



# GeoBlocks: Bausteine zur Quantifizierung von Ungewissheiten in Geologischen Modellen

Marlise Colling Cassel, Francisca Robledo, Friedrich Carl, Carlos Colombo, Jian Yang

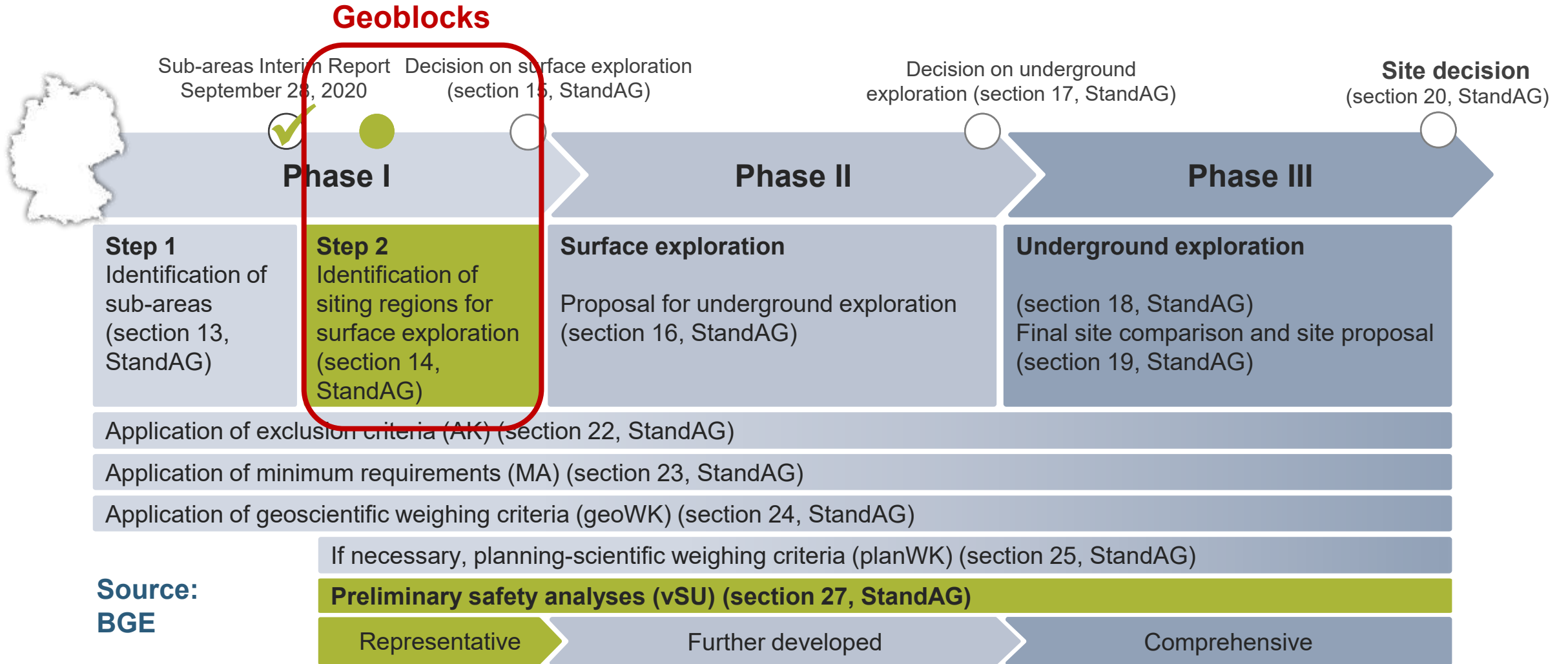
**PIs:** Florian Wellmann, Peter Achtziger, Florian Amann, Peter Kukla, Clare Bond, Stefan Back

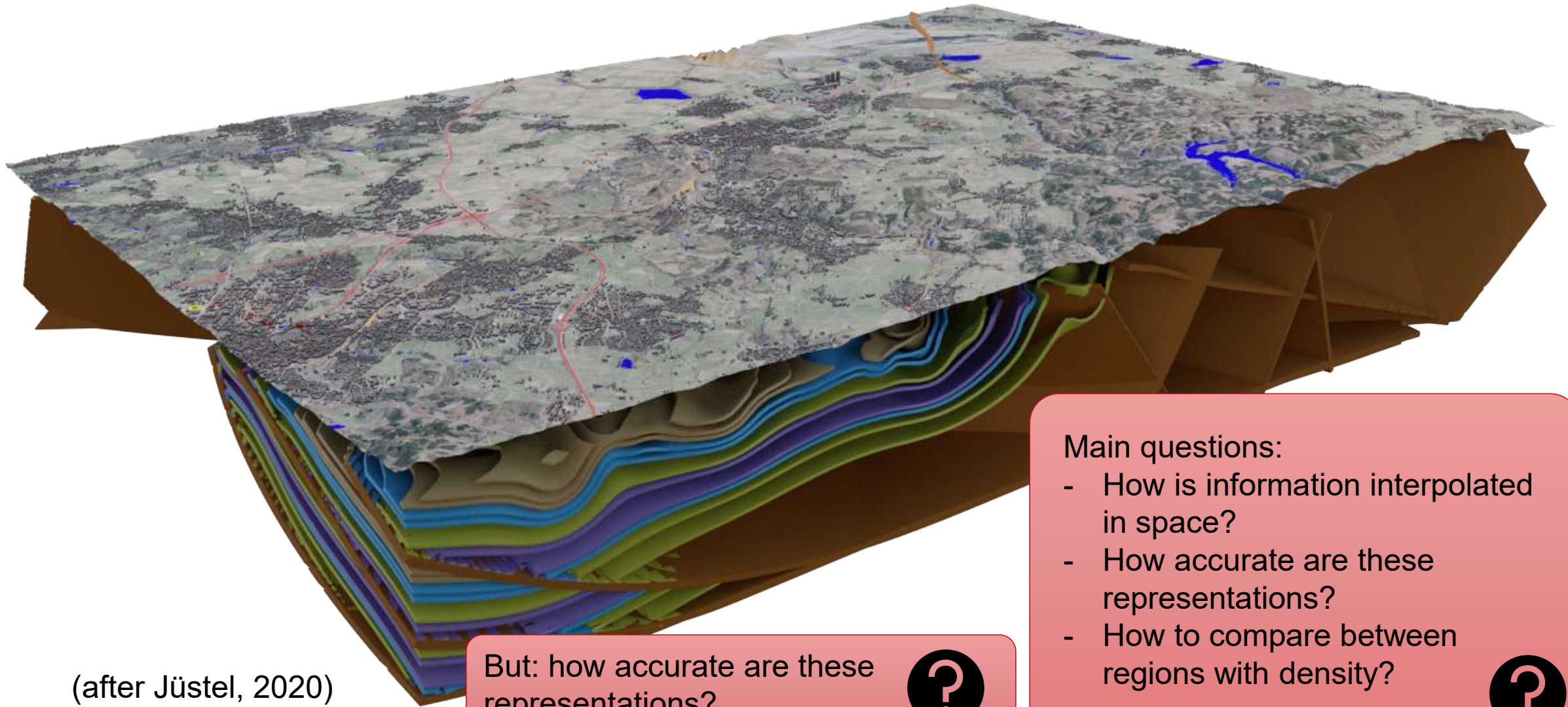
**Associated Partners:**

Frithjof Bense, Fabian Jähne-Klingberg, Heidrun Stück, Björn Zehner<sup>†</sup> (BGR), Michael Kettermann (IEG)

URS Workshop – Potsdam 05.02.-07.02.2025

# Where does the project fit in?





(after Jüstel, 2020)

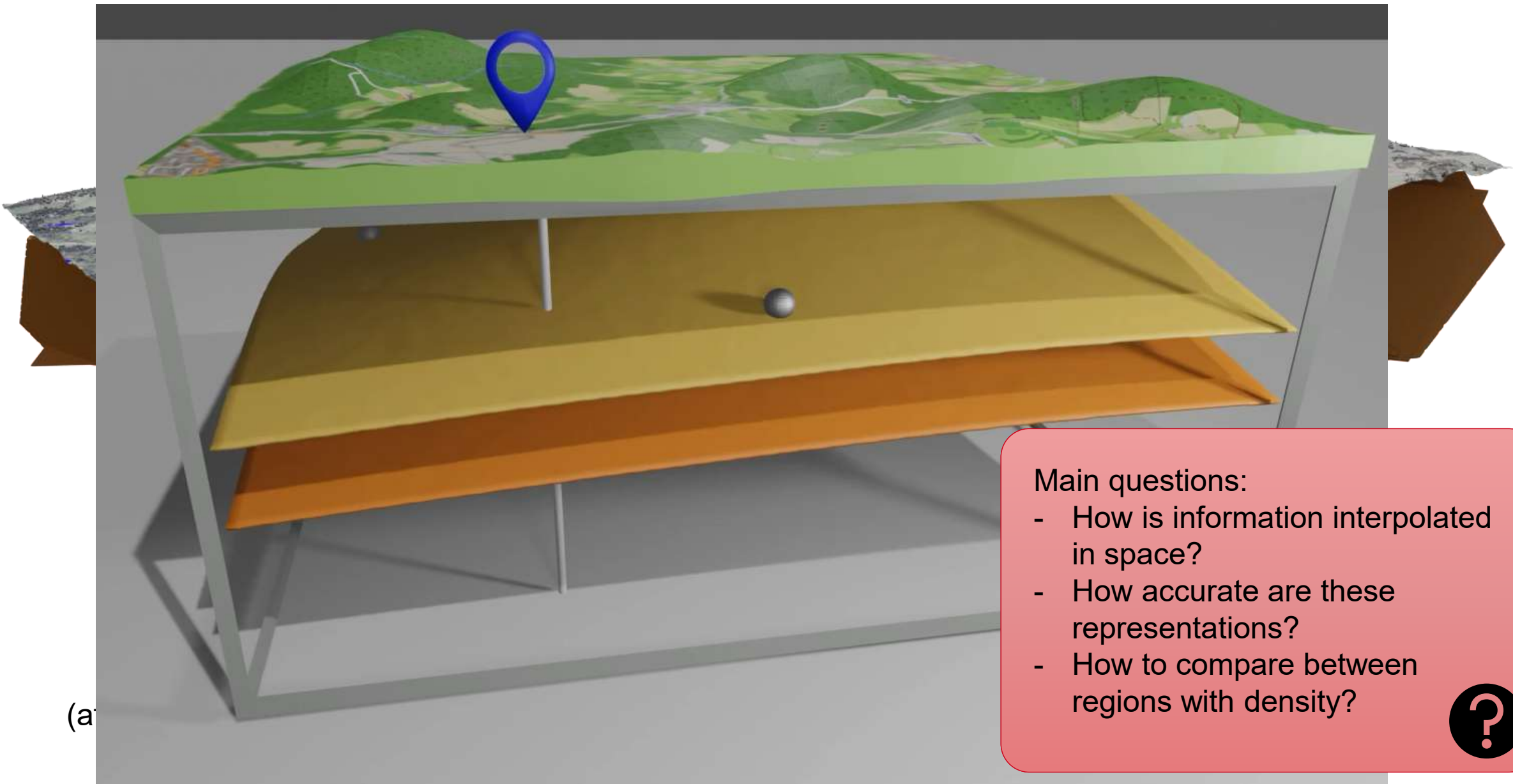
But: how accurate are these representations?



Main questions:

- How is information interpolated in space?
- How accurate are these representations?
- How to compare between regions with density?





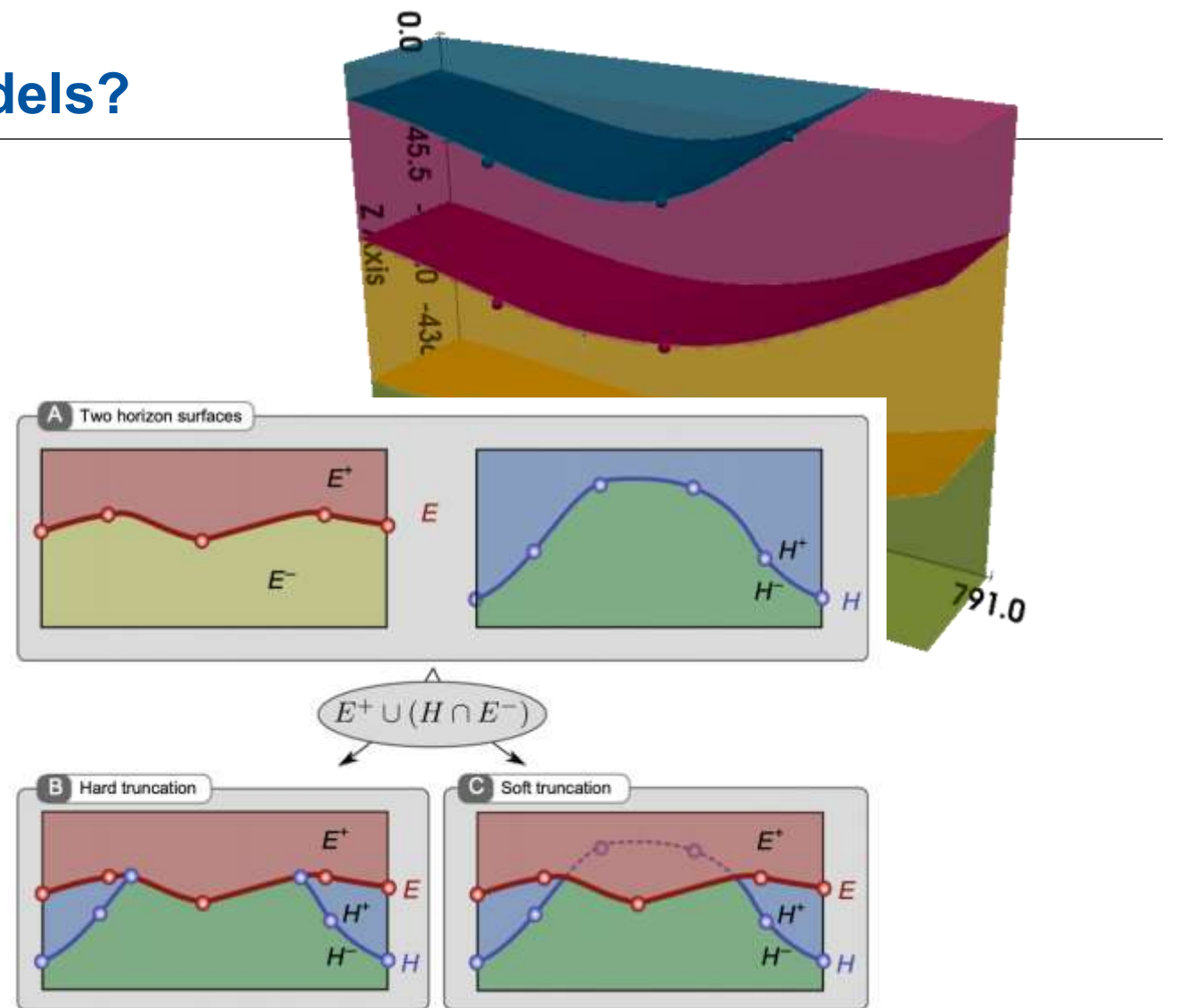
Main questions:

- How is information interpolated in space?
- How accurate are these representations?
- How to compare between regions with density?



# How to construct 3-D geometric models?

- Explicit and implicit geometry representations
- Aim: construct (reasonably) complex 3D geometric representations with a small number of input points => **low-dimensional (geometric) parameterizations**
- Many **tools available** with purpose-built GUIs and sophisticated workflows – but (partly) **at expense of transparency**
- More geological settings can be generated through a **combination of multiple scalar fields** (boolean logic)
- Here: methods implemented in **open-source geological modeling package GemPy**



# How to construct 3-D geometric models?



- Exp
- Aim
- geo
- num
- (ge
- Ma
- GU
- (pa
- Mo
- thro
- fiel
- Her
- geo



## GemPy

<https://www.gempy.org>

**GemPy:** Flexible implicit geological modeling and uncertainty quantification

**Open-source geomodeling**

([www.gempy.org](http://www.gempy.org)), (de la Varga, Schaaf & Wellmann, 2019)



