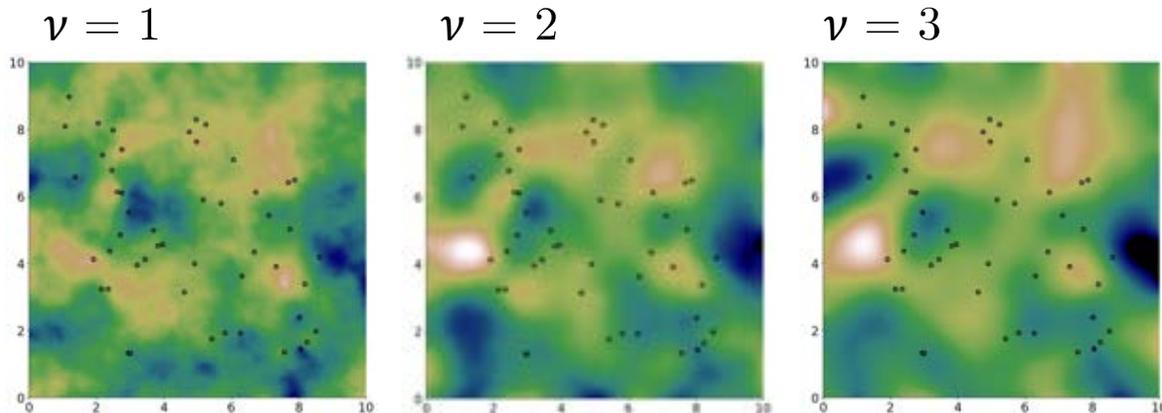
→ Equivalence between the SPDE approach and the covariance approach

(Whittle, 1954)

→ SPDE parameters have a physical interpretation

$\nu = \alpha - \frac{d}{2}$ smoothness | d : dimension

$$\kappa \propto \frac{1}{\text{range}}$$



Computation times benchmark

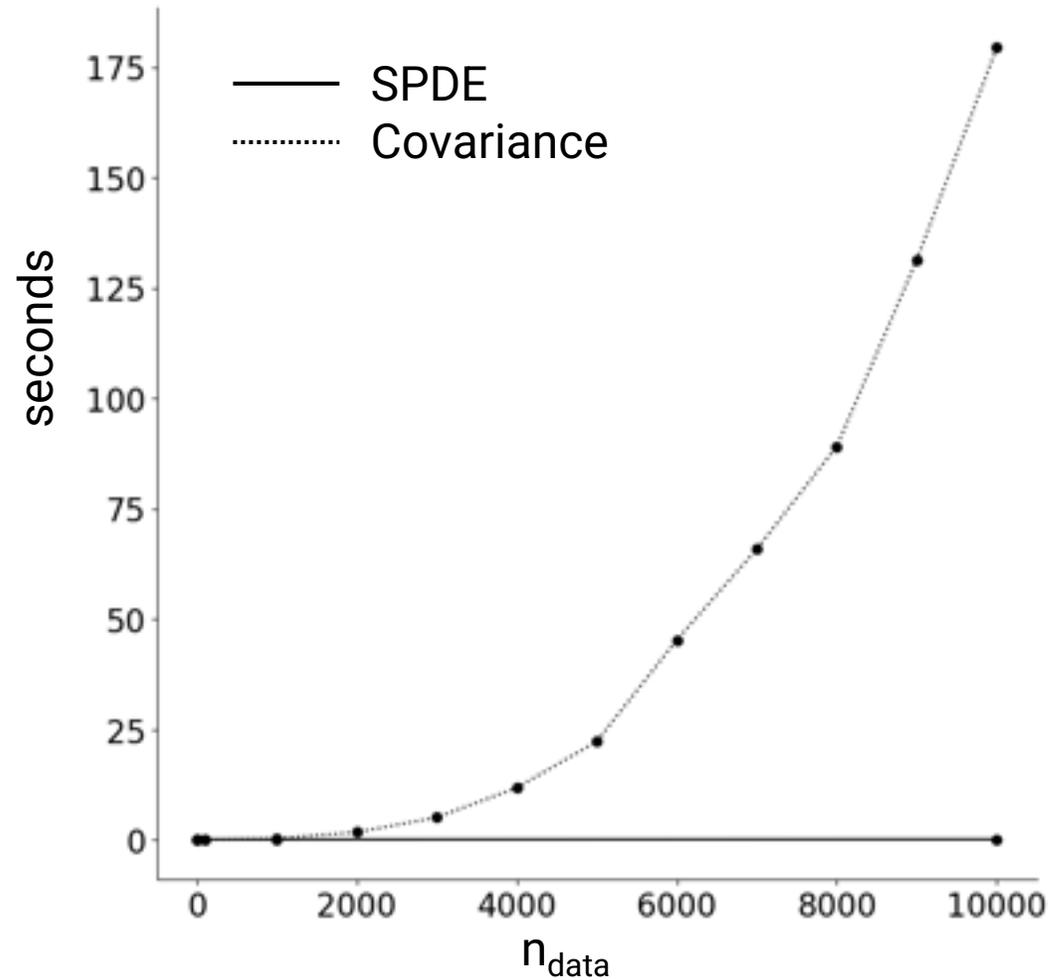
Number of data, SPDE approach vs covariance approach

Underlying mesh resolution

Smoothness param ν

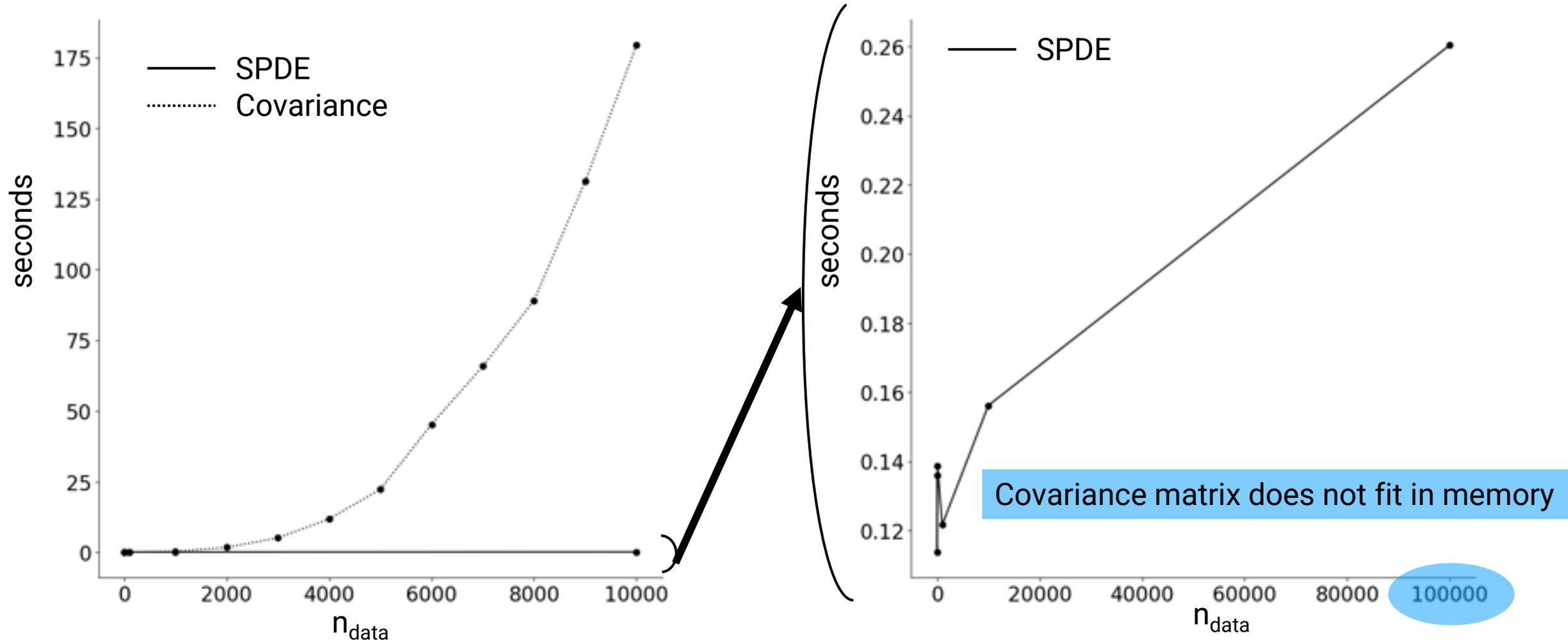
Computation times benchmark, SPDE vs covariance