♡ Sponsor

📌 Edit Pins ▾

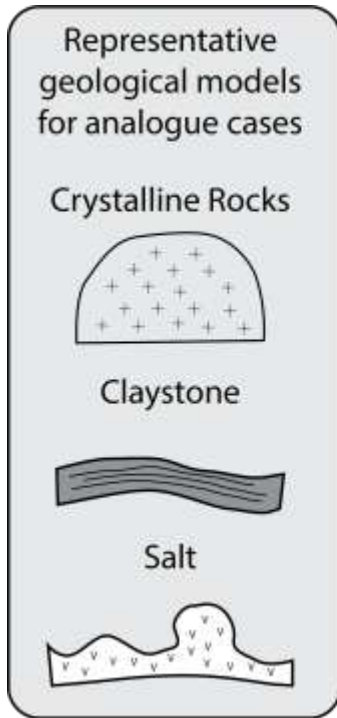
👁️ Unwatch 52 ▾

🍴 Fork 227 ▾

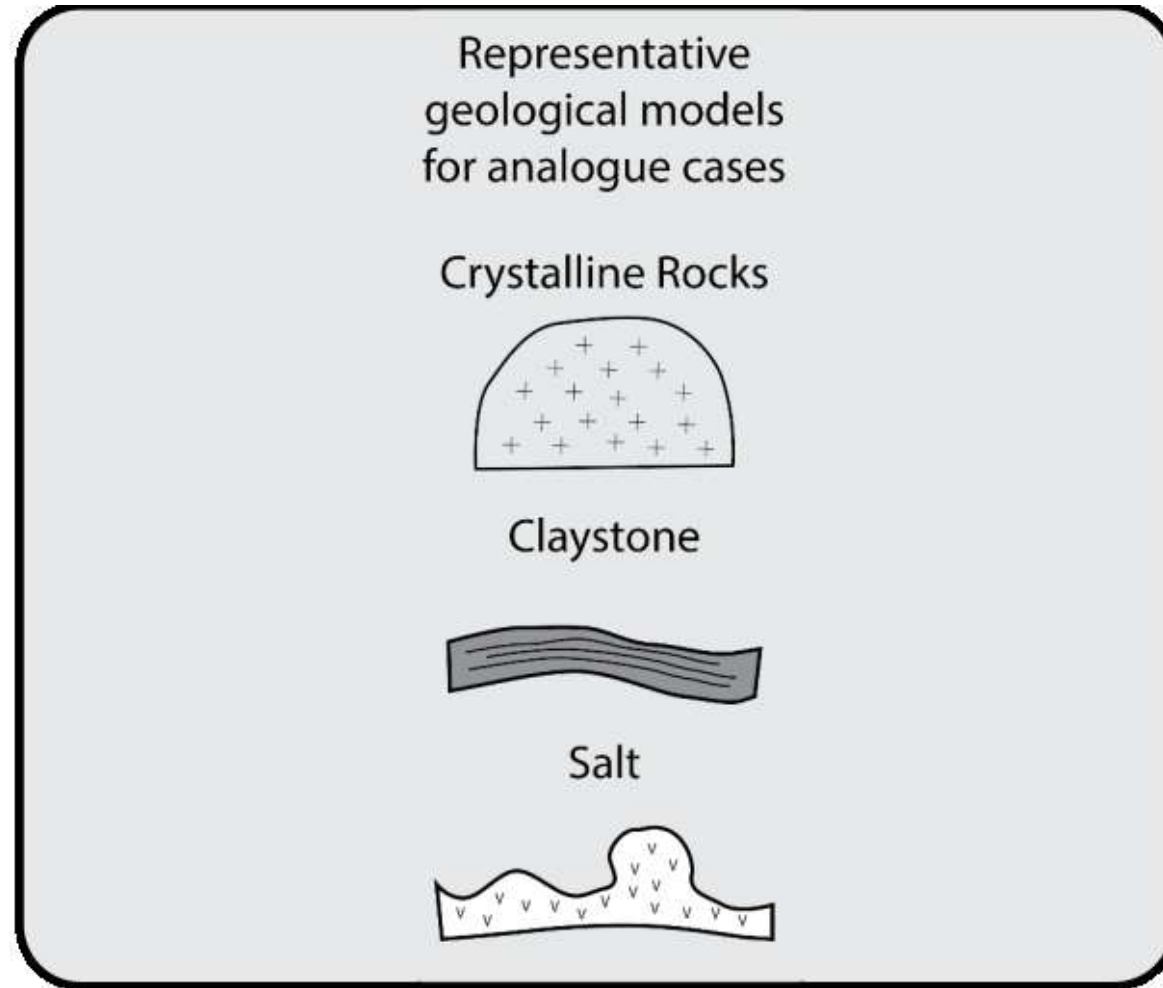
★ Starred 886 ▾

# Geoblocks project overview

---



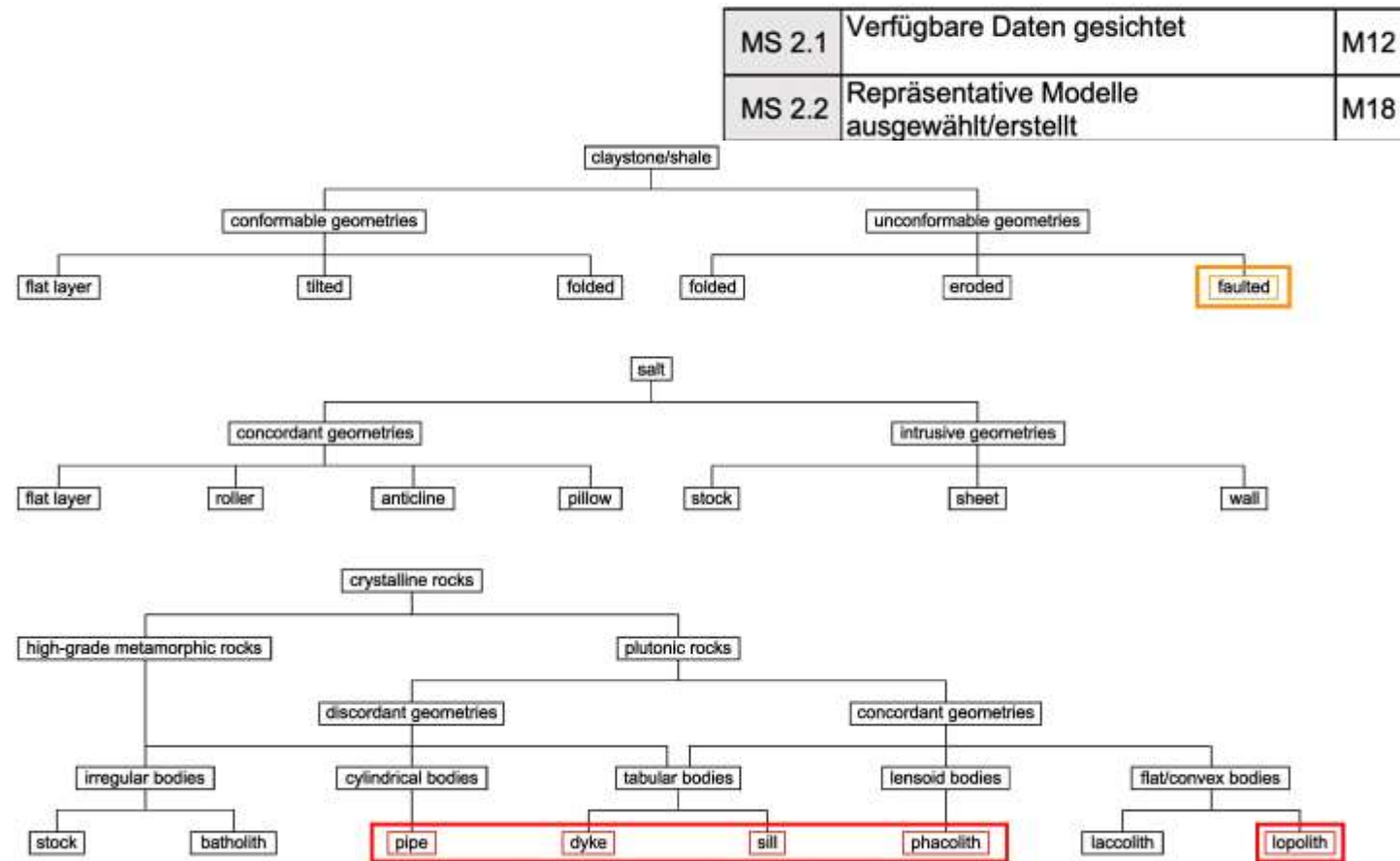
# Part 1: Representative models and quantitative comparison of explicit geomodels





# Library of Geometries: Geometrical Systematization of Geological Bodies

- Workflow for geomodelling has to be applicable to the entire geologically complex range of potential host rocks
- Different geological-tectonic settings and input data sets are available for different parts of the German subsurface
- Potential host rocks show large range of subsurface geometries
- Geological-geometrical systematization into end members for all host rocks





- Based on **geometrical systematization**, using standard geometries for visualization
- Additional subdivision of structural types based on research on German subsurface conditions and **relevance for site selection process**
- **Work in progress**

Examples based on TUNB model, BGR et al. (2022)

Stocks

Walls



cylindrical

hourglass-shape

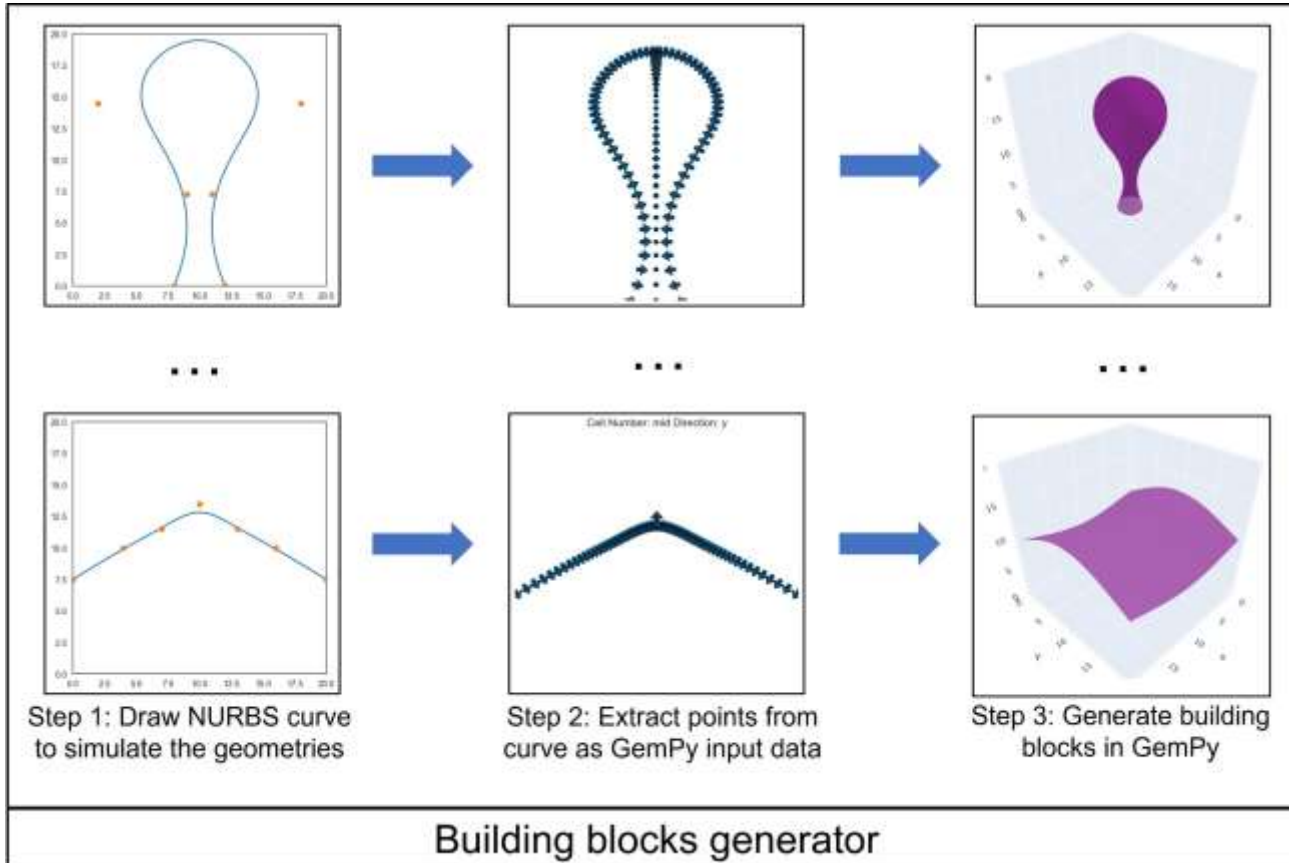
moderately anisotropic

highly anisotropic

# Standard geometries and model generator V1.0

MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18
--------	--------------------------------------------	-----

## GemPy model generator V1.0



- Geometrical systematization is basis for 3D regular geomodels representing geometrical end members
- 2 model generator versions for creation of standard geometries

### Fundamentals V1:

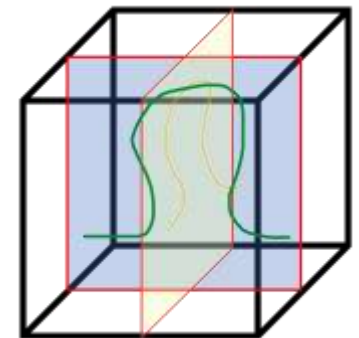
- Use NURBS curves (defined by control points) to interactively set up two orthogonal 2D sections
- interpolating in 3D using implicit surface interpolation method implemented in GemPy

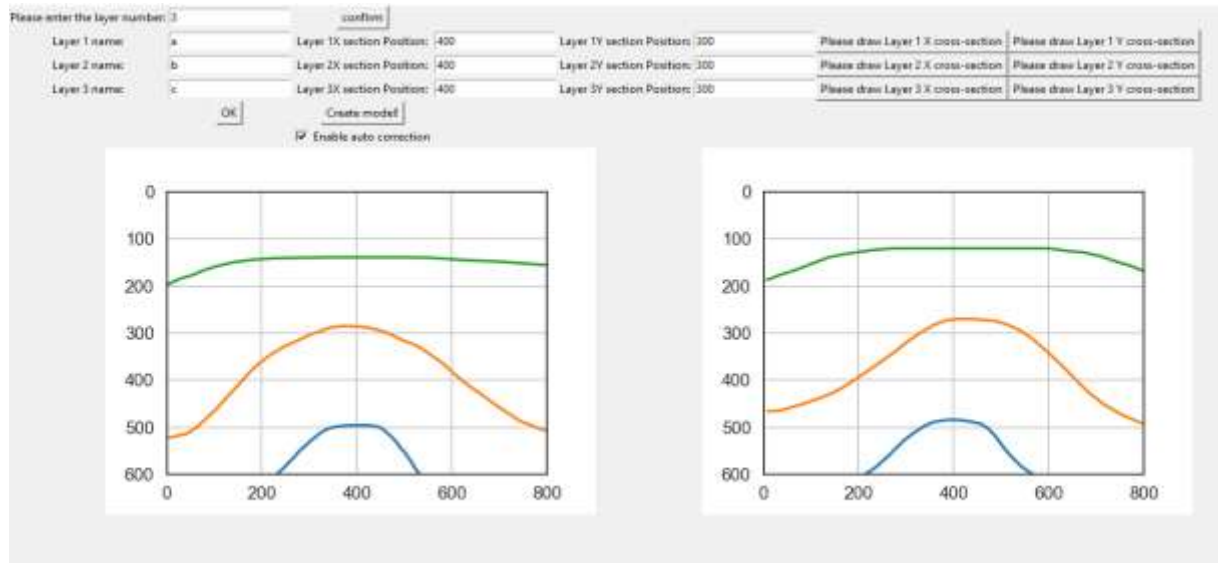
### Advantages:

- Only 7 points to manipulate
- Easy to build a regular geometry

### Disadvantages:

- Only allows one layer and two cross-sections
- Cannot be used to build complex geometries





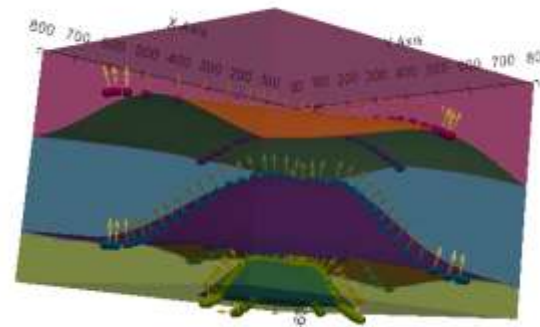
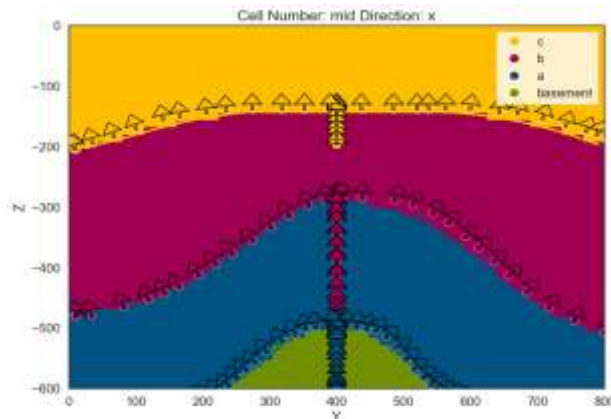
## GemPy model generator V2.0

### Fundamentals:

- Use drawing curve to define the geometry
- Sample points from curve as import

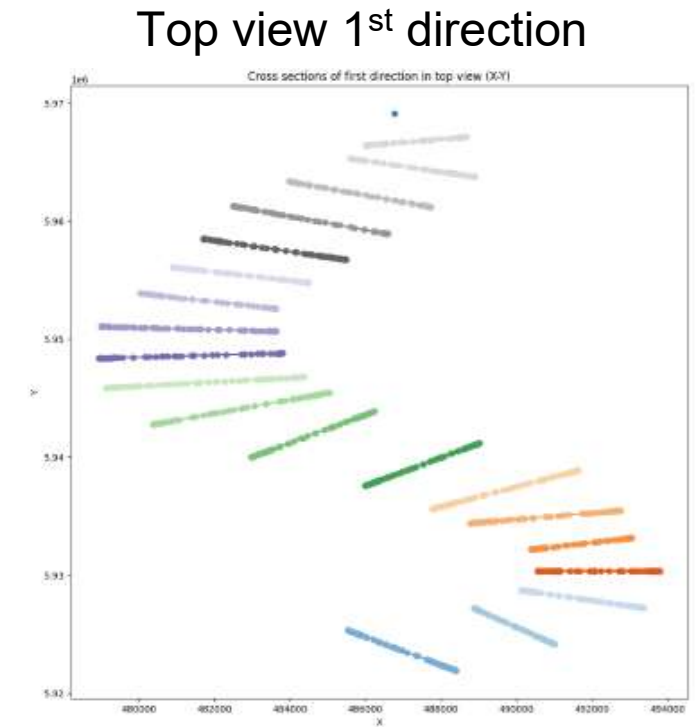
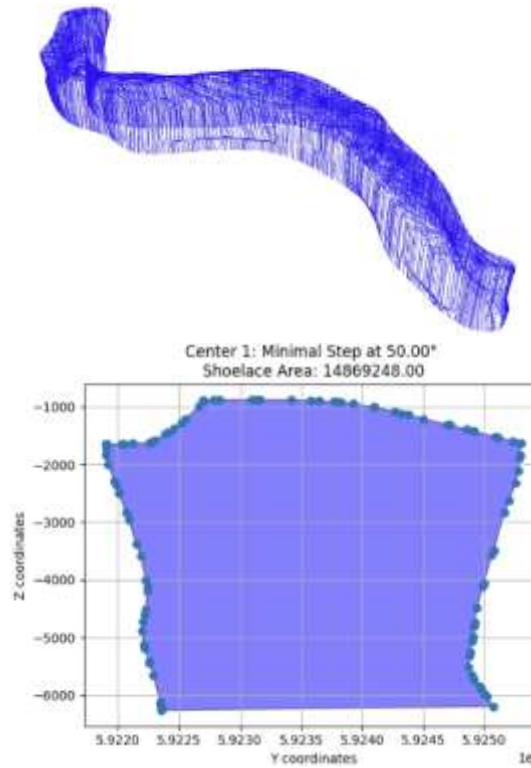
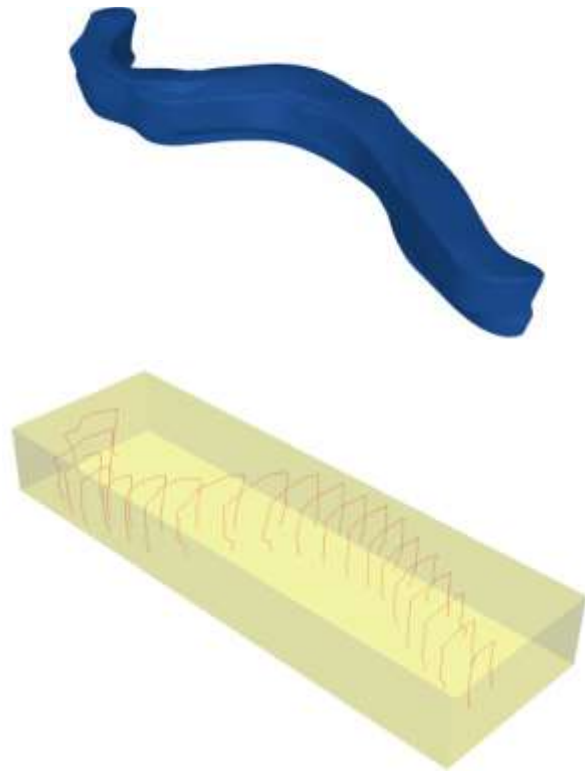
### Advantages:

- Allows multiple layers and cross-sections
- Easy to build complex geometry
- A background image can be used as reference for drawing

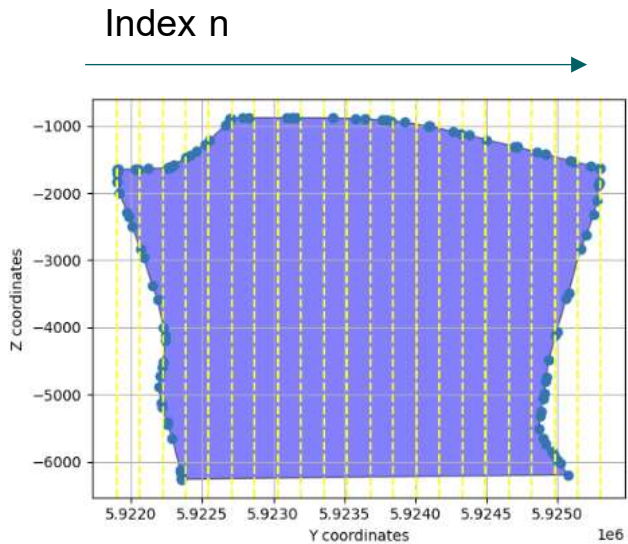


# Quantitative comparison of structures using geometrical/ statistical properties

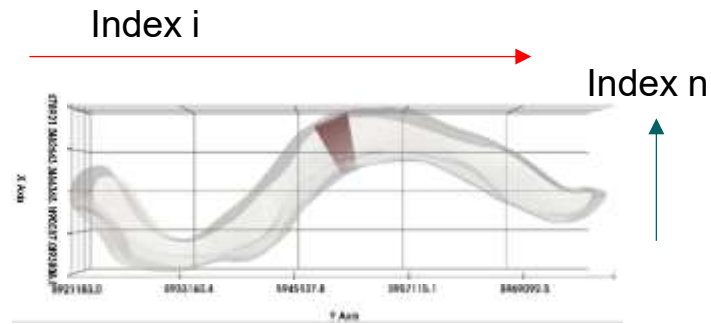
- Question: **how to describe and compare geological structures** quantitatively?
- Measurement of lateral & vertical dimensions and gradient & curvature data on explicit geomodels
- Data analysis can reproduce main geometrical characteristics of input data



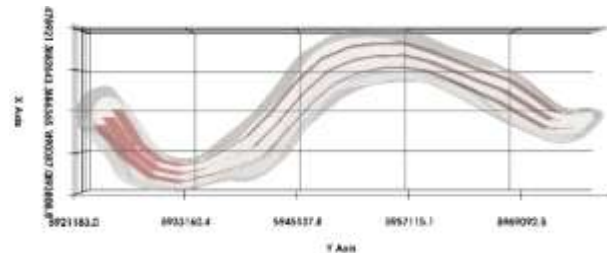
# Determination of orthogonal sections



1. Divide the sections of the first direction (index=i) into  $n=22$  vertical lines  
Extract X, Y and Zmin +Zmax from each line  $\rightarrow$  2 points



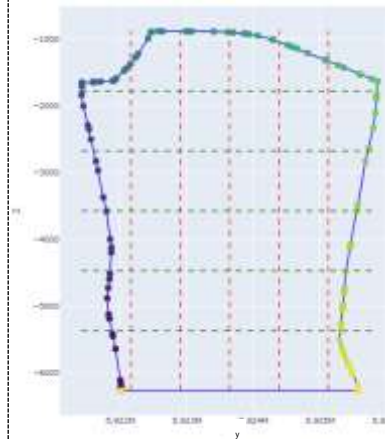
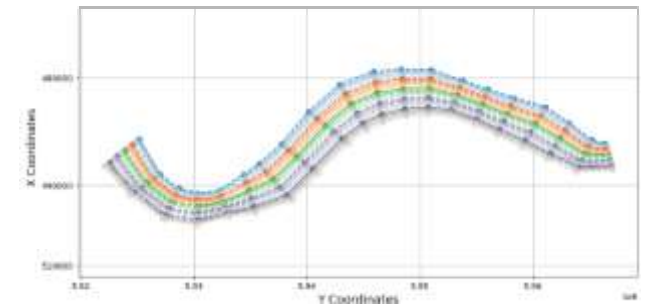
2. Create trapezoidal segments (one from overall 4 points of 2 consecutive sections)



3. Combine trapezoidal segments of given line index=n to form orthogonal section

# Dimensional measurements

Position of orthogonal sections & overall coverage of input mesh



Y-Range at specified Z-values:

```

z = -5368.179341634134: y-range = 5921875.43824438 to 5925584.548996711
z = -5368.179341634134: y-length = 3429.1107619113334
z = -4479.188888268229: y-range = 5921858.2284031815 to 5925588.562831321
z = -4479.188888268229: y-length = 358.341938239833
z = -3572.1942749023438: y-range = 5911791.888266885 to 5925546.478232308
z = -3572.1942749023438: y-length = 3752.777966118297
z = -2674.281741536458: y-range = 5921881.780118897 to 5925728.1128450245
z = -2674.281741536458: y-length = 4118.411928027528
z = -1776.2892881785728: y-range = 5921421.335883854 to 5925801.456872417
z = -1776.2892881785728: y-length = 4428.124788761882

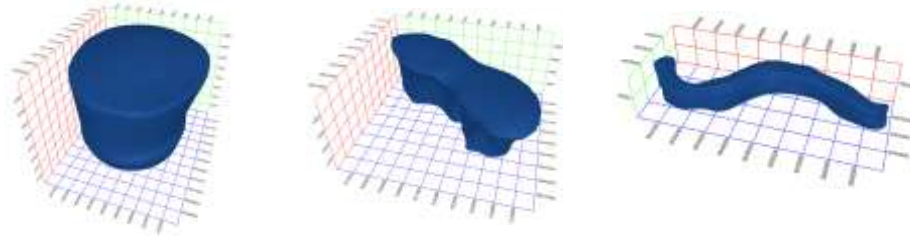
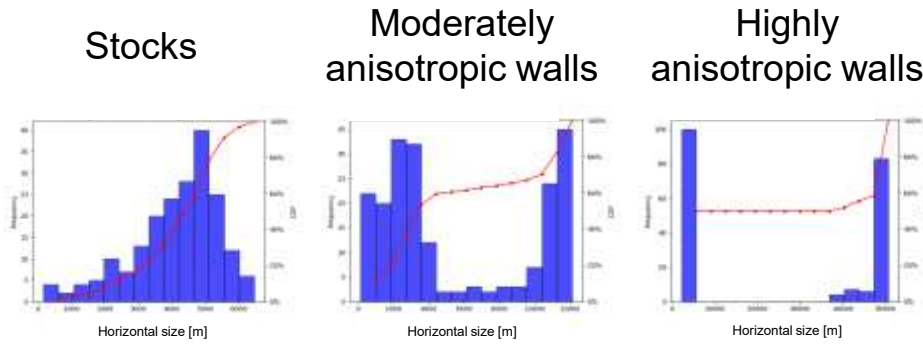
Z-Range at specified Y-values:
y = 5922160.646616397: z-range = -6266.171875 to -1383.852328889313
y = 5922160.646616397: z-length = 4882.319322518899
y = 5922981.1421854385: z-range = -6266.171875 to -883.888151472159
y = 5922981.1421854385: z-length = 5184.313853852784
y = 5923641.6177544245: z-range = -6266.171875 to -895.386838346129
y = 5923641.6177544245: z-length = 5378.775871853187
y = 5924382.13332838: z-range = -6266.171875 to -1833.8838225118806
y = 5924382.13332838: z-length = 5128.37878448891
y = 5925122.6288228515: z-range = -6266.171875 to -1128.1887527755791
y = 5925122.6288228515: z-length = 4937.87322228421
    
```

Measurement results: automatic determination of dimensions

# Data analysis: Comparison and systematization of structures and datasets

## 1) Distributions of combined horizontal dimensional data

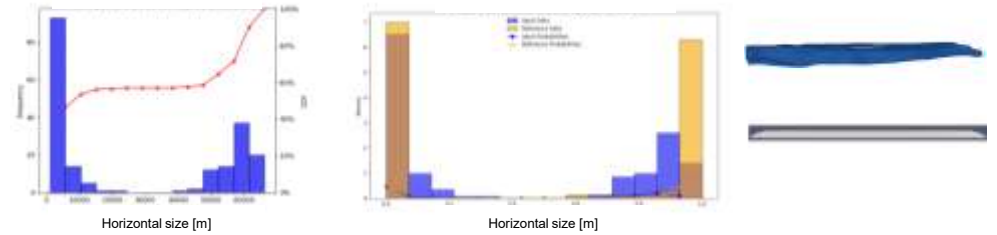
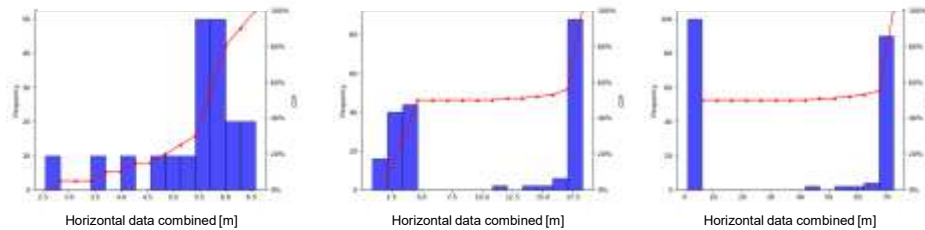
Distribution of horizontal size for selected TUNB models



A)



Distribution of horizontal size for standard geometries from the library



Kullback-Leibler divergence allows to quantify the similarity of a certain body (or slice, e.g. derived from seismics) with the bodies of the library of geometries and select the closest

B)

<2

2-6

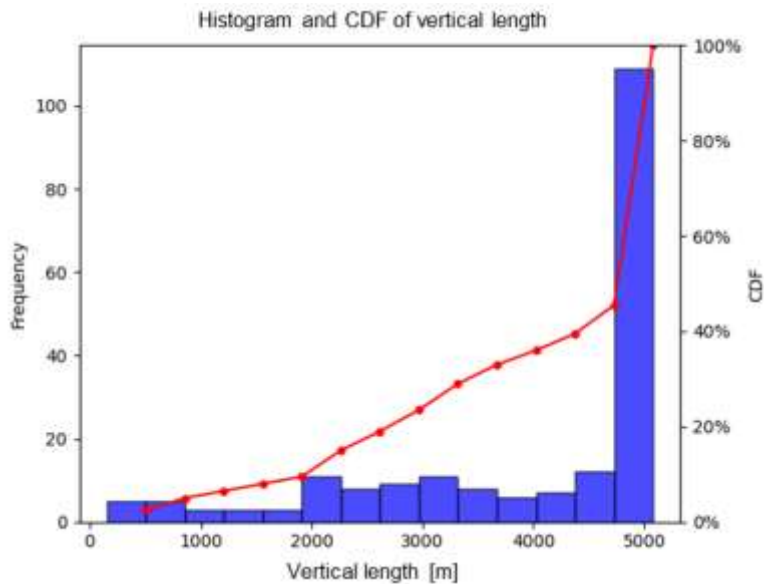
>6



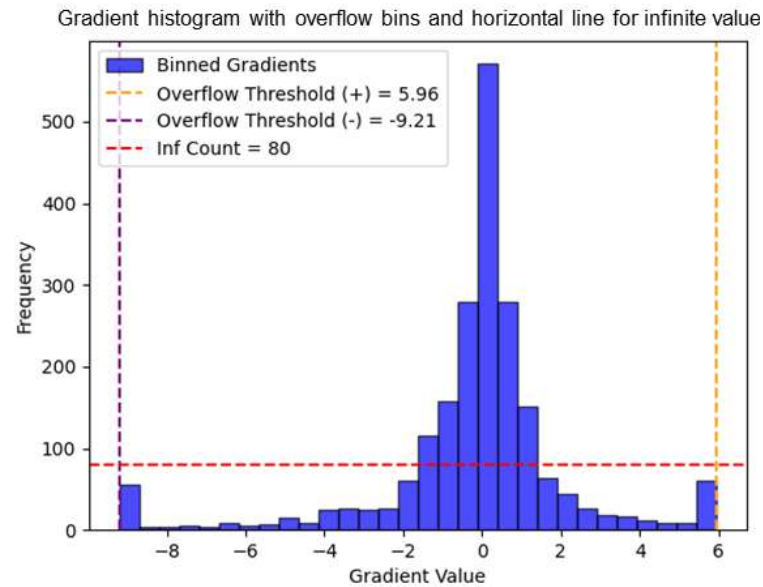
Averaged anisotropy:  
ratio between mean of data of orthogonal direction & mean of data of 1<sup>st</sup> direction

## 2) Histograms of vertical data and gradient & curvature data

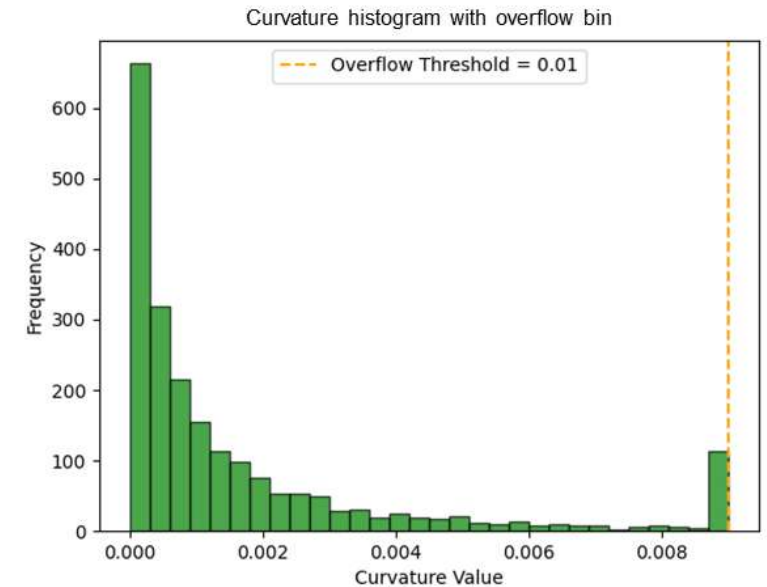
Typical distribution of vertical data



Example of gradient data

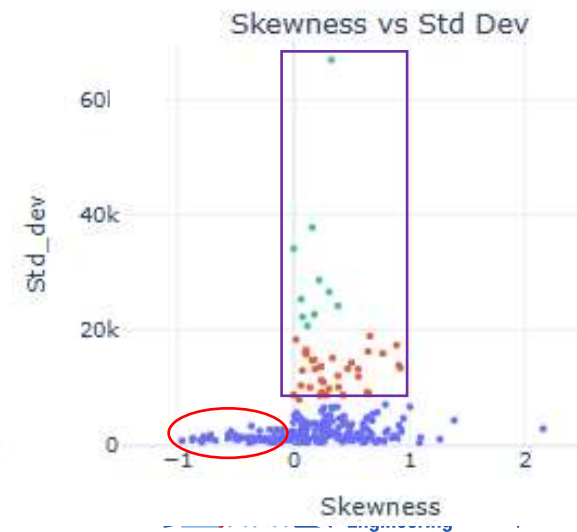
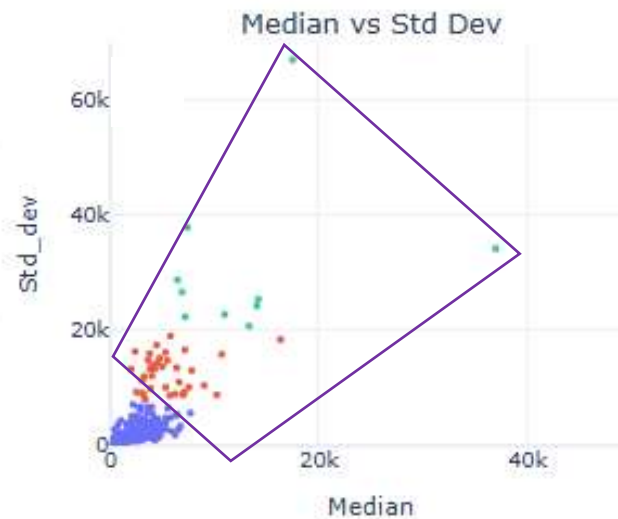
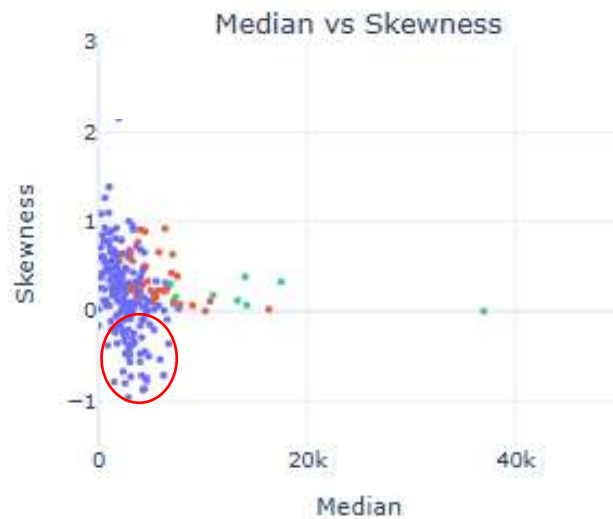
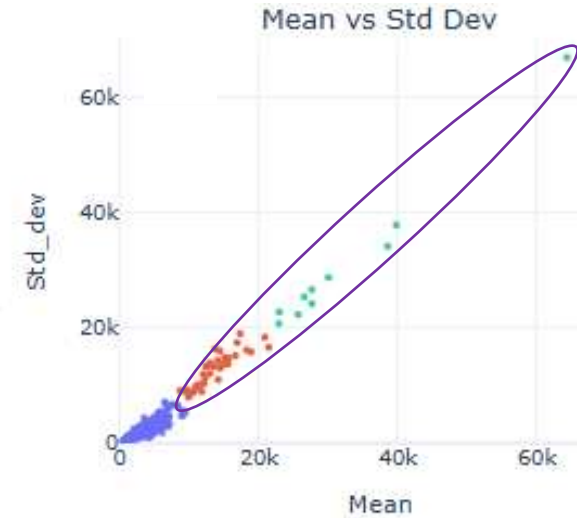
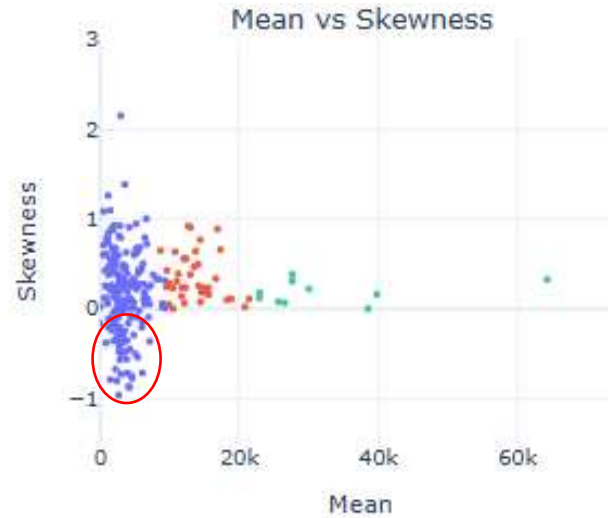
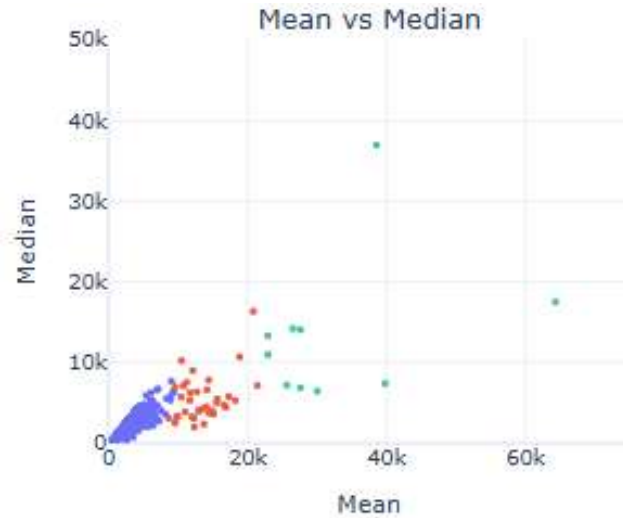