Kriging computation time



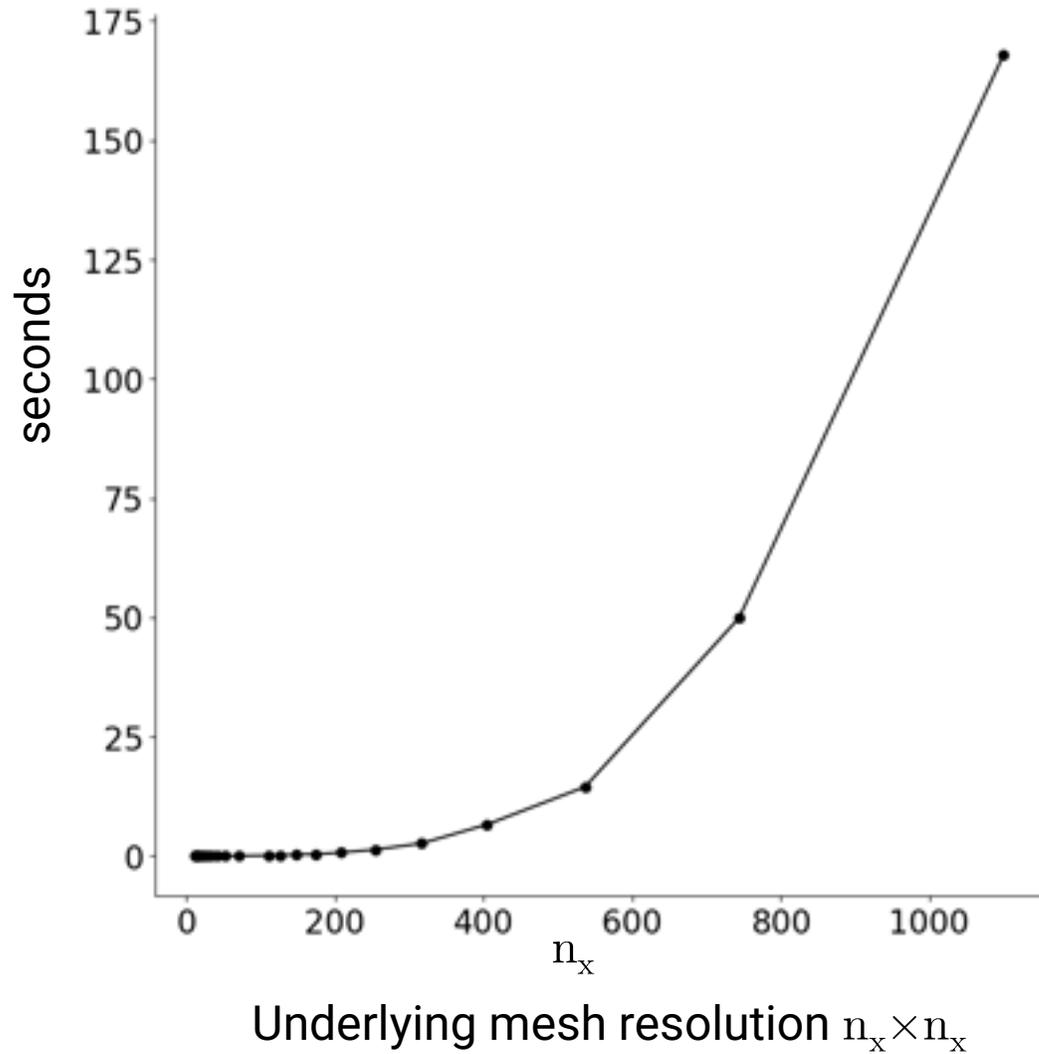
Computation times benchmark, SPDE vs covariance

Kriging computation time



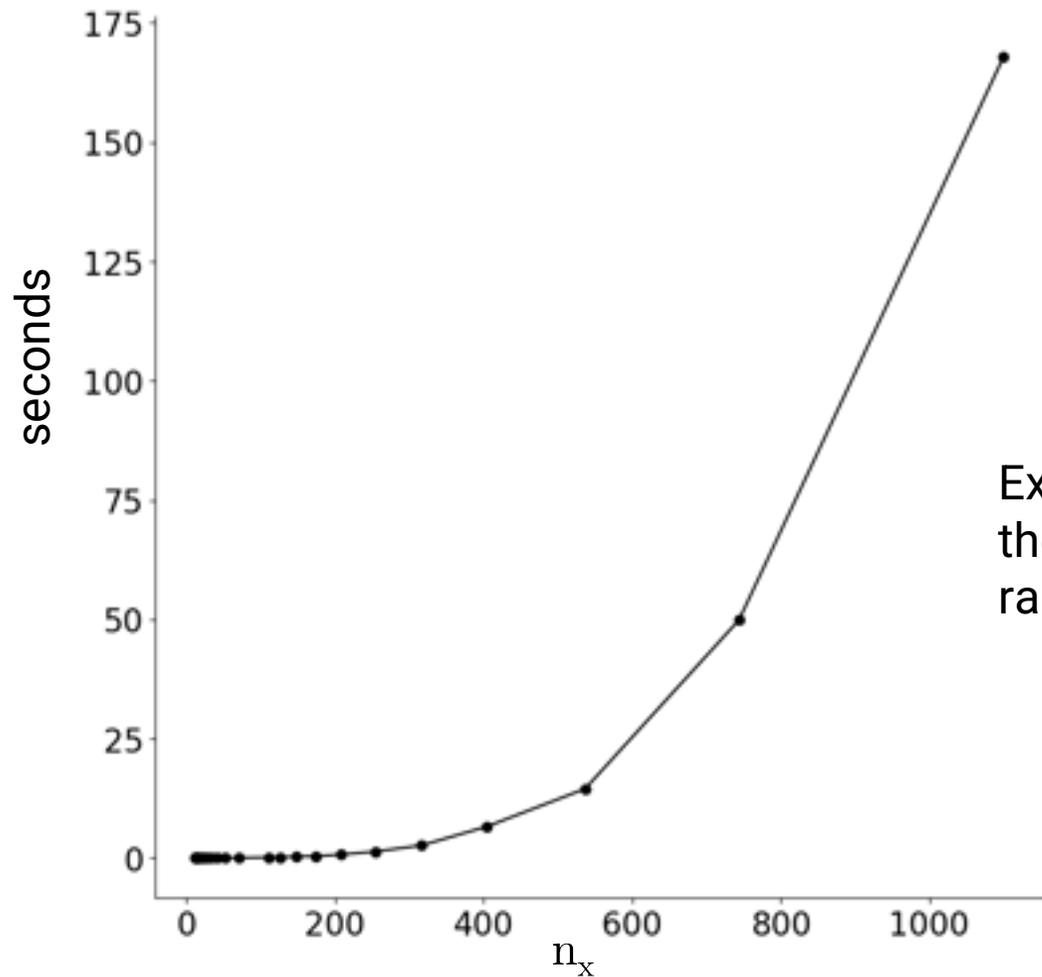
Computation times benchmark, underlying mesh