Example of curvature data



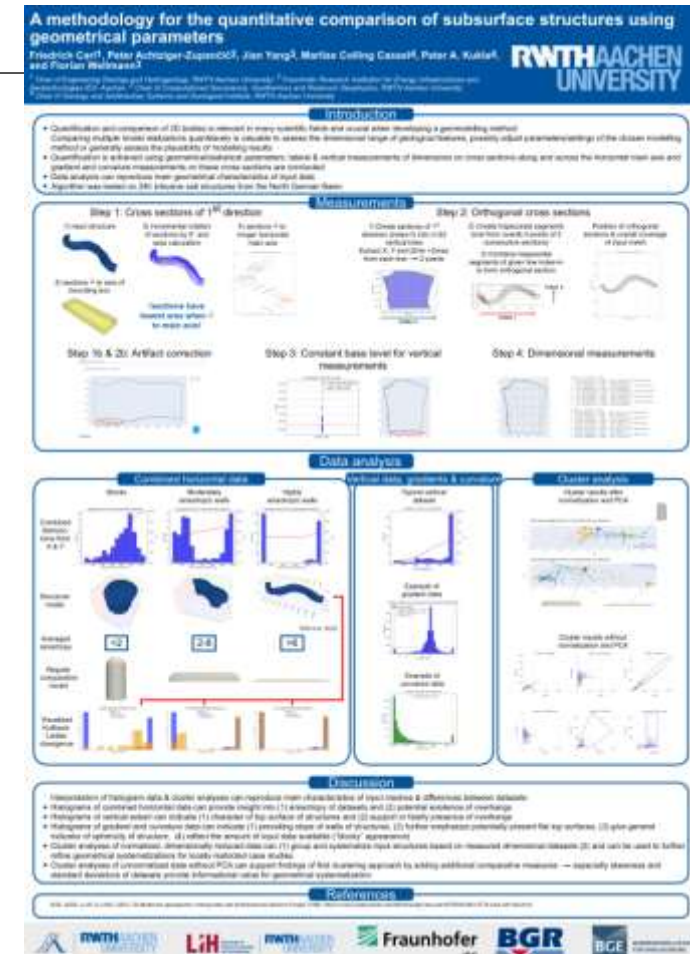


# Cluster analysis on statistical moments (Example: horizontal data)



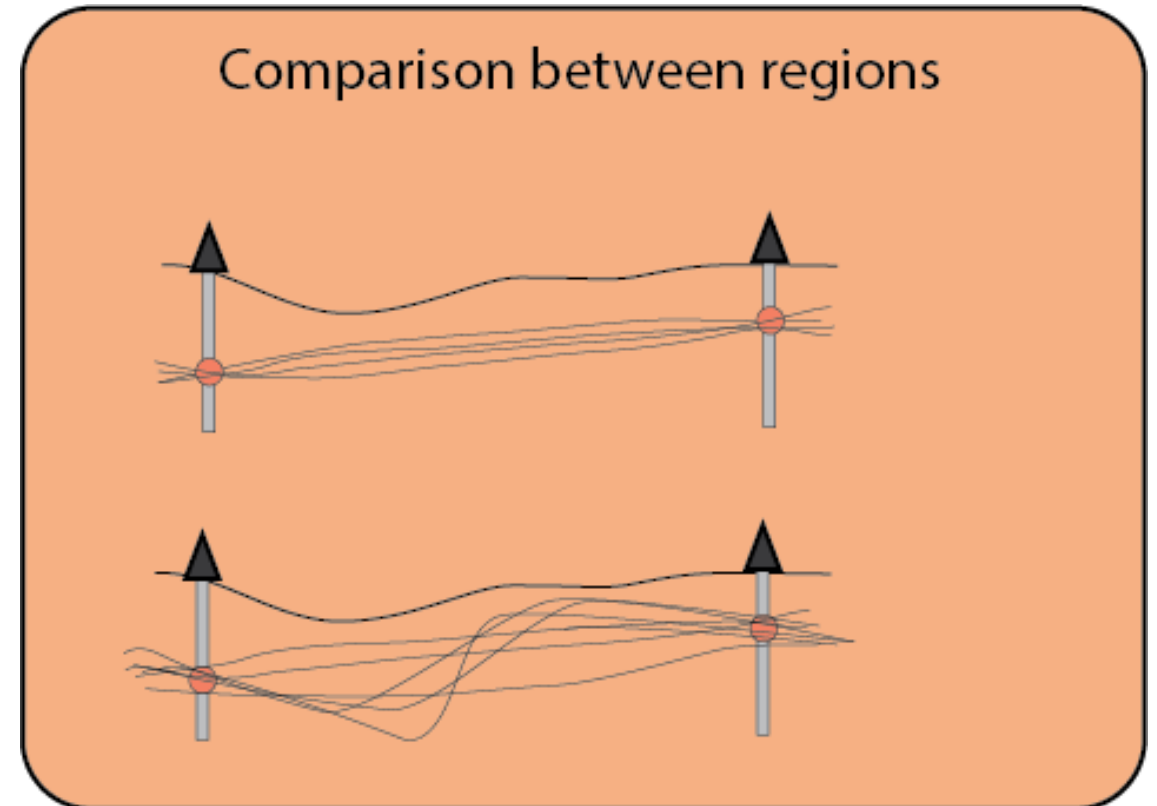
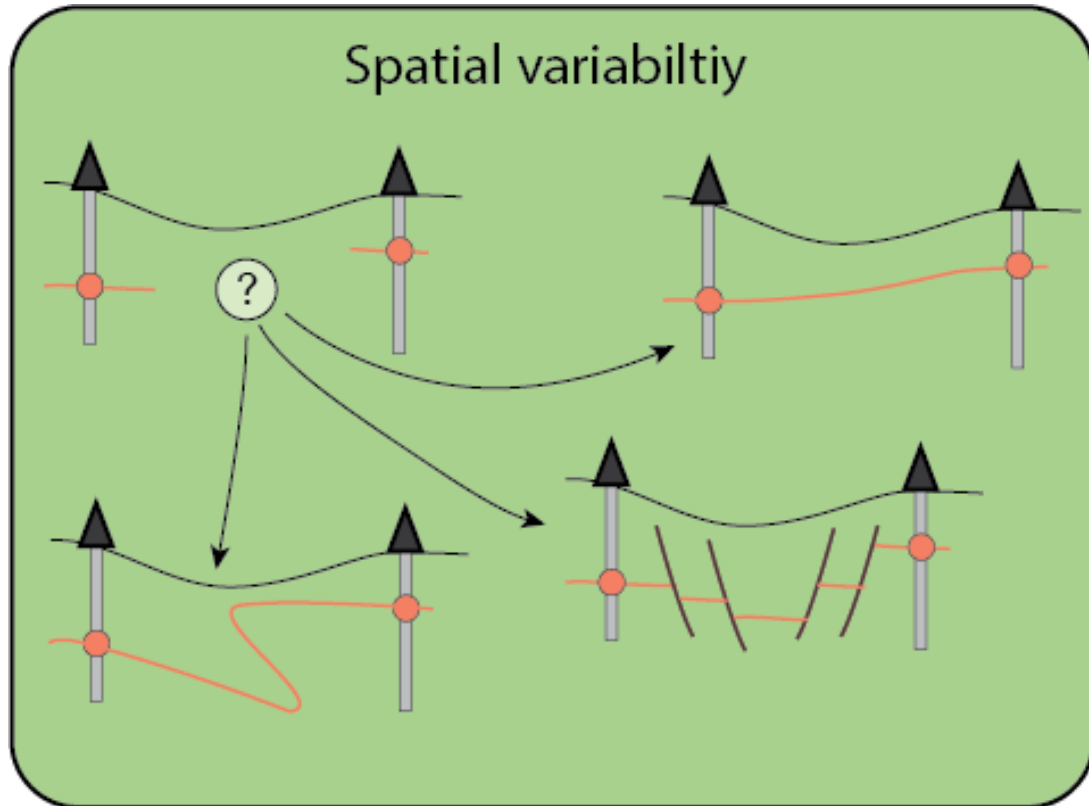
# Classification of geometries: intermediate result

- Interpretation of **histogram data & cluster analyses** can reproduce main characteristics of input meshes & differences between datasets:
  - Statistics of “simple” geometrical properties **allows to describe each salt body** and carries information on: anisotropy of the model, existence of overhangs, character of top (→dissolution), prevailing slope of walls, sphericity
  - The **distribution of the geometrical properties** allows to classify models by e.g. clustering and select the most likely standard geometry (-ies)
  - It allows to test **how much data** (e.g. shapes from seismics or thickness from borelogs) **is needed to identify a structure** and select the appropriate model from the library of geometries and might restrict the suit of hyperparameters used for modeling
  - The distribution allows simple falsification of prior models



MS 2.3	Testdatensätze extrahiert	M18
MS 4.2	Auf Testdatensätze angewendet	M30

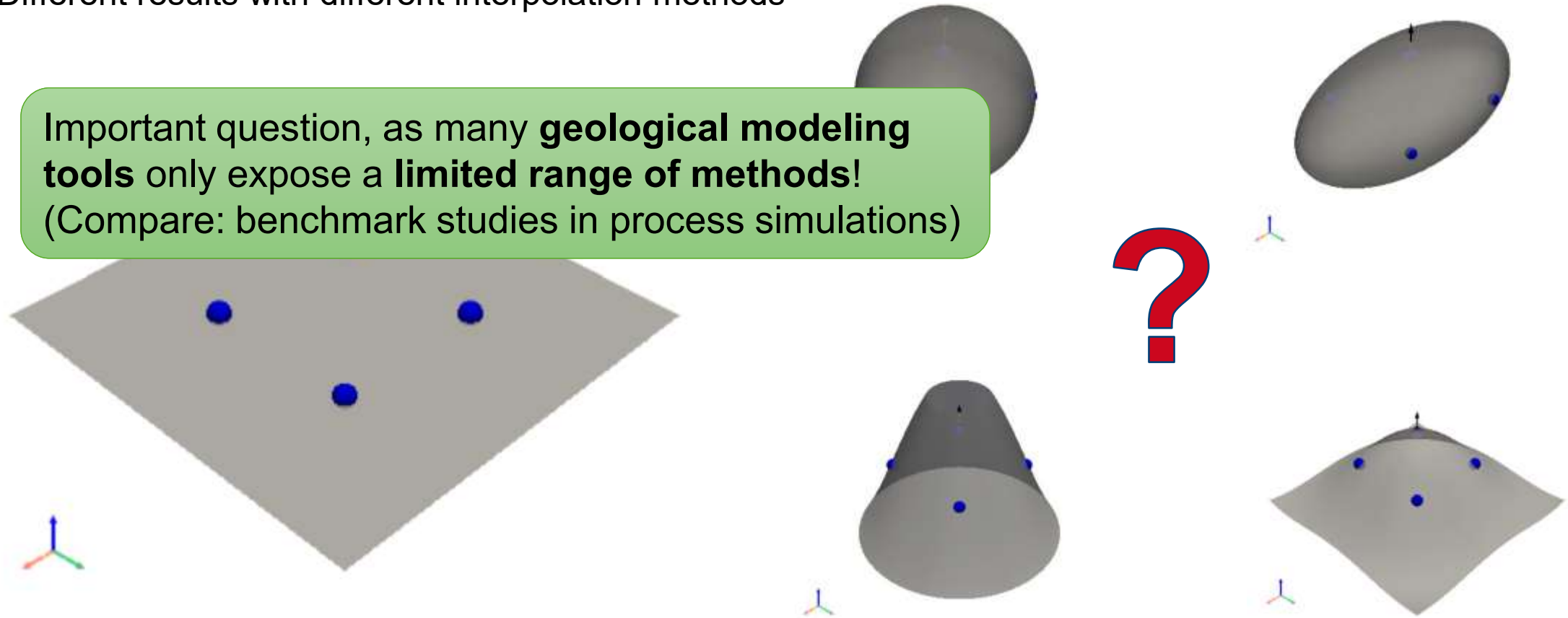
## Part 2: How to interpolate between known points?



# Which model is better?

- Same input data
- Different results with different interpolation methods

Important question, as many **geological modeling tools** only expose a **limited range of methods!**  
(Compare: benchmark studies in process simulations)



# Step1: Comparison between widely used interpolation

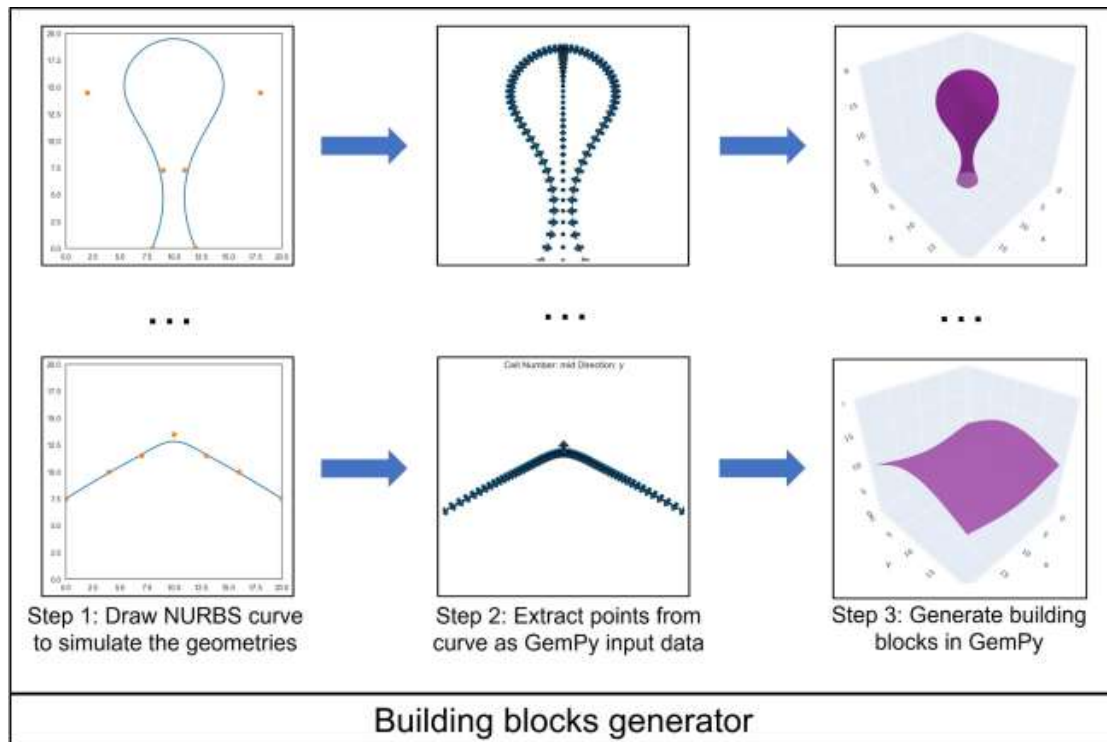
MS 4.1

Quantitativer Vergleich Interpolations-  
methoden für Strukturmodell

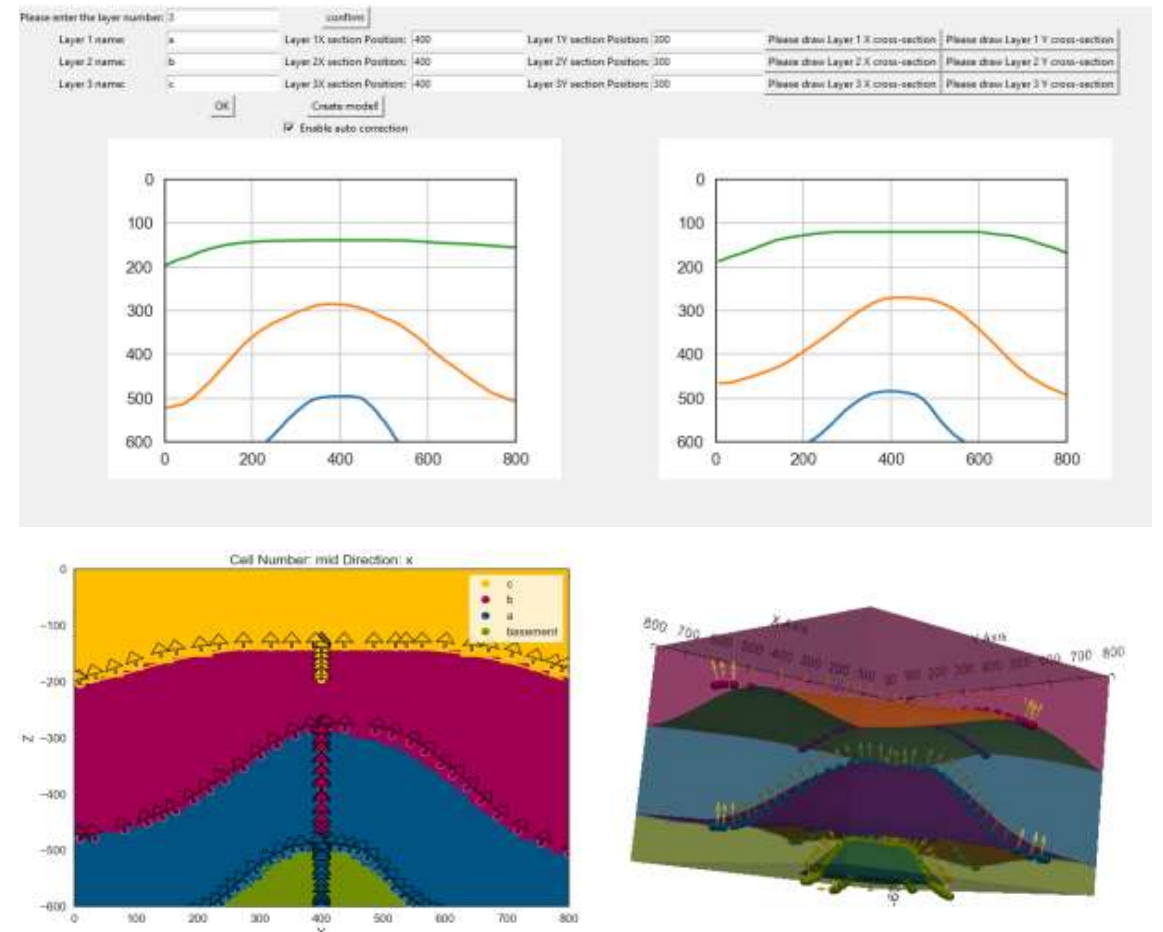
M18

Question: Which interpolation method should be applied to model host rock

## GemPy model generator V1.0

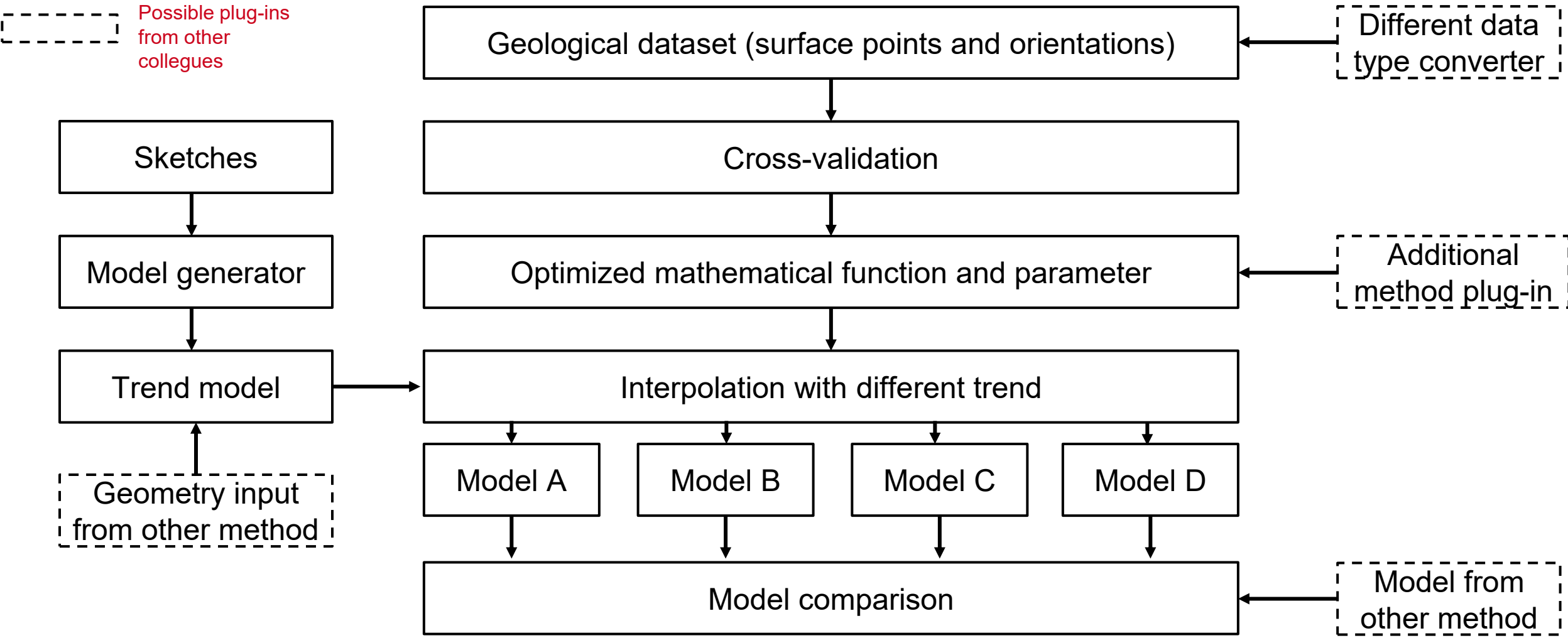


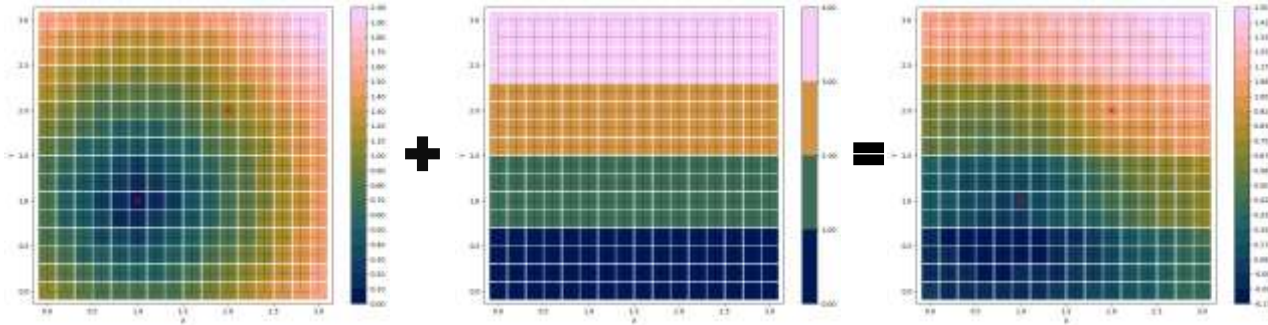
## GemPy model generator V2.0



# Step2: Comparison between optimized kernel functions and drift

MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18
MS 4.3	Hierarchische Modellkonstruktion implementiert	M33

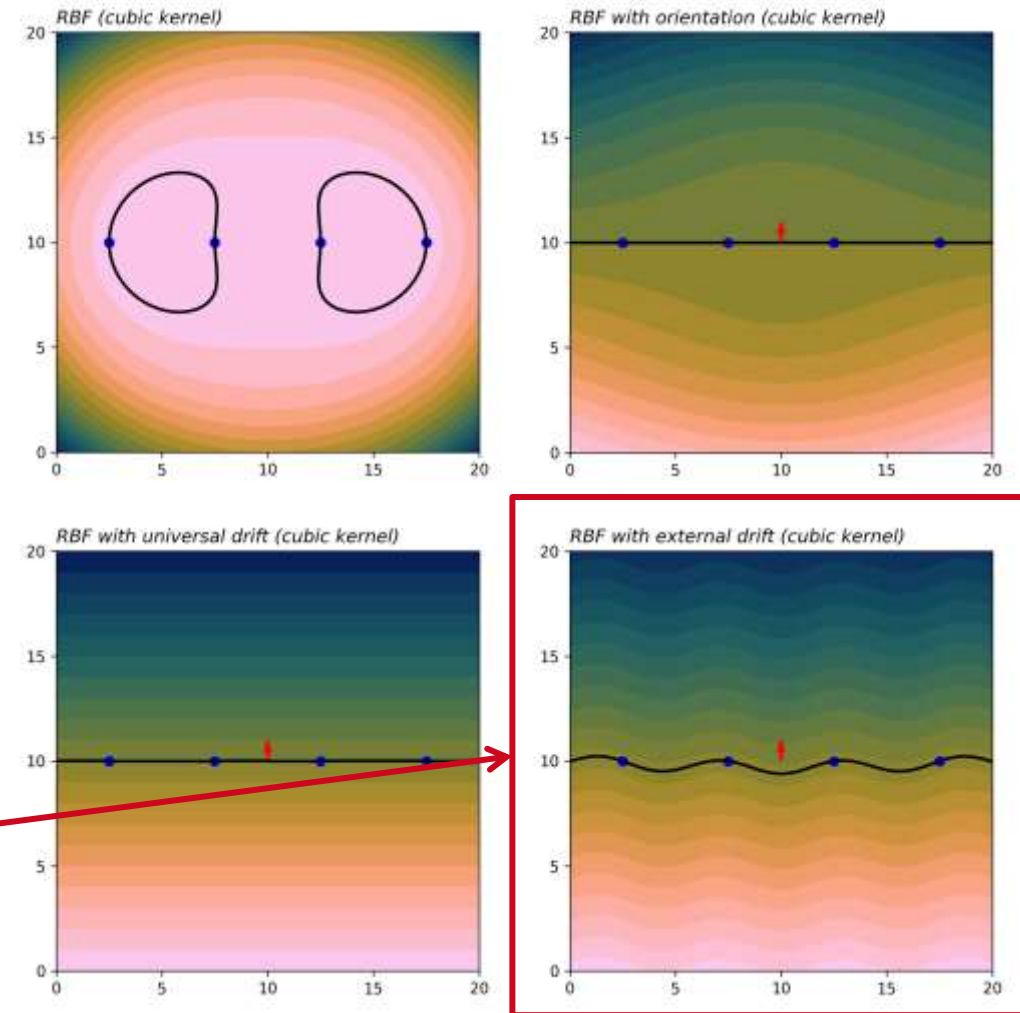




From simple to complex:

- Only surface points
- Surface points + Orientations
- Surface points + Orientations + Universal Drift
- Surface points + Orientations + External Drift

The suitable interpolation method for different geological geometries when we know additional information



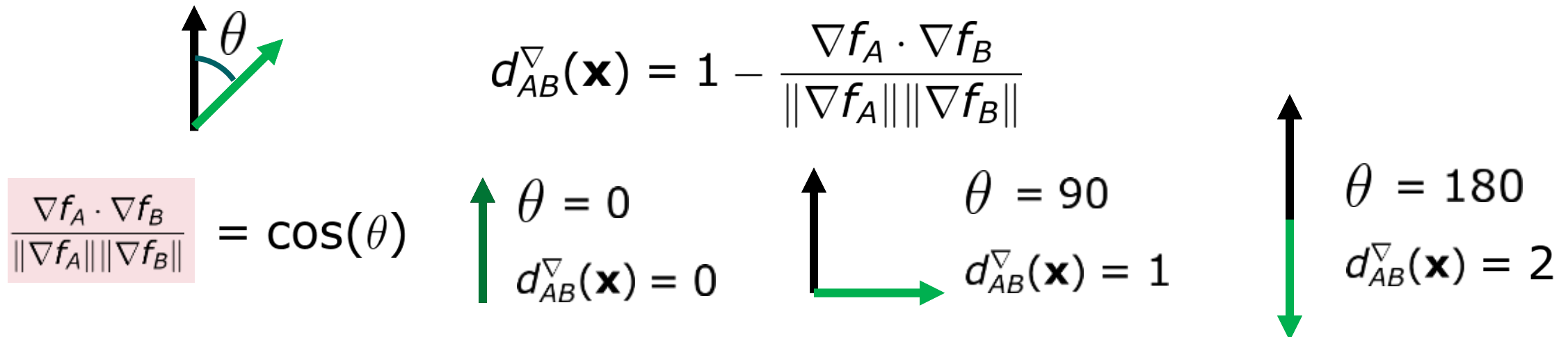


Several approaches:

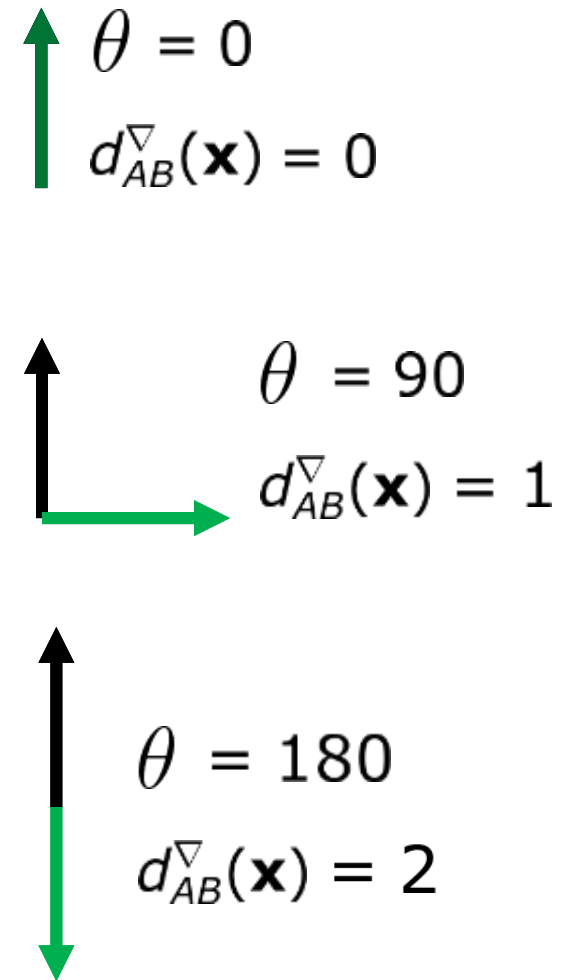
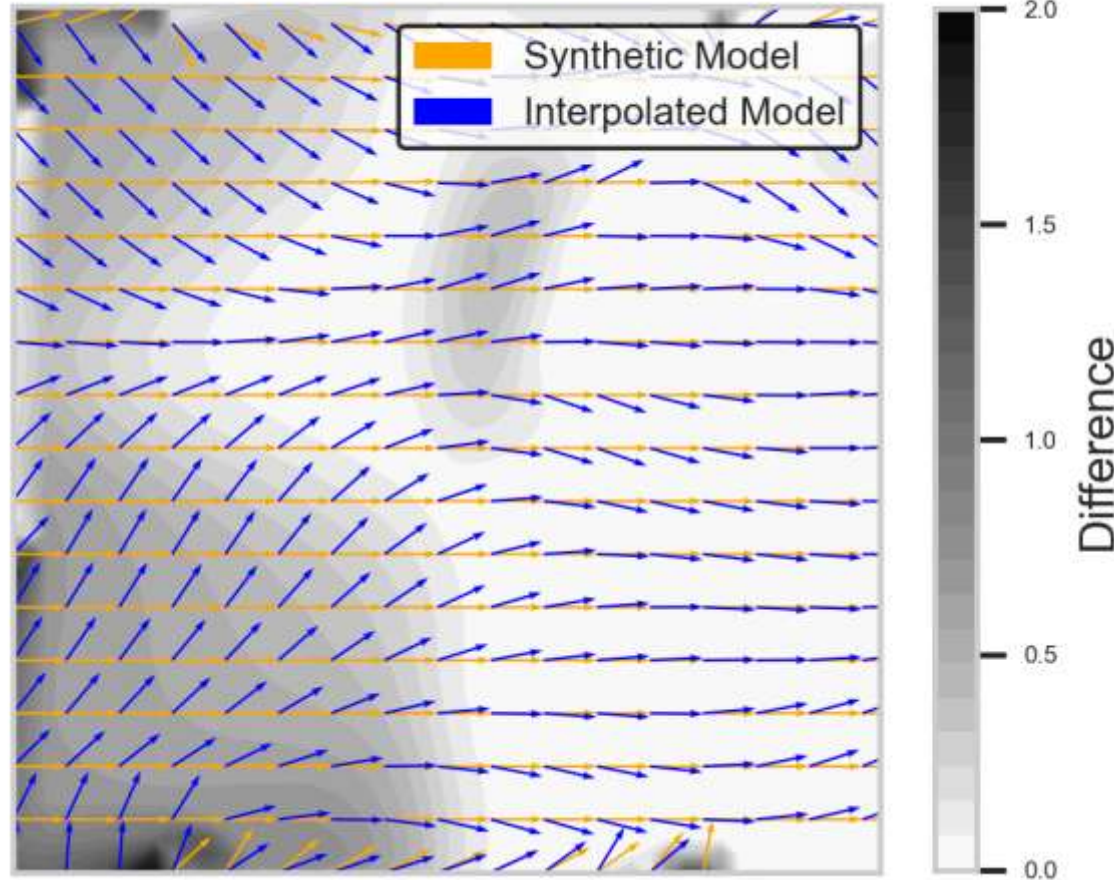
- Computing **summary metrics for each model** and the difference between these metrics (e.g., connectivity metrics, Thiele et al, 2016)
- Computing **distance between models** (e.g., The Hausdorff distance between layers, Suzuki et al, 2008)
- **Characterizing an ensemble** of models (e.g., with information entropy from rock units indicators, Wellmann & Regenauer-Lieb, 2012)

Here we focus on subsets of **implicit models** described by one **continuous scalar field**

(Guillaume Caumon. On some comparison metrics between 3D implicit structural models. IAMG 21st annual conference, 2022, Nancy, France. (hal-04165710))

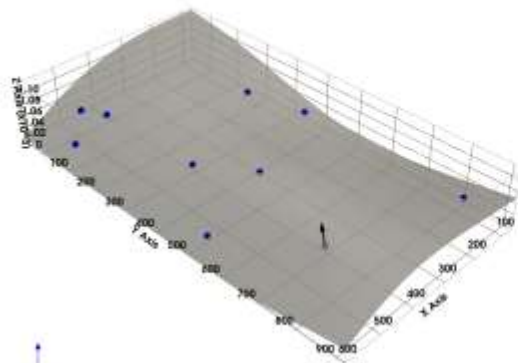


10 points difference plot (slice x=25)