Kriging computation time

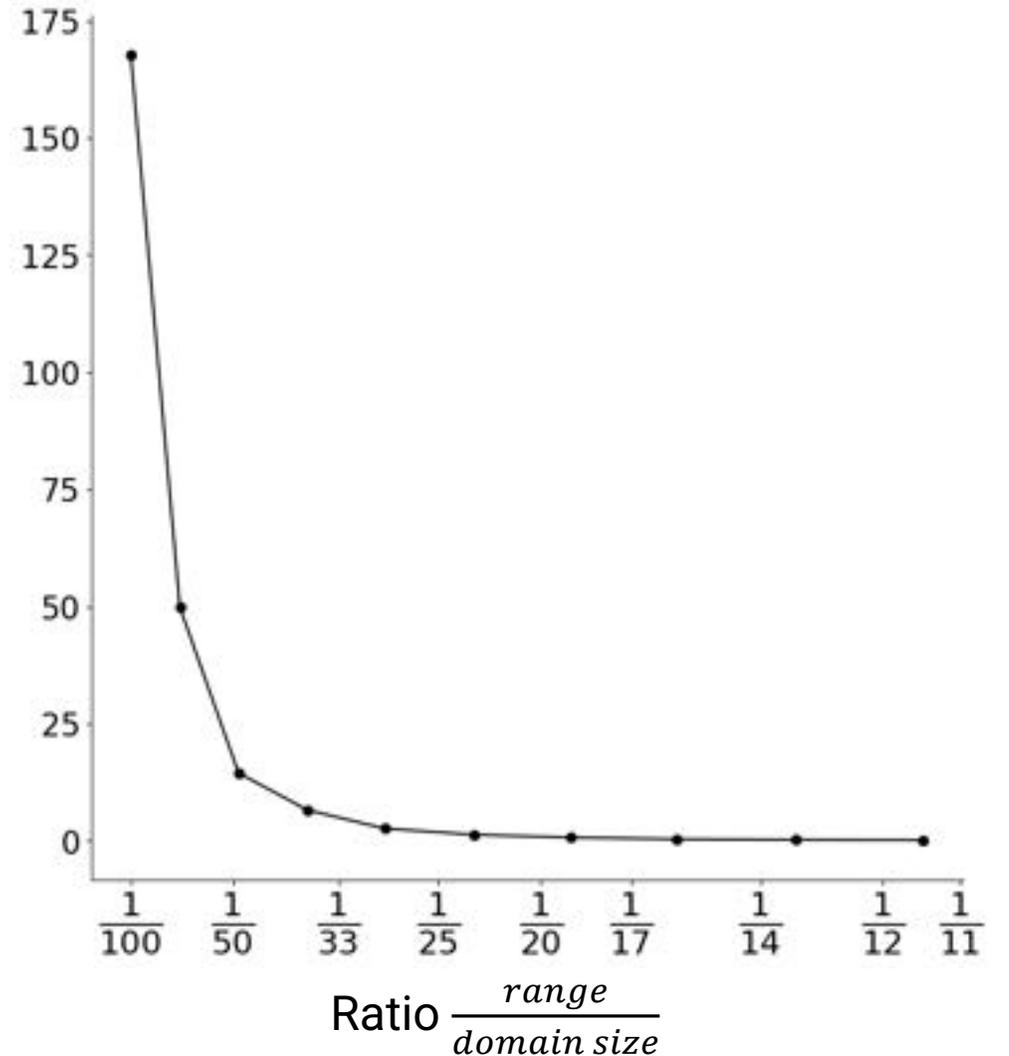


Computation times benchmark, underlying mesh

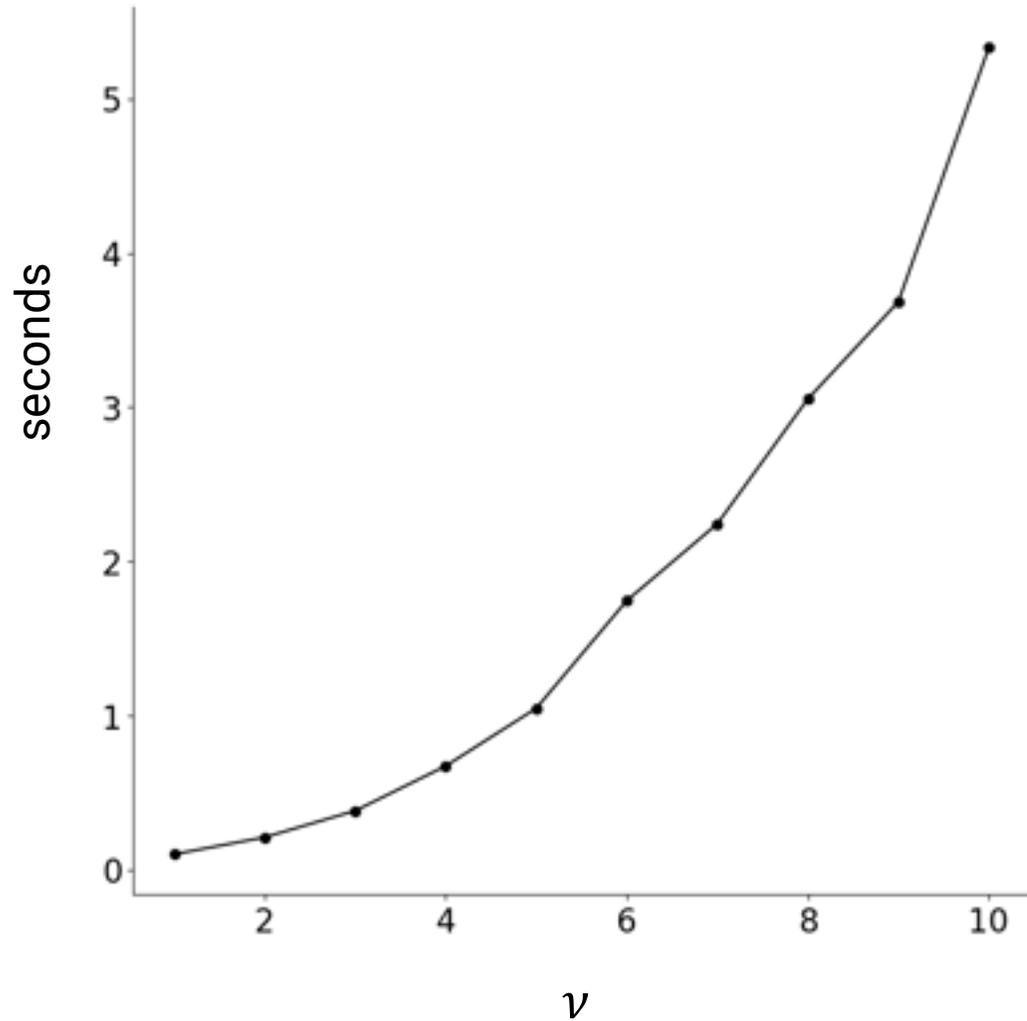
Kriging computation time



Expressed with the variogram range



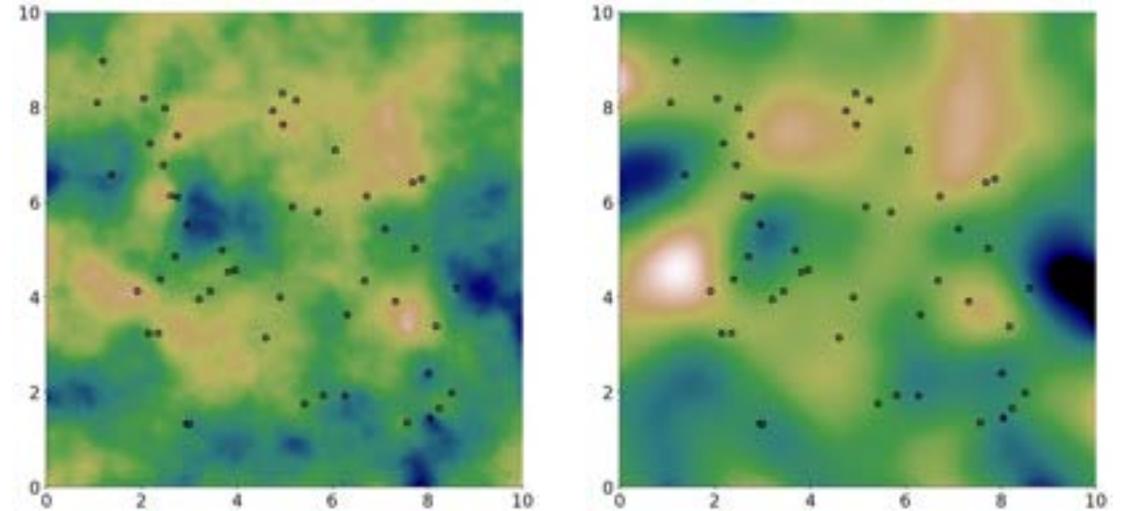
Computation times benchmark, smoothness param ν



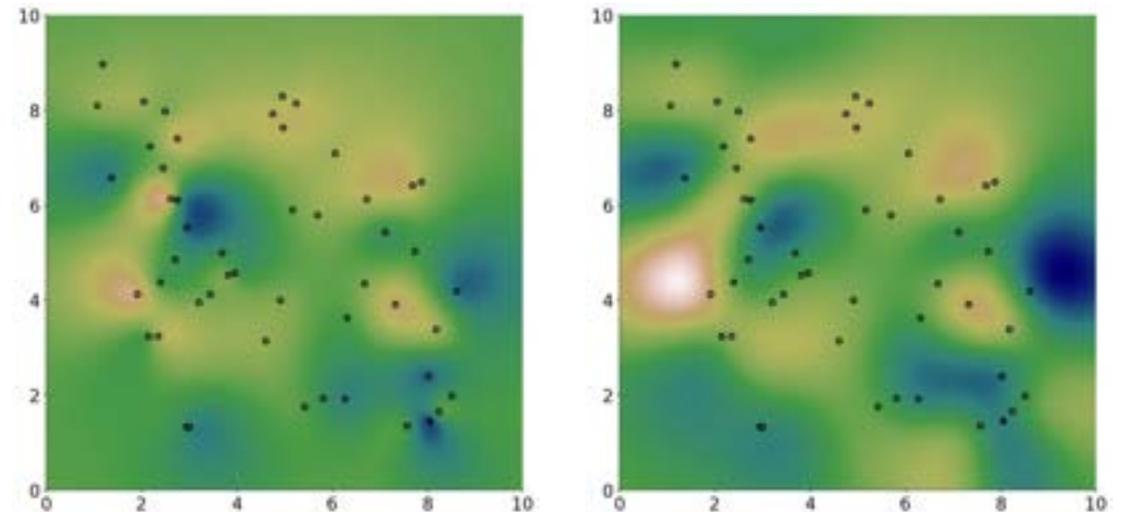
$\nu = 1$

$\nu = 3$

Simulation



Kriging



To summarize

- The SPDE approach allows to compute GRF
 - Computation times:
 - is independent of the amount of data
 - decrease when the variogram range of the GRF increase
 - increase when the smoothness of the GRF increase
- ➔ SPDE approach is faster than the covariance approach

To summarize

- The SPDE approach allows to compute GRF
- Computation times:
 - is independent of the amount of data
 - decrease when the variogram range of the GRF increase
 - increase when the smoothness of the GRF increase
- ➔ SPDE approach is faster than the covariance approach
- Let's go back to

How are uncertainties taken into account in the SPDE process?

Data uncertainties with the SPDE

- Assumption: data are coming from a GRF with a measurement noise

$$\rightarrow Y_{\text{data}} = AZ_{\text{GRF}} + \varepsilon$$

A: mask matrix, ε : Gaussian noise with variance σ^2

Data uncertainties with the SPDE

- Assumption: data are coming from a GRF with a measurement noise

$$\rightarrow Y_{\text{data}} = AZ_{\text{GRF}} + \varepsilon$$

A: mask matrix, ε : Gaussian noise with variance σ^2

- Due to this measurement noise/assumption, the conditioning is not exact*

*however, post-processing is possible for exact conditioning

Data uncertainties with the SPDE

- Assumption: data are coming from a GRF with a measurement noise

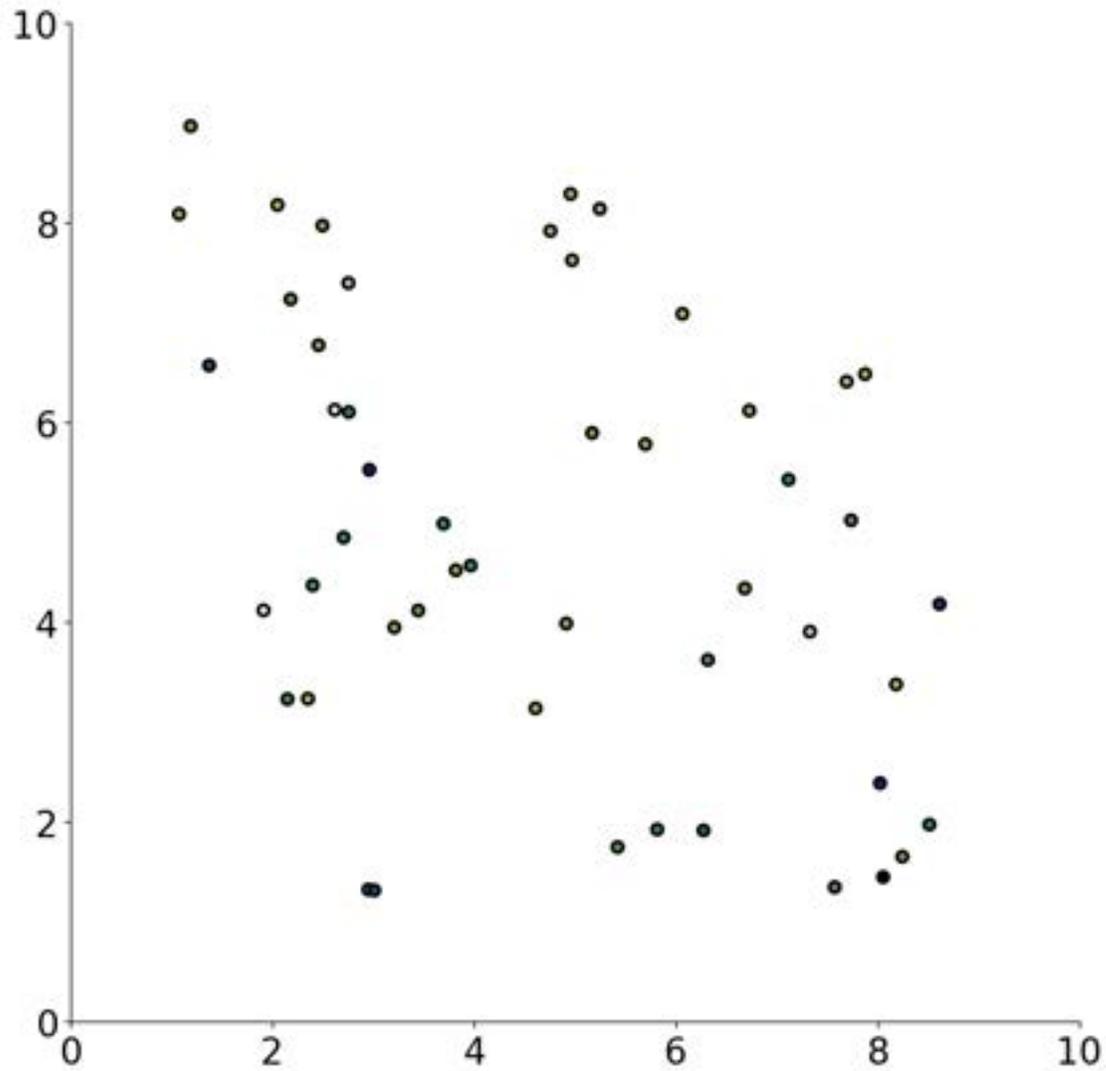
$$\rightarrow Y_{\text{data}} = AZ_{\text{GRF}} + \varepsilon$$

A: mask matrix, ε : Gaussian noise with variance σ^2

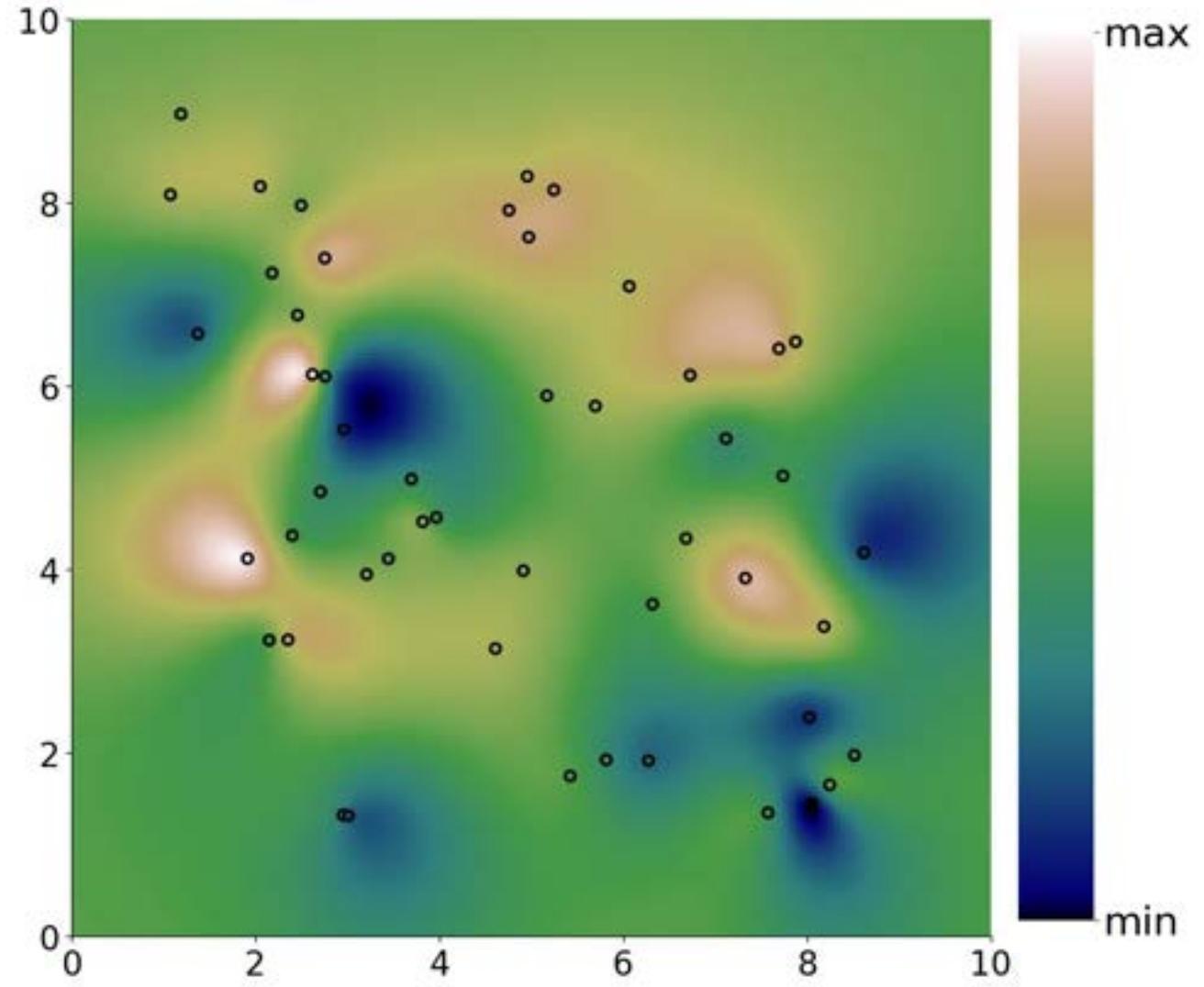
- Due to this measurement noise/assumption, the conditioning is not exact*
*however, post-processing is possible for exact conditioning
- Simulations are fast to compute, allowing to access uncertainties
- The simulations mean is the kriging