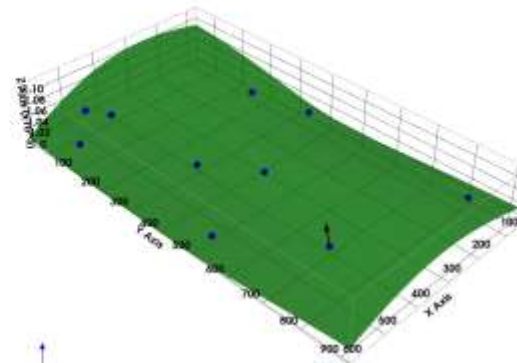


# Case study (synthetic fold model)

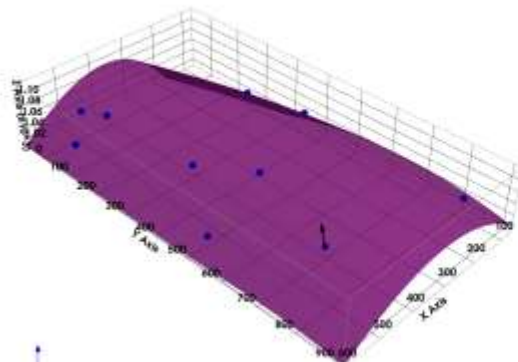
10 surface points and 1 orientation



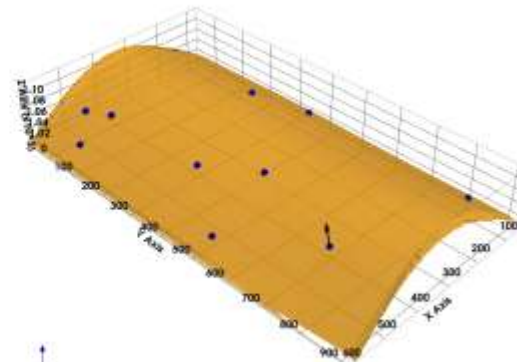
a) without drift



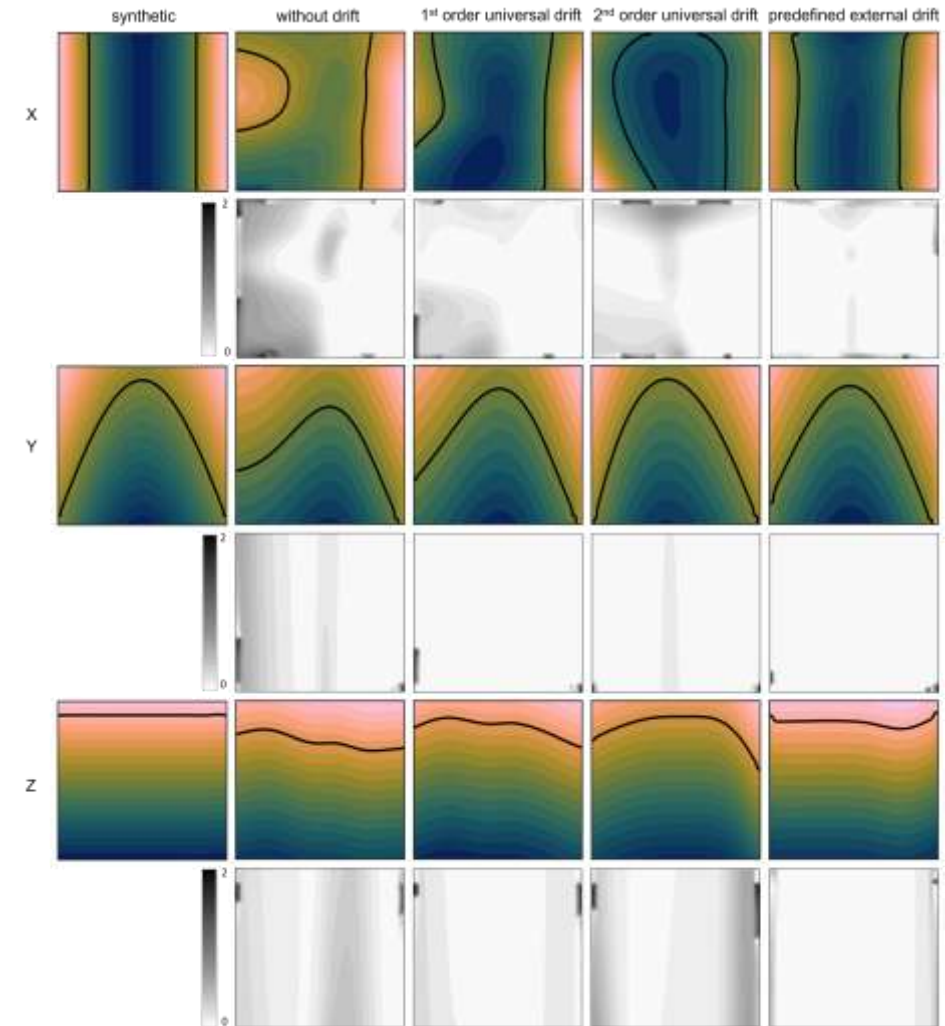
b) 1<sup>st</sup> order universal drift



c) 2<sup>nd</sup> order universal drift



d) predefined external drift



# Case study (synthetic fold model)

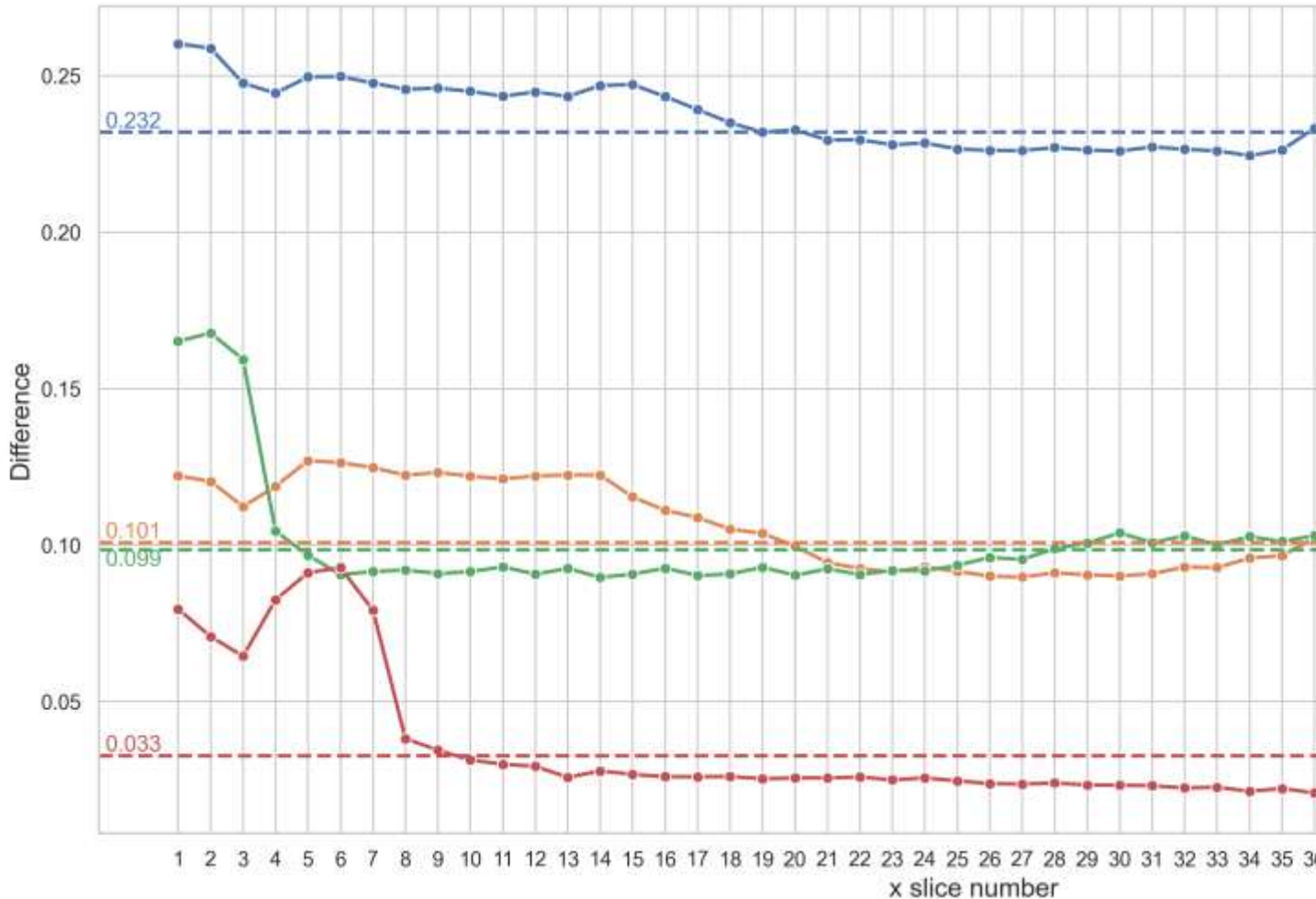
## Three-Dimensional Modeling of Geological Bodies Using Radial Basis Function with External Drift Function

Jian Yang<sup>1</sup>, Friedrich Carl<sup>2</sup>, Peter Aichtziger<sup>3</sup> and Florian Wellmann<sup>1</sup>

<sup>1</sup> Chair of Computational Geoscience, Geothermics and Reservoir Geophysics, RWTH Aachen University

<sup>2</sup> Chair of Engineering Geology and Hydrogeology, RWTH Aachen University

<sup>3</sup> Fraunhofer Research Institution for Energy Infrastructures and Geothermal Systems IEG, Aachen



CS difference plot between synthetic model and interpolate

### Motivation

This study addresses the gap by proposing an innovative approach that integrates geometrical external drift into the RBF framework. This enhancement allows the RBF models to be purposefully biased towards desired geometric configurations, significantly improving their ability to accurately model various geological structures such as planar strata, folded formations, and salt domes.

Sources and types of uncertainty related to different modeling steps (Wellmann & Caumon, 2018)

### Method and Result

**Step 1: Get formatted dataset**

**Step 2: Use cross-validation to optimize kernel functions and parameters**

**Step 3: Interpolate with different drift**

**Step 4: Model comparison (cosine similarity)**

CS difference plot between synthetic model and interpolated model with different drift functions. The difference is shown in 50 slices on the x-axis.

Differences in interpolations across various data densities (10, 30, and 50 points) under different drift models: without drift, 1st order universal drift, 2nd order universal drift, and external drift.

Differences between random sampling and statistical sampling methods across various drift models (without drift, 1st order universal drift, 2nd order universal drift, and external drift) using 10 data points.

### Reference

- Wellmann, F. and Caumon, G. (2018). 3-d structural geological models: concepts, methods, and uncertainties. *Advances in Geophysics*, 1-121. <https://doi.org/10.1016/bs.agph.2018.09.001>
- Caumon G (2022) On some comparison metrics between 3D implicit structural models. In IAMG 21<sup>st</sup> annual conference, Nancy, France

# Part 3. Input data and uncertainties in interpretation

## Specific objectives:

- Identify data types used in site selection
- Identify uncertainties in these input data types
- Assess errors in seismic data interpretation
- Characterise qualitative irreducible uncertainties.

- Expected contribution to site selection:

Provide a method to determine the spatial uncertainty in geological models and geometries.

- Input Data Types:

### **Seismic data**

Velocity models –depth conversion

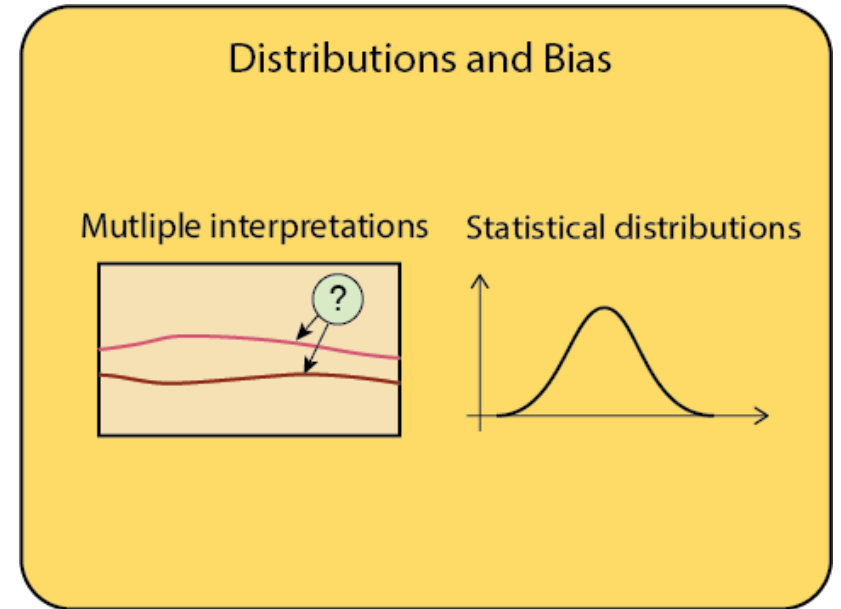
Seismic raw data – acquisition/processing

Borehole data

Geological maps and reports

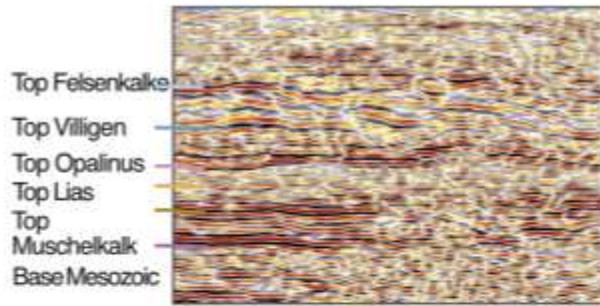
Other geophysical and well data

\*Expert prior knowledge\*



# Mutiple interpretations of seismic image data

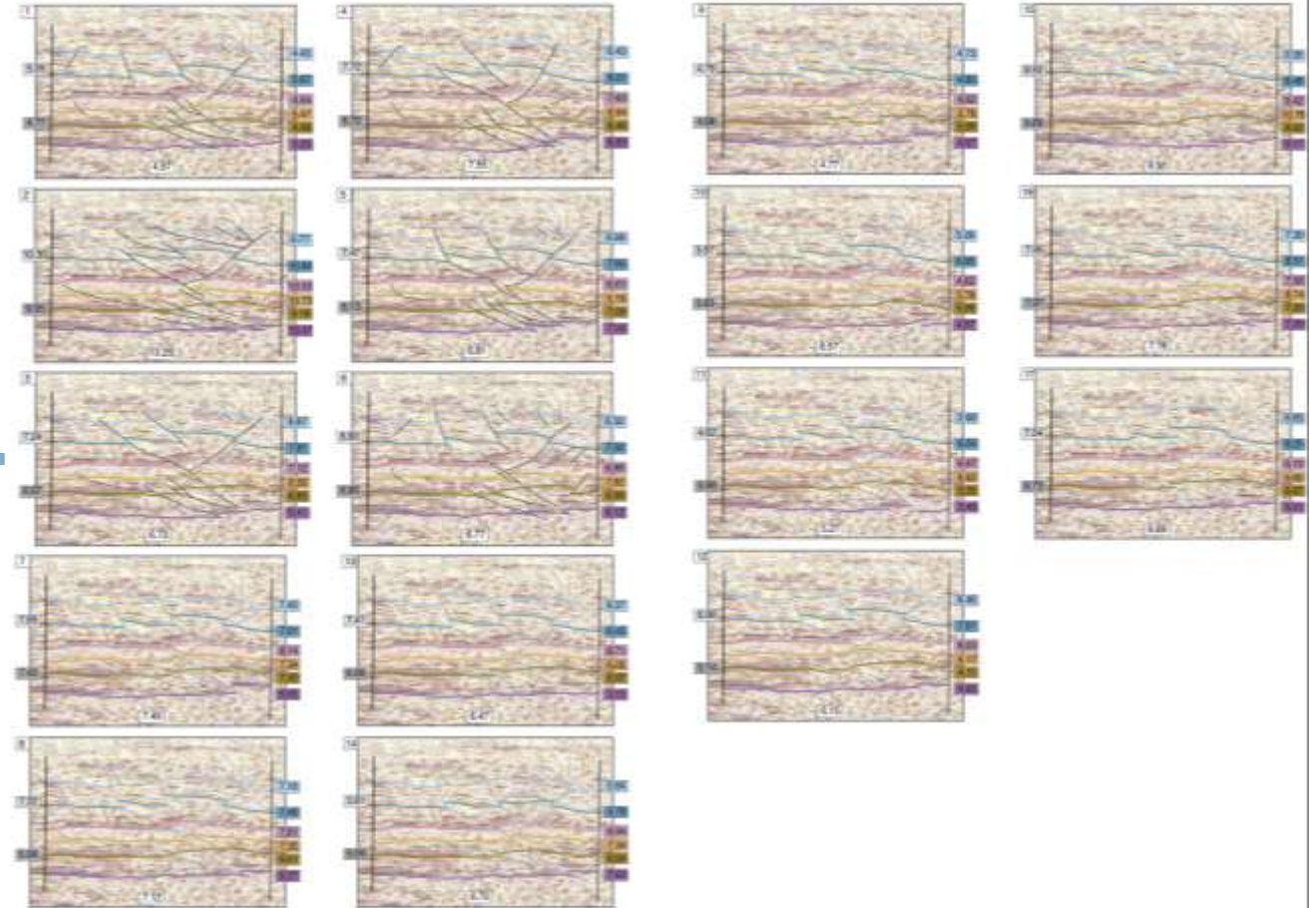
## Interpretational Uncertainty:



Interpretation Workflow 1: Horizon interpretation leaving fault gaps.

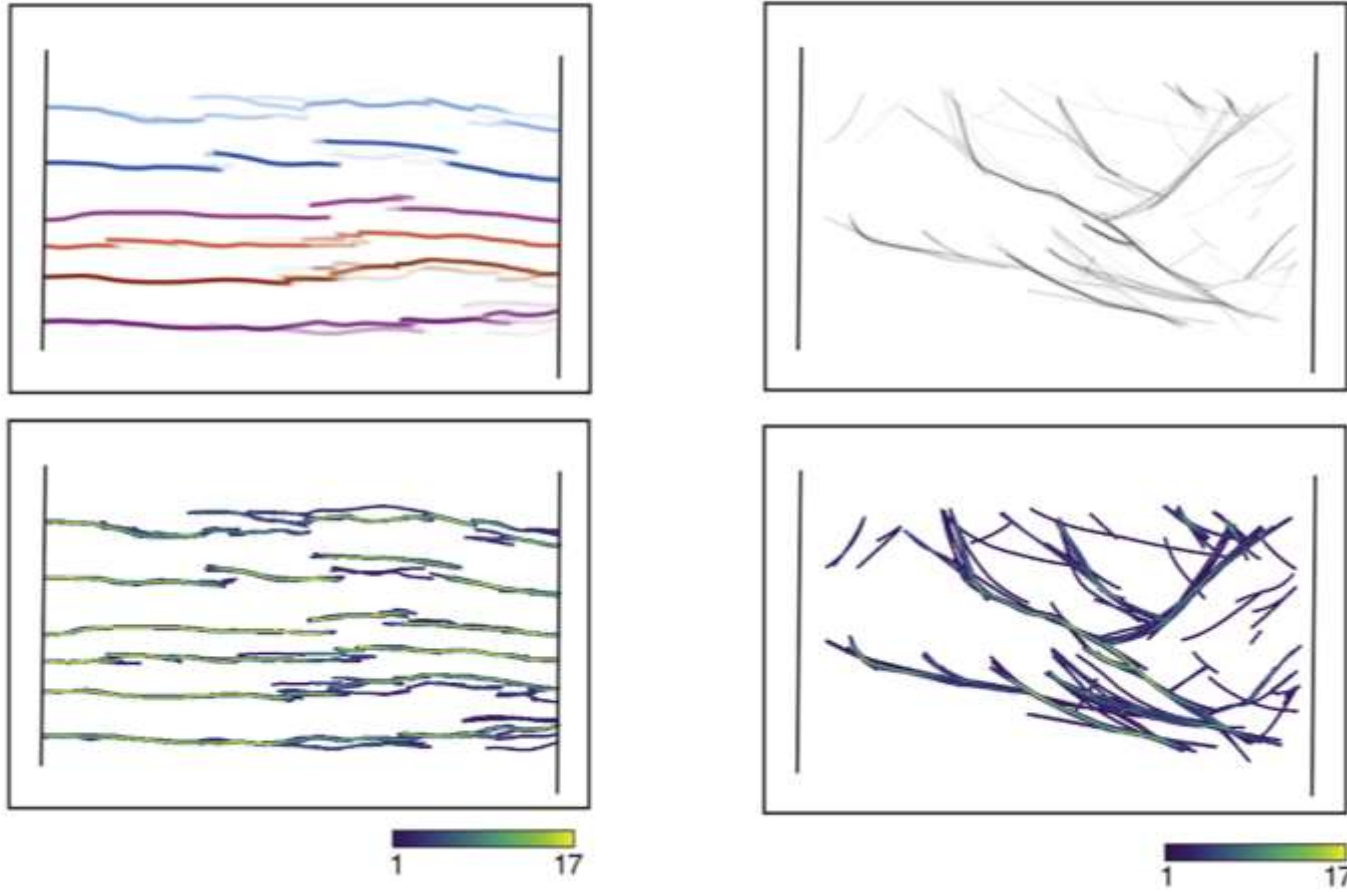


Interpretation Workflow 2: Fault interpretation linking gaps.