The kriging with SPDE

Conditioning data, e.g. altitude measurement

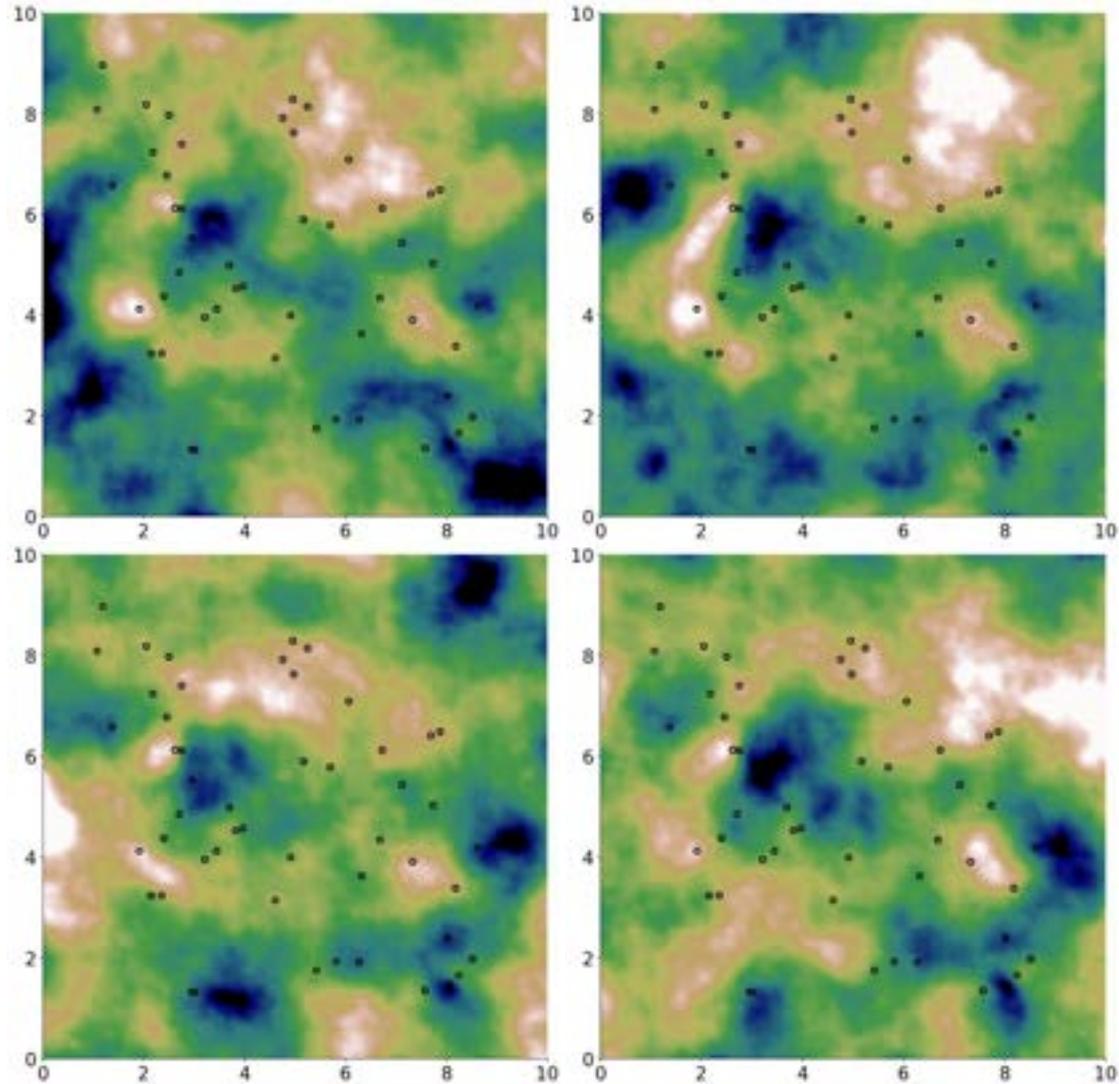


Kriging (simulations means)