# Mutiple interpretations of seismic image data

## Interpretational Uncertainty:



**SUBJECTIVE UNCERTAINTY IN SEISMIC INTERPRETATION**  
 Rubinski, Franuska, Bond, Carr  
 Department of Geology and Petroleum Geology, School of Earth and Atmospheric Sciences, University of Alberta, 1-16A Science Building, 8350-136 Avenue, Edmonton, AB T6G 2G1

**UNCERTAINTY IN FAULT SEISMIC INTERPRETATION**

Seismic reflection image processing and interpretation involves a complex workflow. The interpretation of seismic reflection data is a subjective process, influenced by interpreter experience, knowledge, and the quality of the data. This paper discusses the sources of uncertainty in fault seismic interpretation and provides a framework for understanding and managing this uncertainty.

**SUBJECTIVE UNCERTAINTY**

Subjective uncertainty in fault seismic interpretation arises from the interpreter's personal experience, knowledge, and the quality of the data. This uncertainty is influenced by the interpreter's experience, knowledge, and the quality of the data.

- The interpretation of a fault varies between different interpreters.
- The interpretation of a fault varies over time, influenced by the interpreter's experience, knowledge, and level of confidence.
- Multiple fault interpretations for a single profile.

**DIFFERENT FAULT INTERPRETATION WORKFLOWS LEAD TO DIFFERENT FAULT MAPS**

Figure 1: Workflow diagram showing the process from seismic data to fault maps. It illustrates how different interpretation workflows lead to different fault maps.

**UNDERSTANDING AND MANAGING UNCERTAINTY IN SEISMIC INTERPRETATION FOR SUBSURFACE MODELLING**

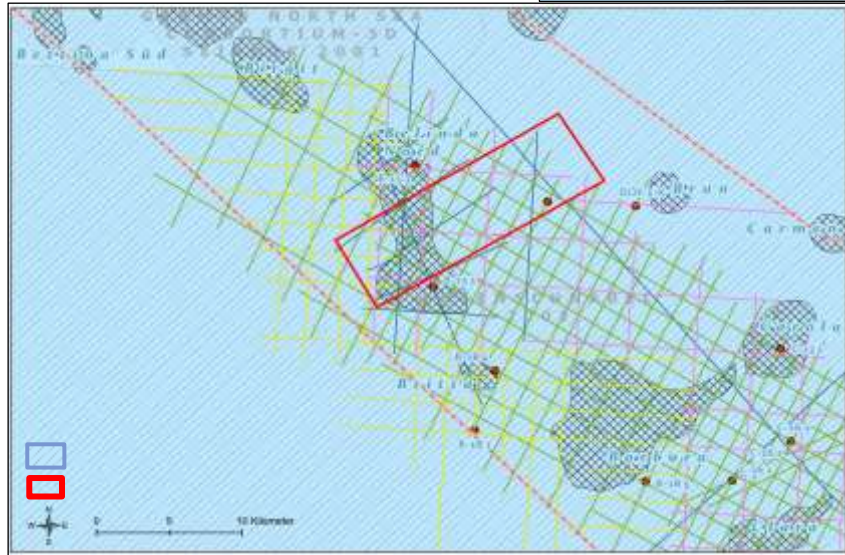
The subsurface model is a representation of the earth's subsurface. It is used to predict the behavior of the subsurface. The model is built based on the interpretation of seismic data. The model is used to predict the behavior of the subsurface.

# Uncertainty Quantification from Seismic Data (Current Work)

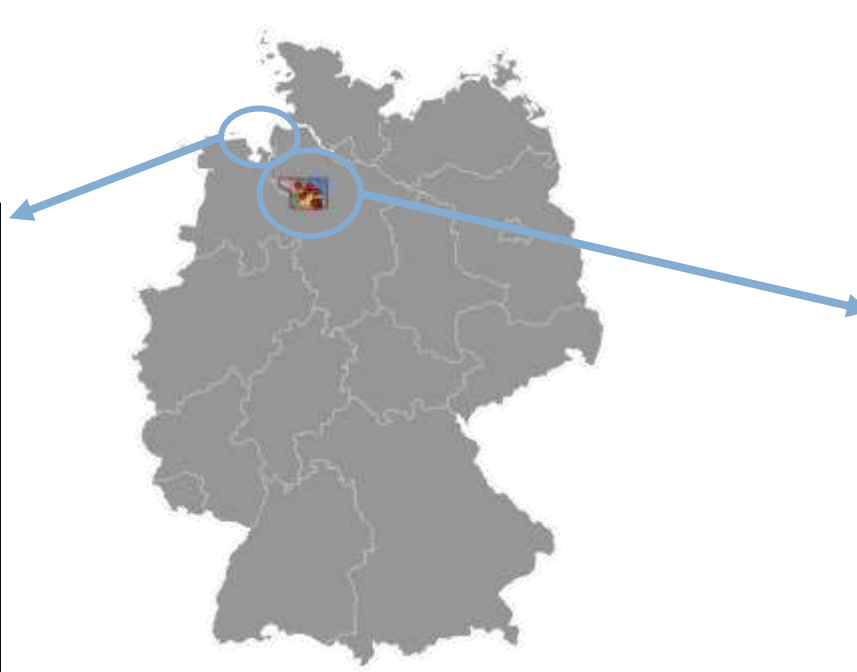
MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

## Input Data

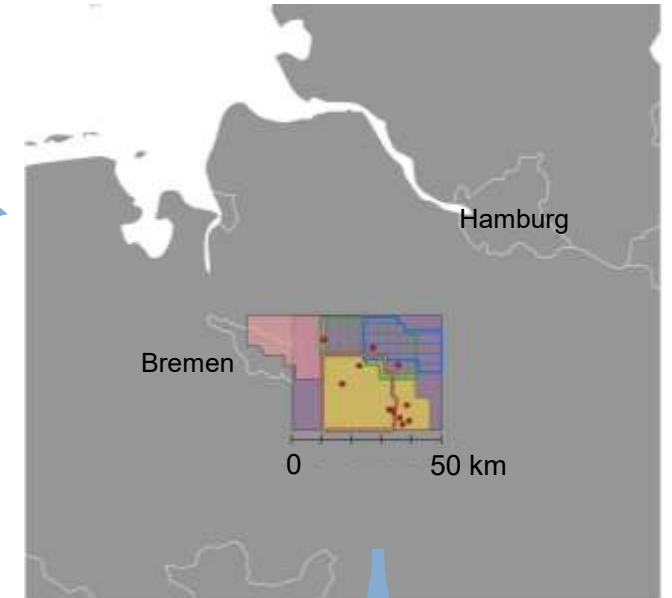
### North Sea Area



1974-2002 (+ reprocessing)



### North Hannover Area



NH 032\_Wedehof-Hamwiede 93-94

NH 2040\_Wedehof-Hamwiede Repr 2010

NH 2012\_Völkersen Repr 2012

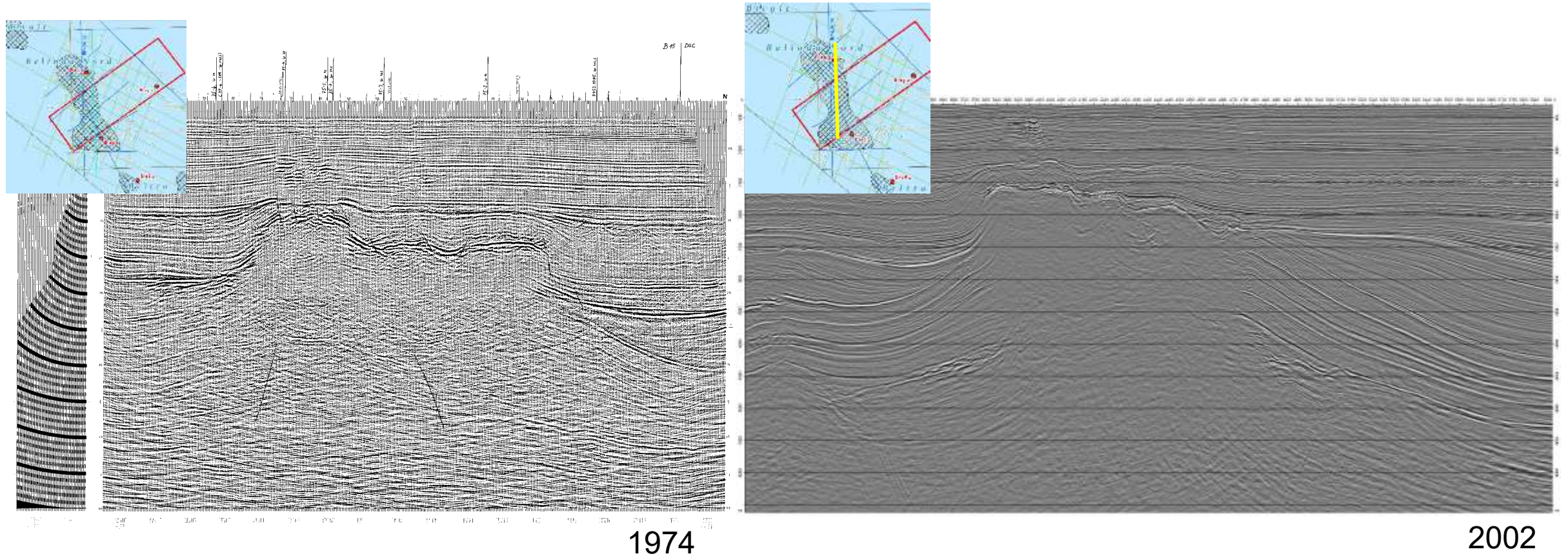
NH 2039\_NordHannover Repr 2004



# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

Input Data – vintages of seismic, reprocessing – data quality and bias/prior knowledge



MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

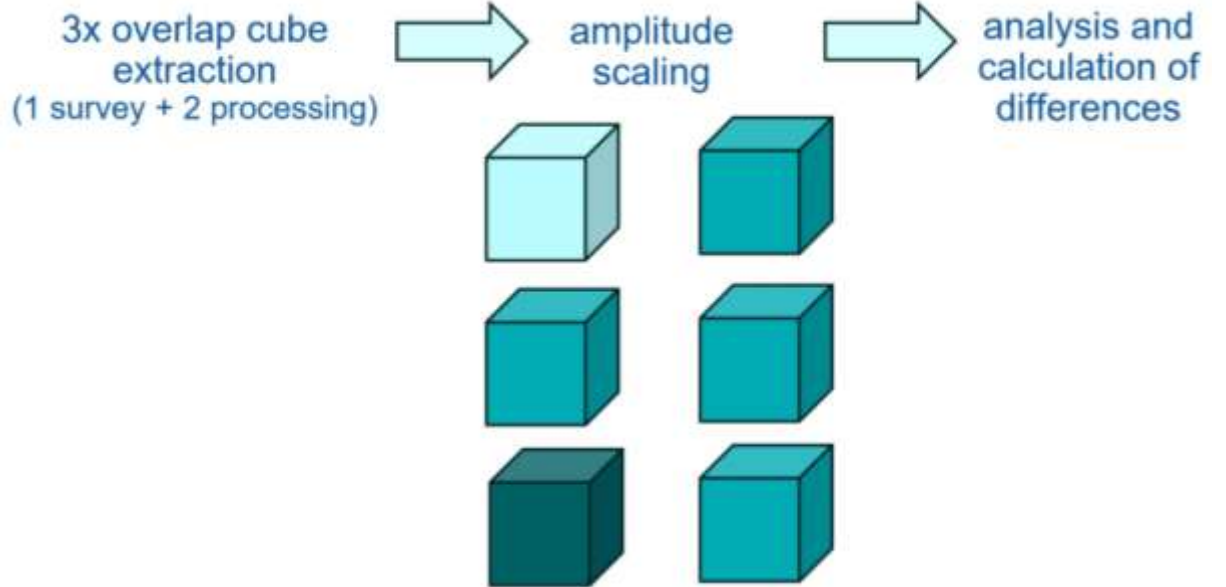
# Uncertainty Quantification from Seismic Data (Current Work)

## Seismic Workflow 1 Uncertainties in Processing

### Seismic Data – Hannover Data Focus

- Quality control of data received from Landesamt für Bergbau, Energie und Geologie and contributed by BGR.
- Database set-up in Petrel.

- Amplitude scaling to create comparable amplitude volumes from different reprocessings.
- Calculate differences between comparable amplitude volumes using calculator (different amplitudes and extract attributes - frequency, reflectivity, coherence)
- Compare different depth-conversion results/approached/techniques



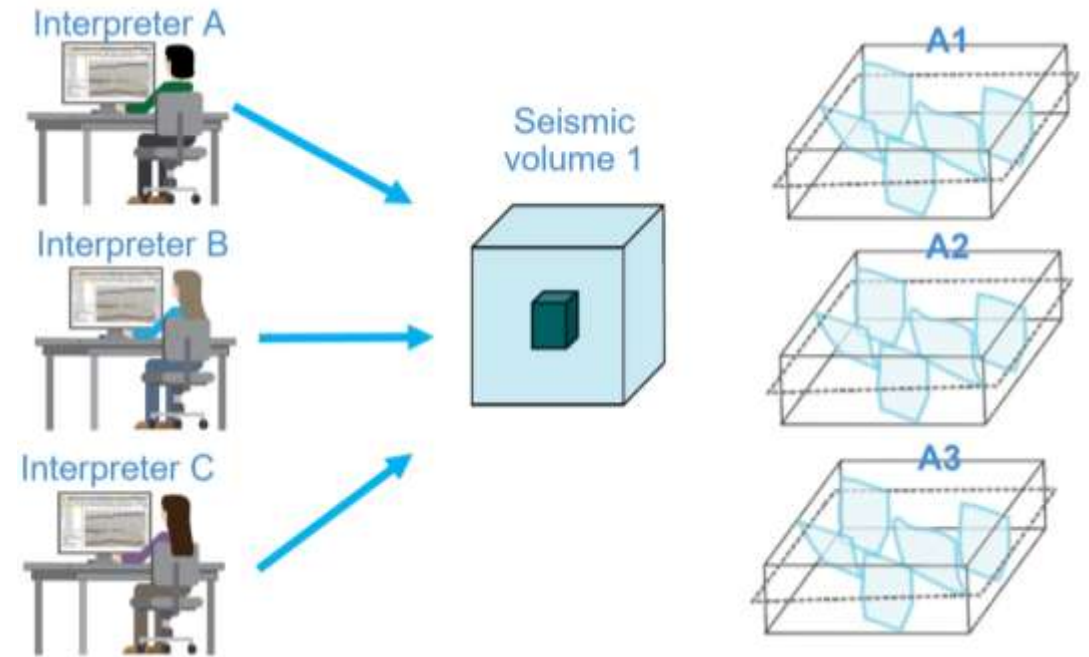
MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