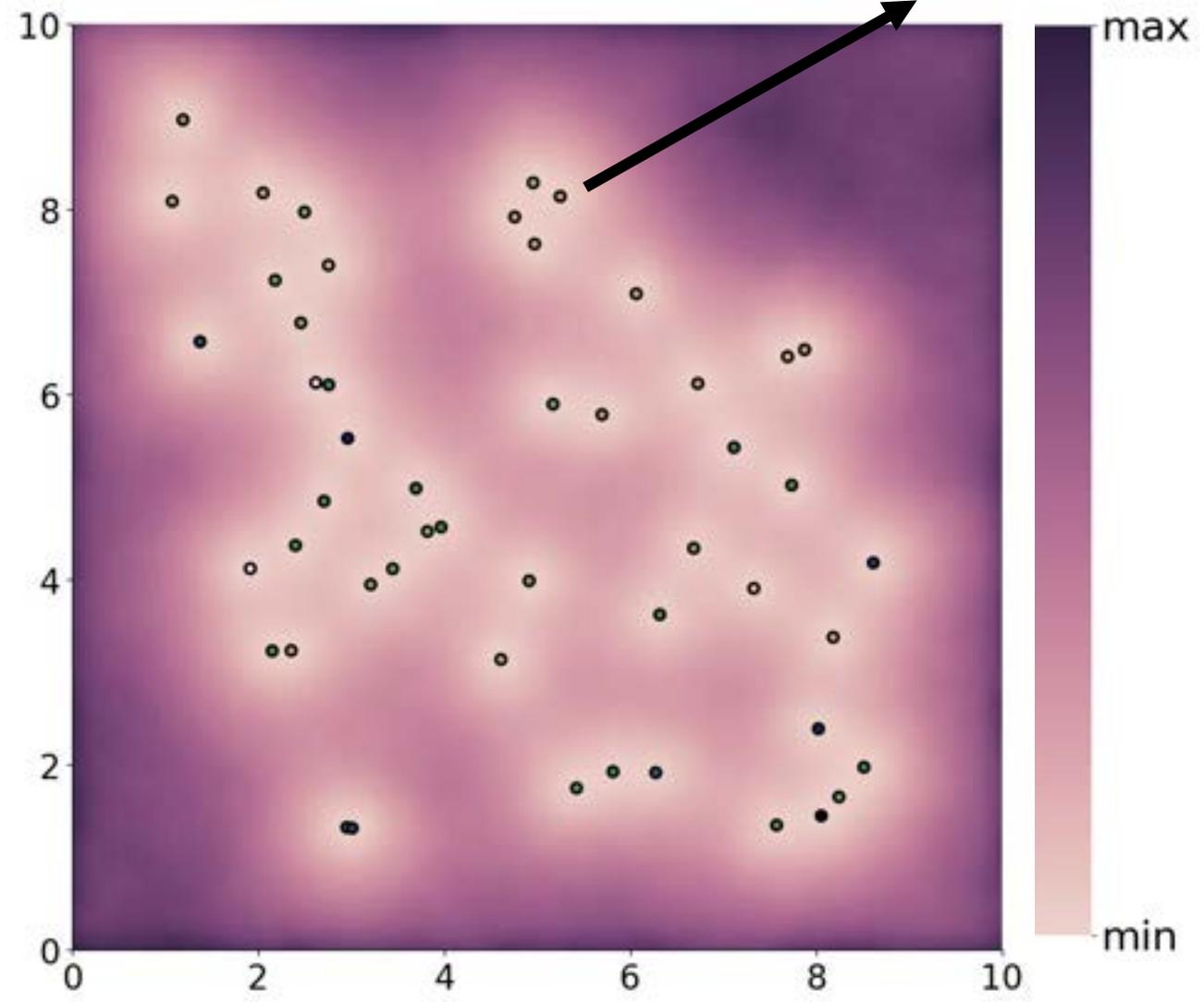


Conditional simulations with SPDE

4 conditional simulations



Variance



Increase with the distance with data

Ongoing work

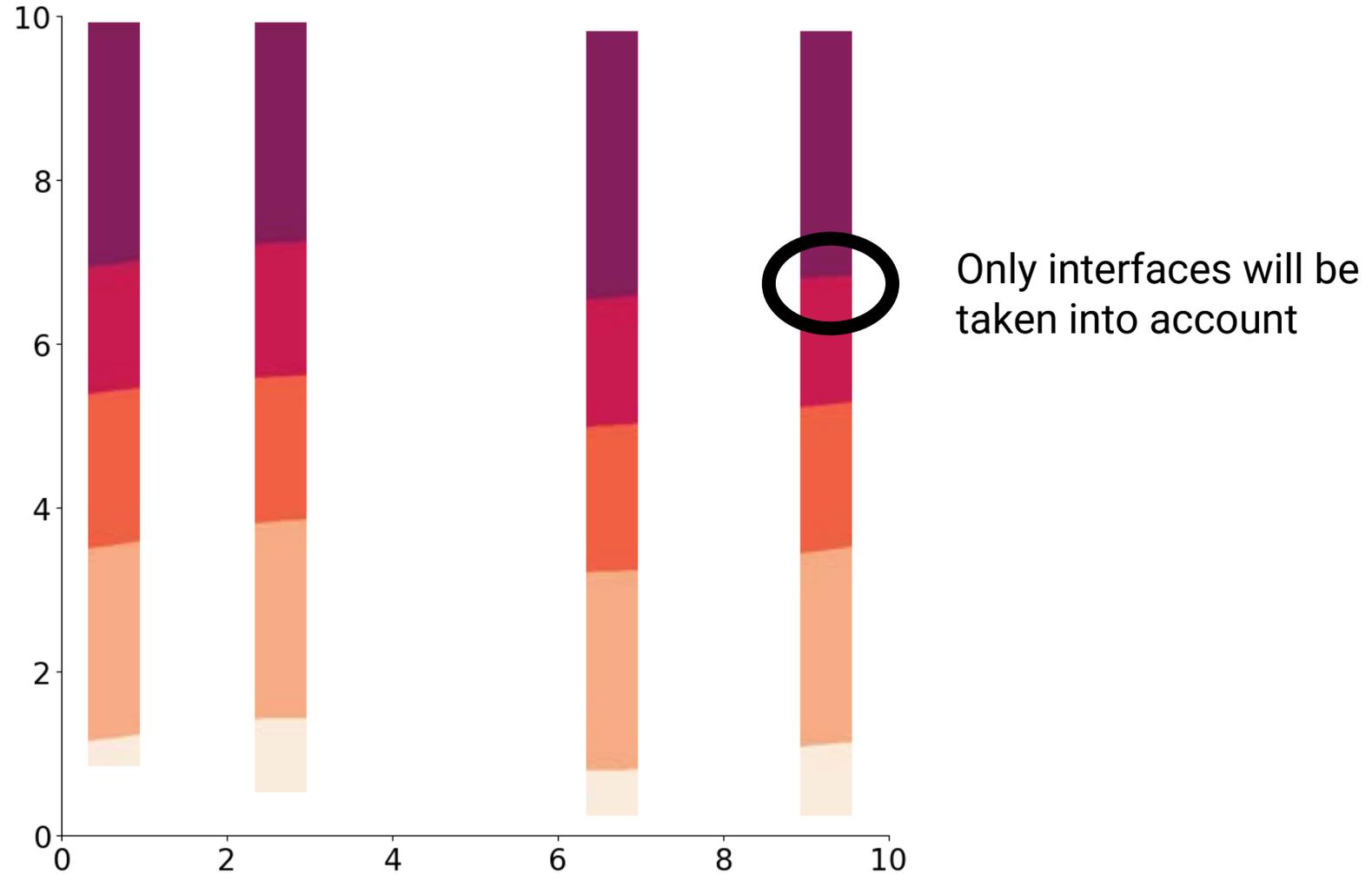
Geological modeling with SPDE, based on the potential field method (Lajaunie et al., 1997)

Inequality data $Z(\mathbf{x}) \in [A, B]$, based on Gibbs sampling (Freulon, 1992)

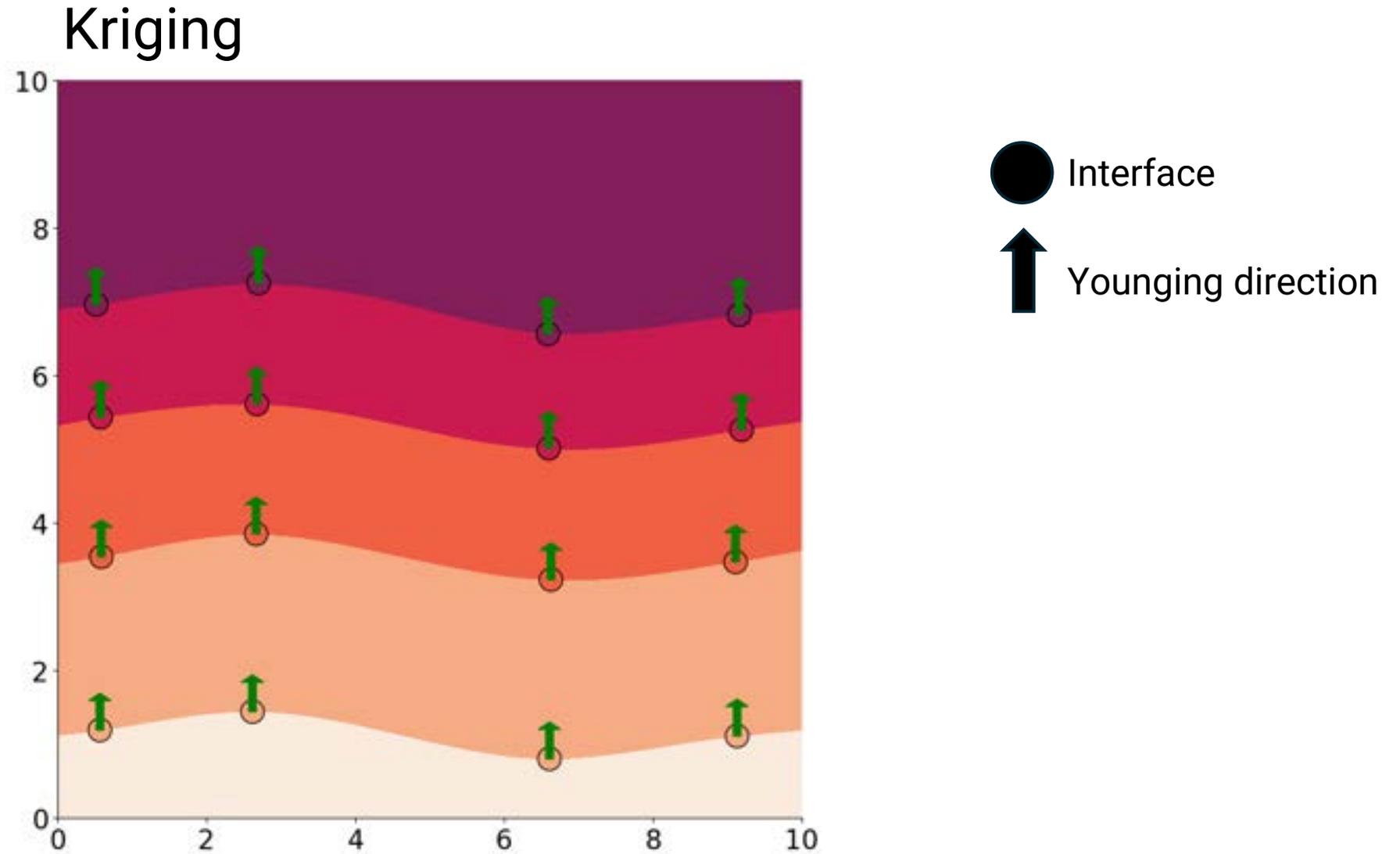
SPDE kriging and simulations with a fault (based on the work of Marechal, 1984)

Geological modeling with SPDE, based on the potential field method

Drillholes data



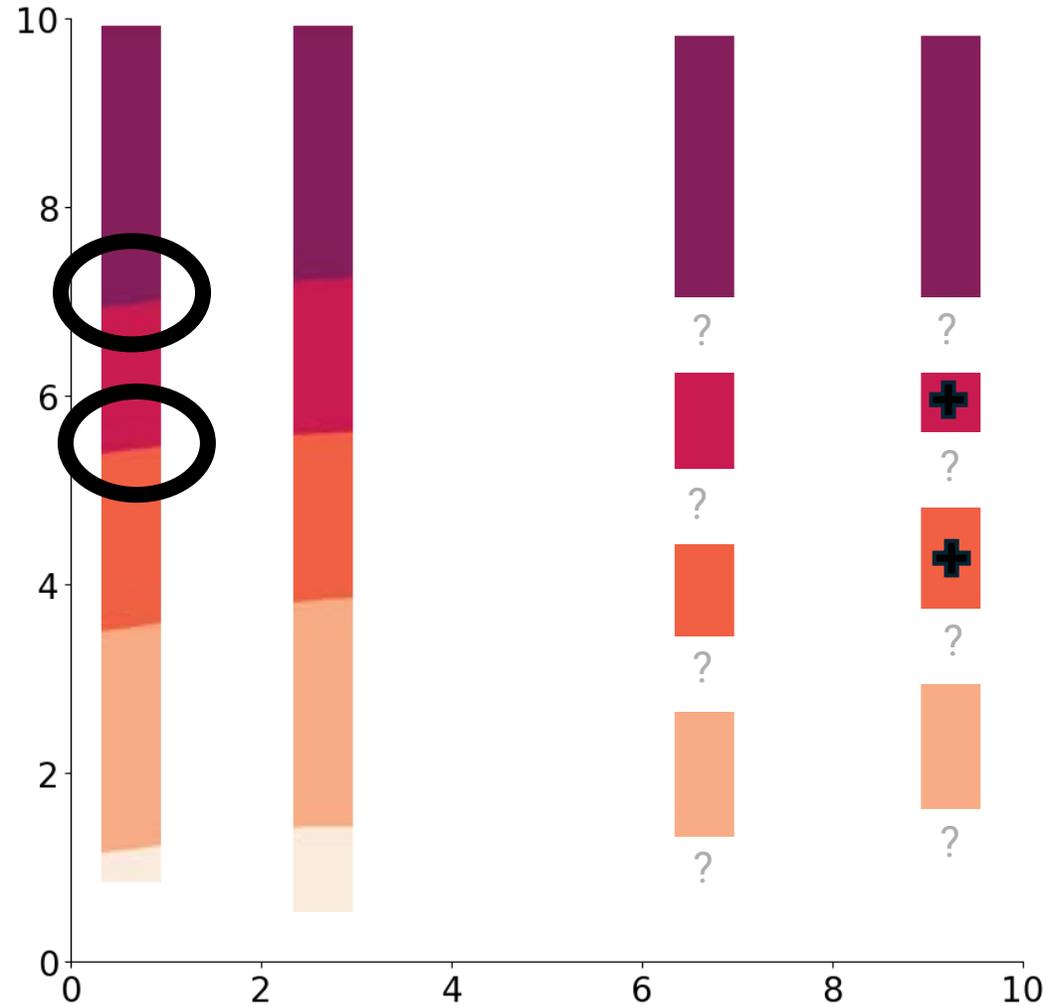
Geological modeling with SPDE, based on the potential field method



Inequality data, based on Gibbs sampling

Drillholes data with missing interfaces

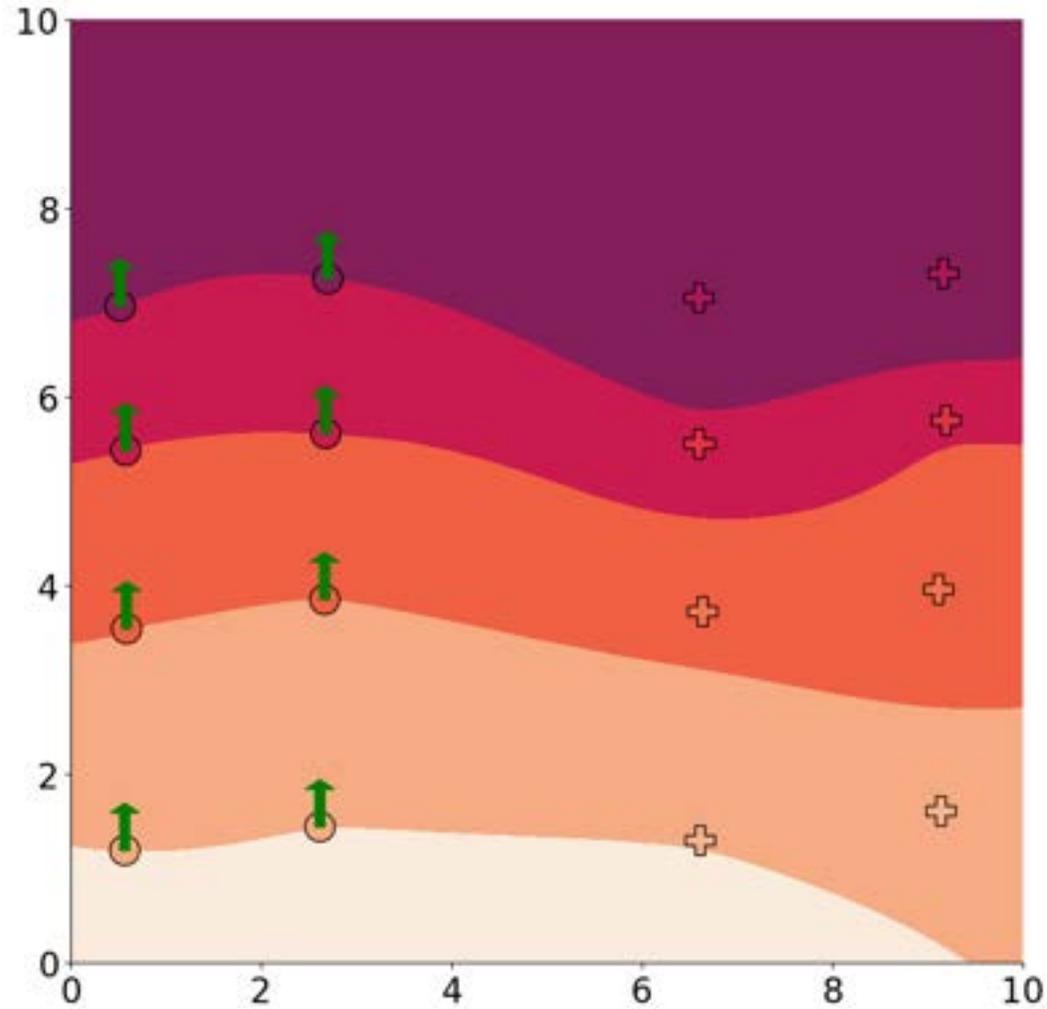
Here interfaces will be taken into account



- Here, interfaces are not observed
- But we have information of points between two interfaces " $Z(x) \in [A, B]$ "

Inequality data, based on Gibbs sampling

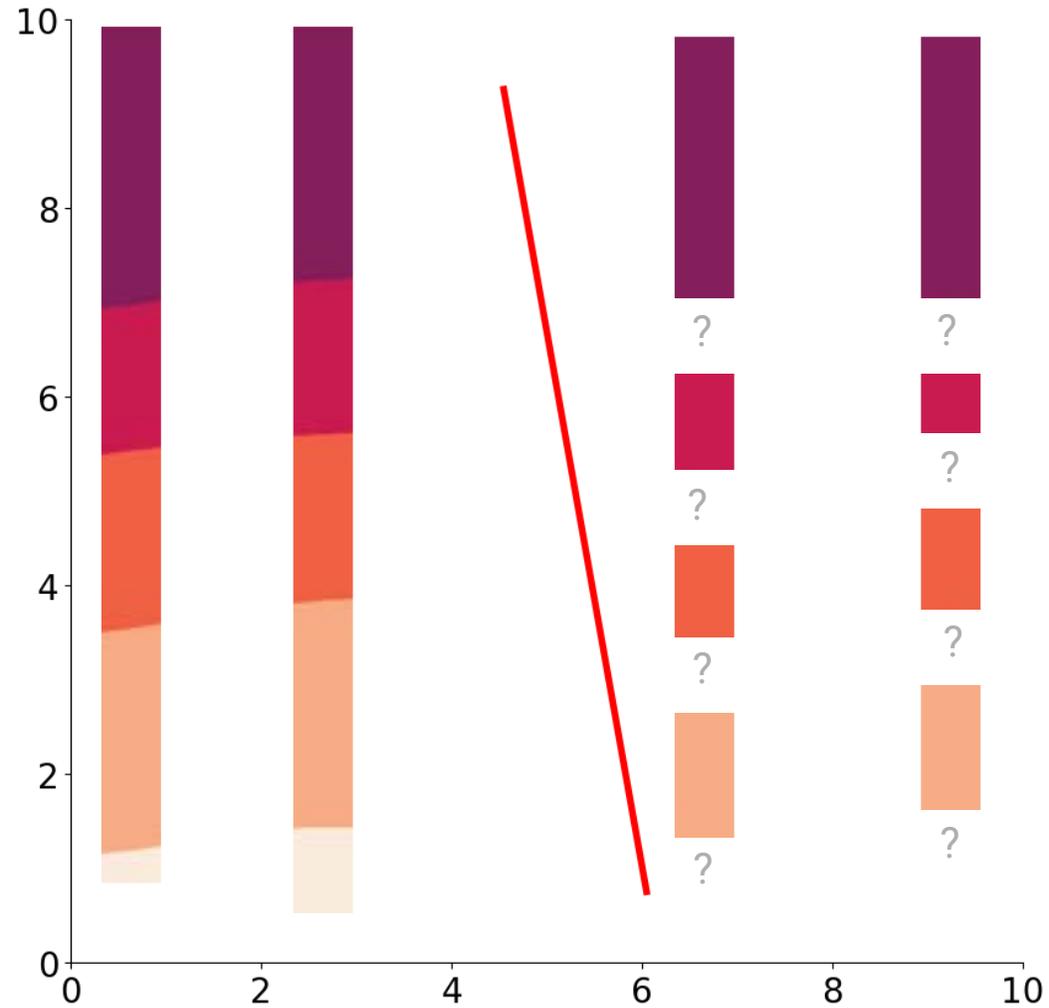
Kriging with inequality



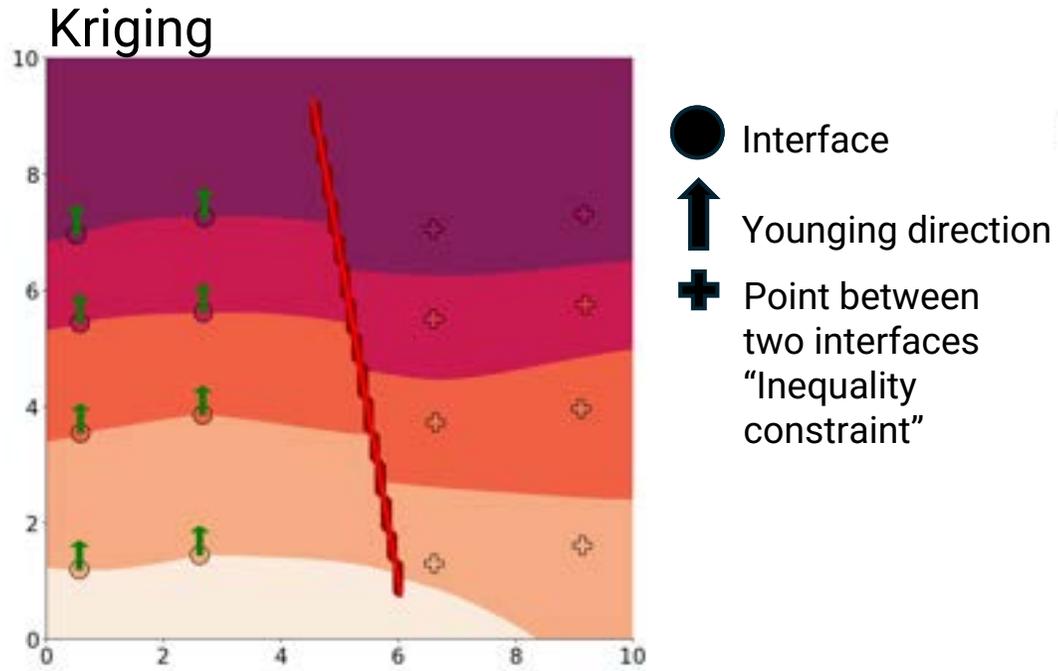
- Interface
- ↑ Younging direction
- ⊕ Point between two interfaces
“Inequality constraint”

What happens when there is a geological fault?

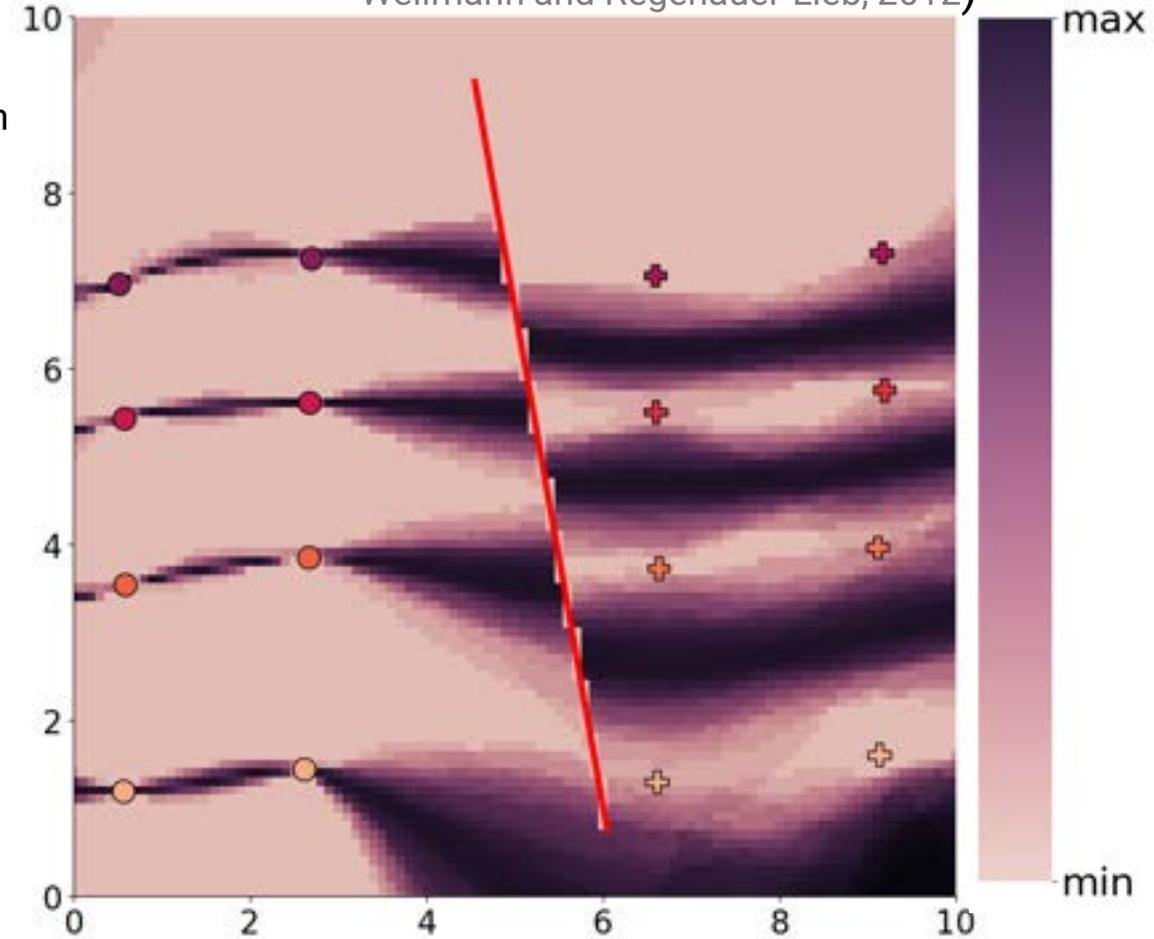
Drillholes data with a known fault



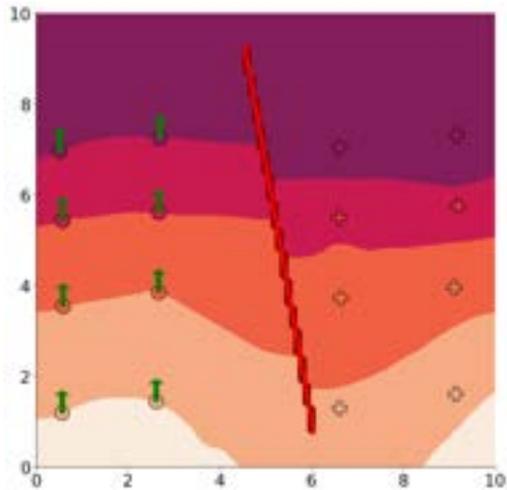
Kriging and simulations with a fault



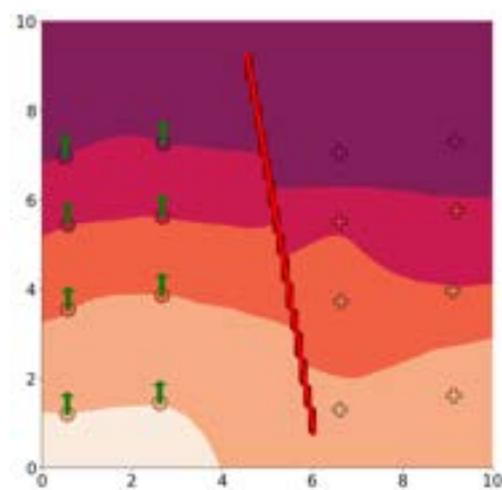
Uncertainty map (Shannon entropy,
Wellmann and Regenauer-Lieb, 2012)



Simulation 1



Simulation 2



Conclusion

- The SPDE approach is an alternative of the “classic” covariance approach for geostatistics
- The SPDE approach is faster, and computation time don't depend on the amount of data

Conclusion

- The SPDE approach is an alternative of the “classic” covariance approach for geostatistics
- The SPDE approach is faster, and computation time don't depend on the amount of data
- Ongoing works to be able to:
 - Use SPDE for geological modeling
 - Use inequality constraints with SPDE
 - Take into account the presence of faults

Conclusion

- The SPDE approach is an alternative of the “classic” covariance approach for geostatistics
- The SPDE approach is faster, and computation time don't depend on the amount of data
- Ongoing works to be able to:
 - Use SPDE for geological modeling
 - Use inequality constraints with SPDE
 - Take into account the presence of faults
- Computations done with the open-source library gstlearn (C++, R, python)

github.com/gstlearn/gstlearn

Thank you for your attention!

Charlie GARAYT^(1,2), Nicolas Desassis⁽¹⁾, Nicolas Clausolles⁽²⁾, Simon Lopez⁽²⁾

(1) Mines Paris, PSL (2) BRGM

References

- Lindgren, F., Rue, H., & Lindström, J. (2011). An Explicit Link between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 73(4), 423–498. <https://doi.org/10.1111/j.1467-9868.2011.00777.x>
- Whittle, P. (1954). ON STATIONARY PROCESSES IN THE PLANE. *Biometrika*, 41(3-4), 434-449. <https://doi.org/10.1093/biomet/41.3-4.434>
- Lajaunie, C., Courrioux, G., & Manuel, L. (1997). Foliation fields and 3D cartography in geology: Principles of a method based on potential interpolation. *Mathematical Geosciences*, 29(4), 571–584. <https://doi.org/10.1007/bf02775087>
- Freulon, X. (1992). *Conditionnement du modele gaussien par des inegalites ou des randomisees*. <http://www.theses.fr/1992ENMP0333>
- Wellmann, J. F., & Regenauer-Lieb, K. (2012). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. *Tectonophysics*, 526–529, 207–216. <https://doi.org/10.1016/j.tecto.2011.05.001>
- Marechal, A. (1984). Kriging Seismic Data in Presence of Faults. Dans *Geostatistics for Natural Resources Characterization* (p. 271-294). https://doi.org/10.1007/978-94-009-3699-7_17
- University, M. P.-. P. (2026). *gstlearn/gstlearn: Stable Release v1.10.1*. *Zenodo (CERN European Organization for Nuclear Research)*. <https://doi.org/10.5281/zenodo.18352742>