# Uncertainty Quantification from Seismic Data (Current Work)

## Seismic Workflow 2 Uncertainties in Interpretation

### Seismic Interpretation – Hannover and North Sea

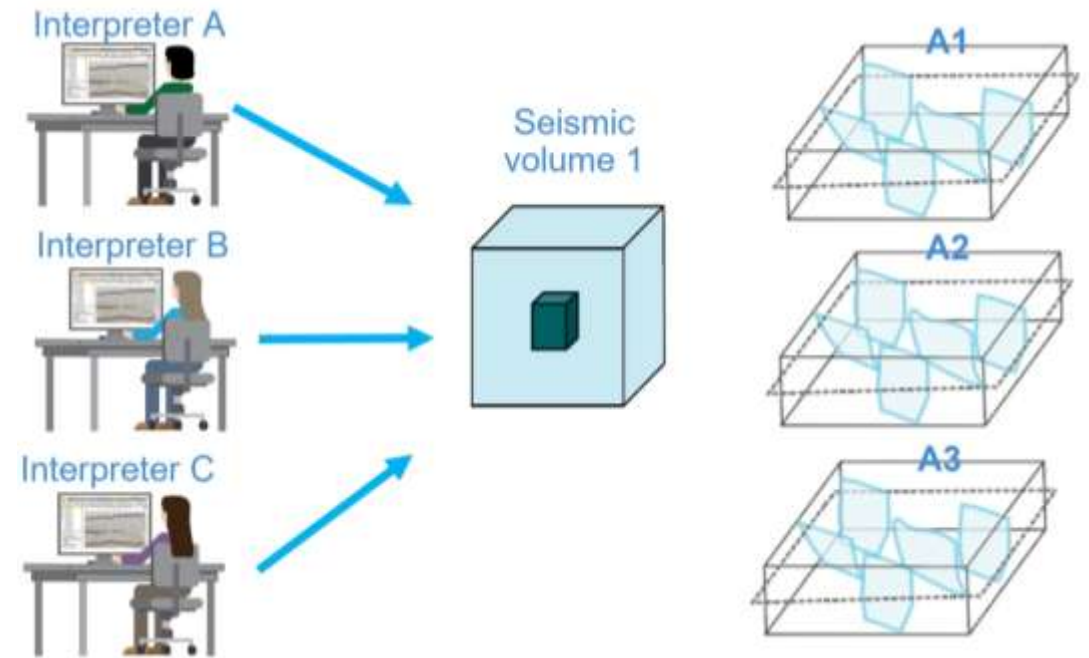
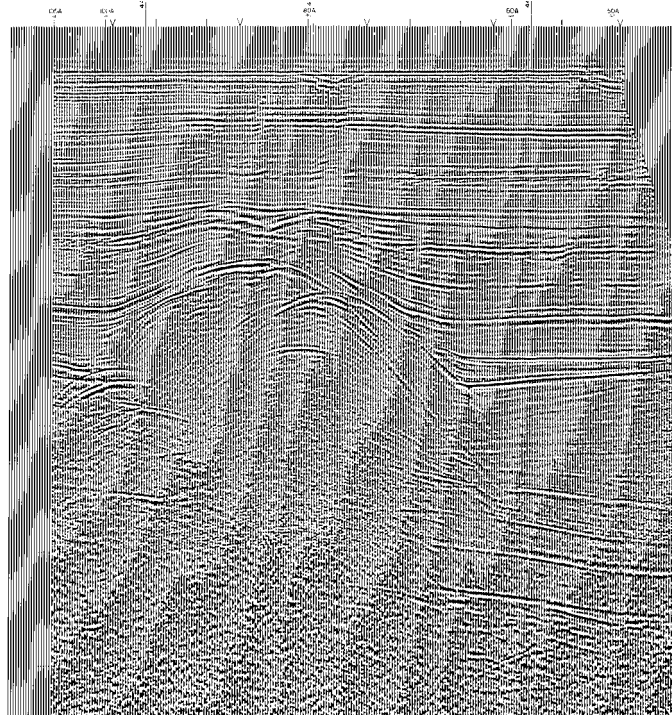
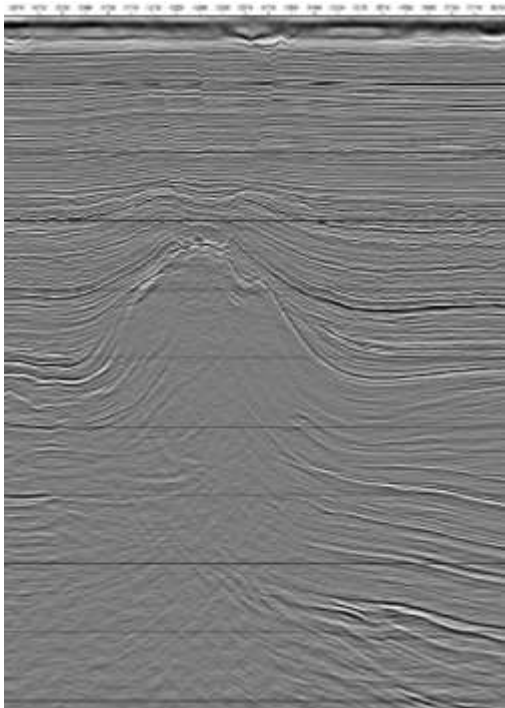
- Multi-interpreter experiment of seismic profiles and volumes (3D and 2D)
- Seismic Interpreters with different levels of expertise:
  - BGR and Fraunhofer IEG, Bachelor's and Master's students from University of Aberdeen, RWTH Aachen University, Leibniz University, Conference attendees.



# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

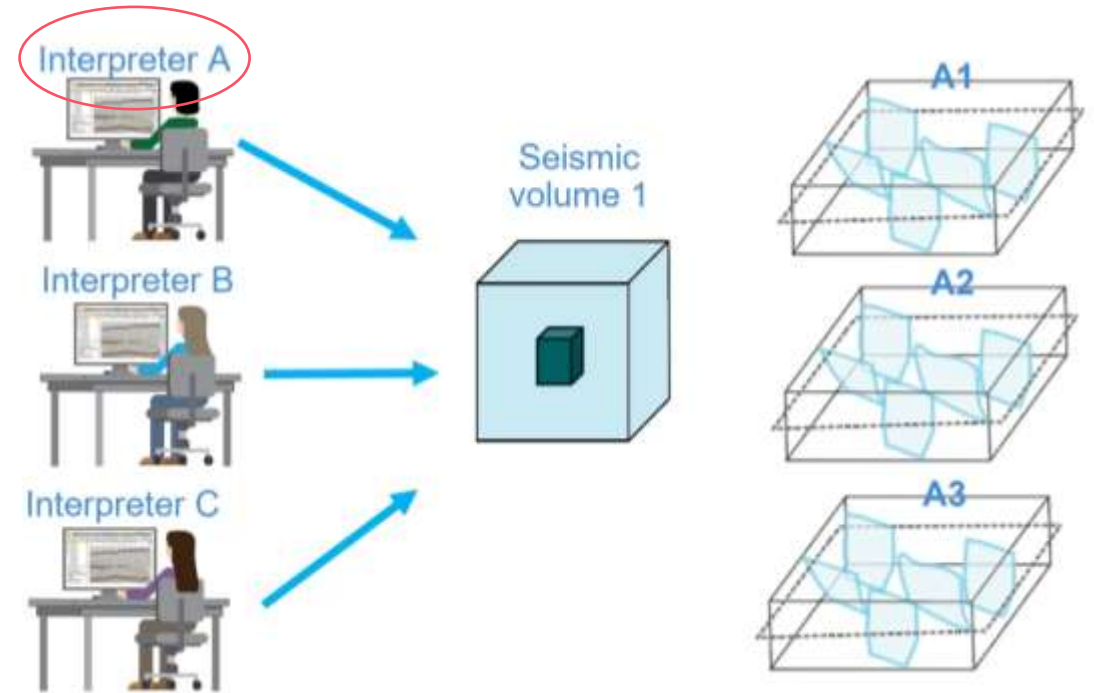
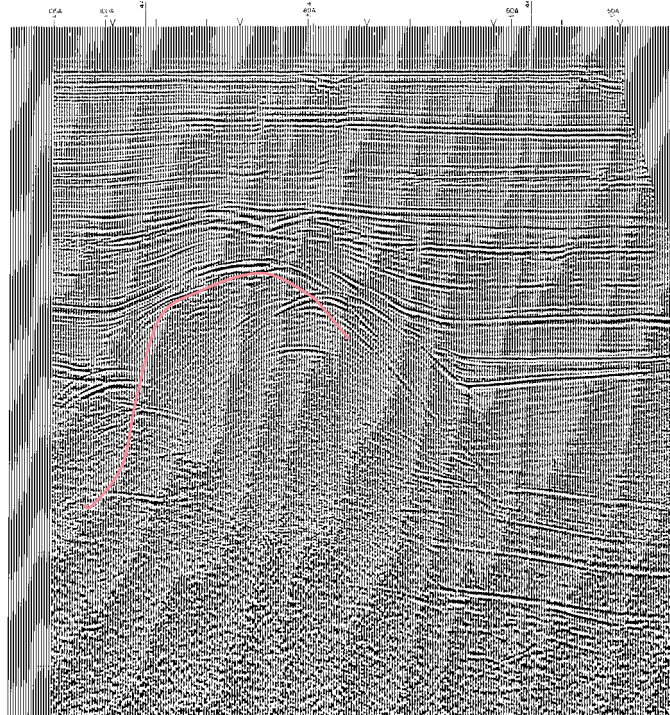
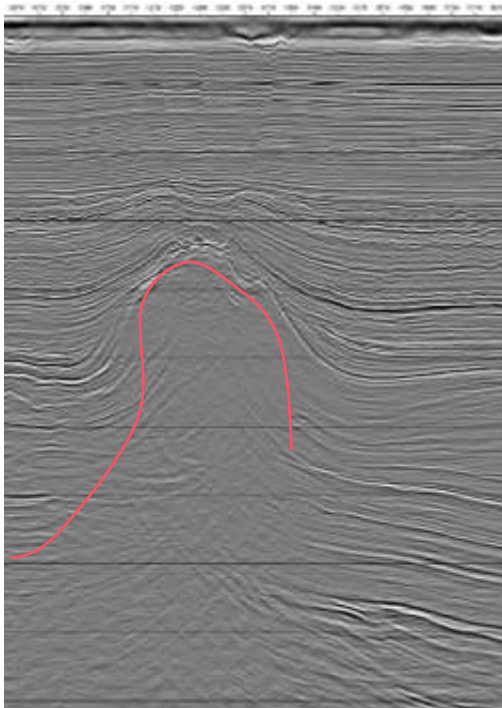
## Seismic Workflow 2 Uncertainties in Interpretation



MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

# Uncertainty Quantification from Seismic Data (Current Work)

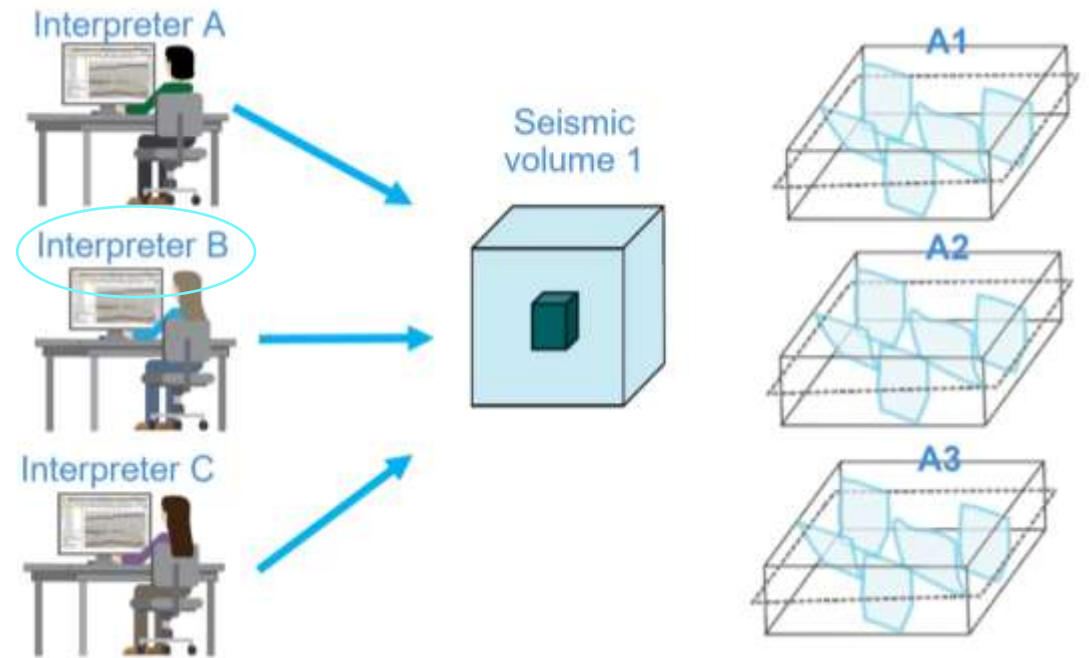
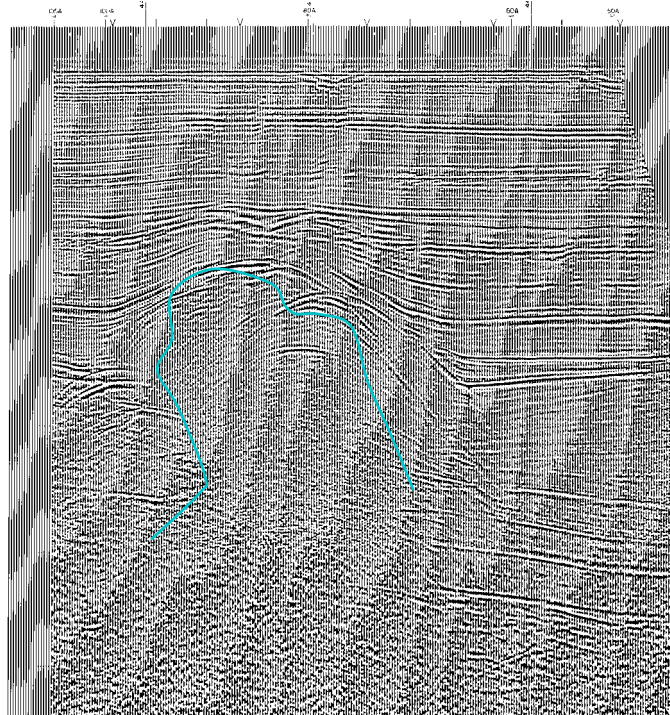
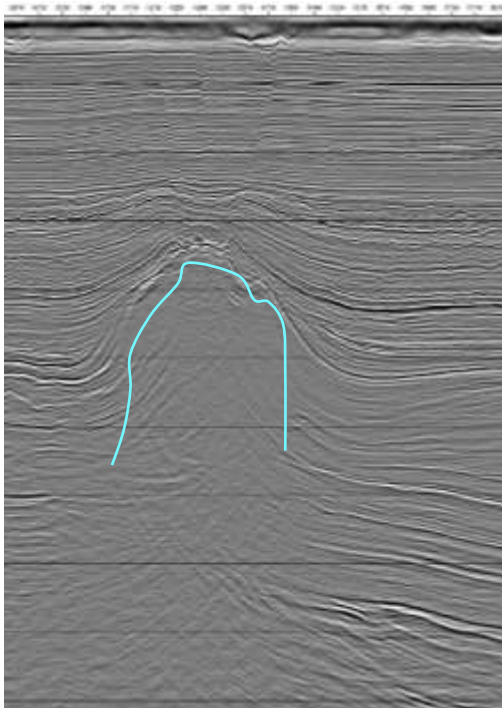
## Seismic Workflow 2 Uncertainties in Interpretation



MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

# Uncertainty Quantification from Seismic Data (Current Work)

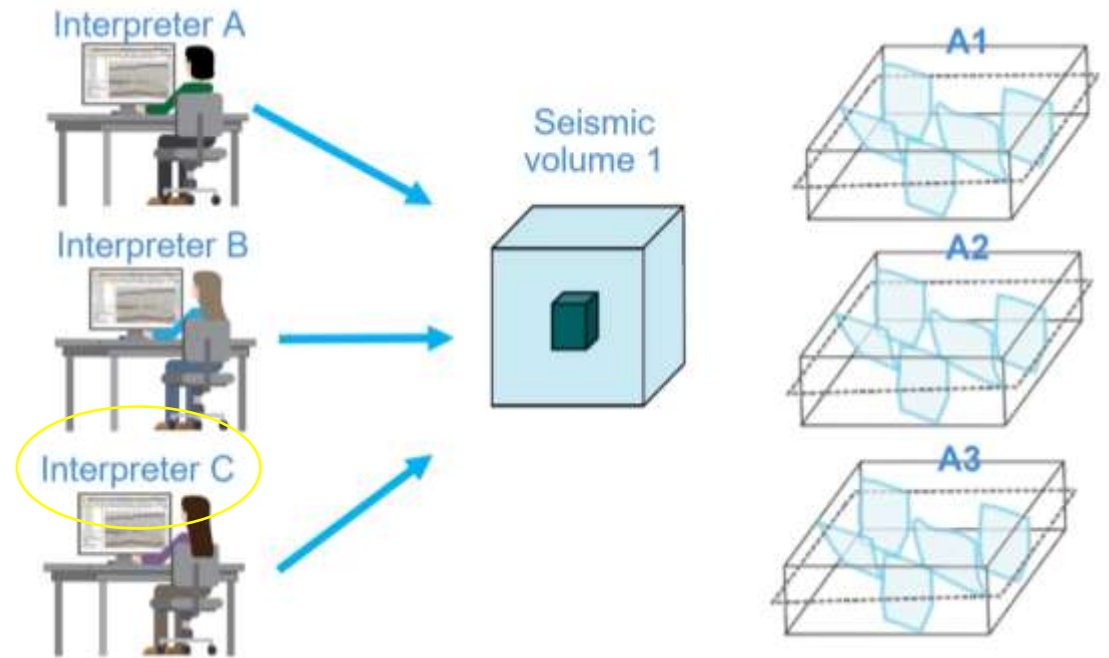
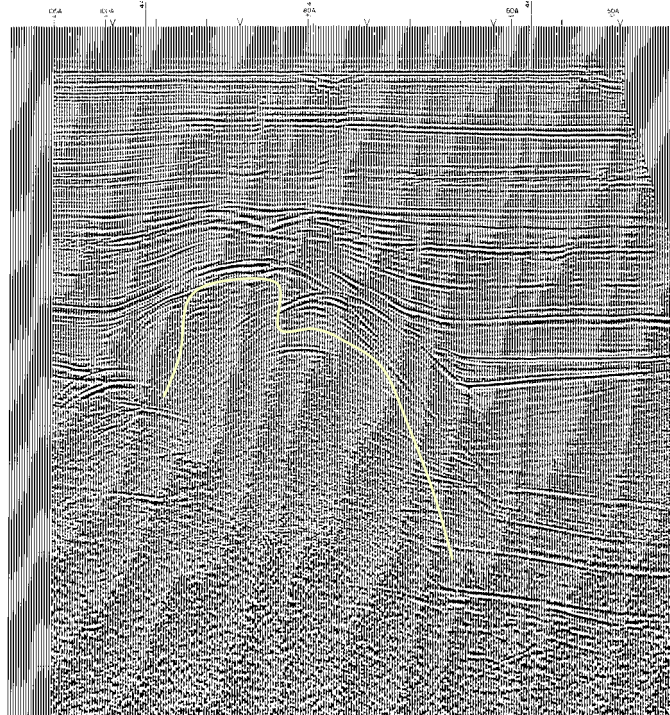
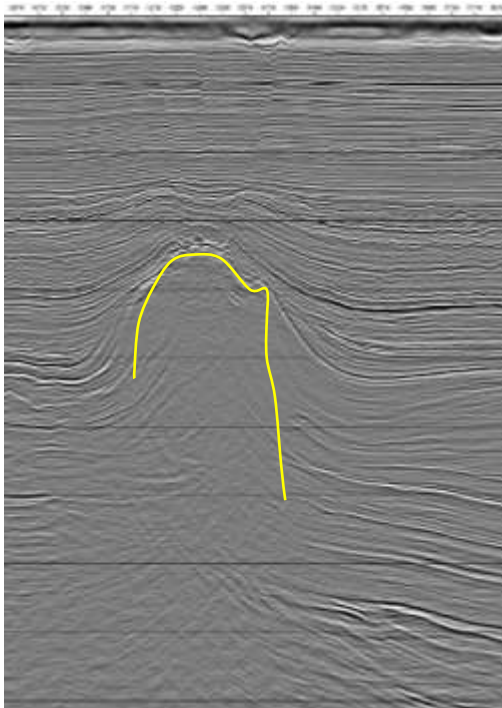
## Seismic Workflow 2 Uncertainties in Interpretation



MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

# Uncertainty Quantification from Seismic Data (Current Work)

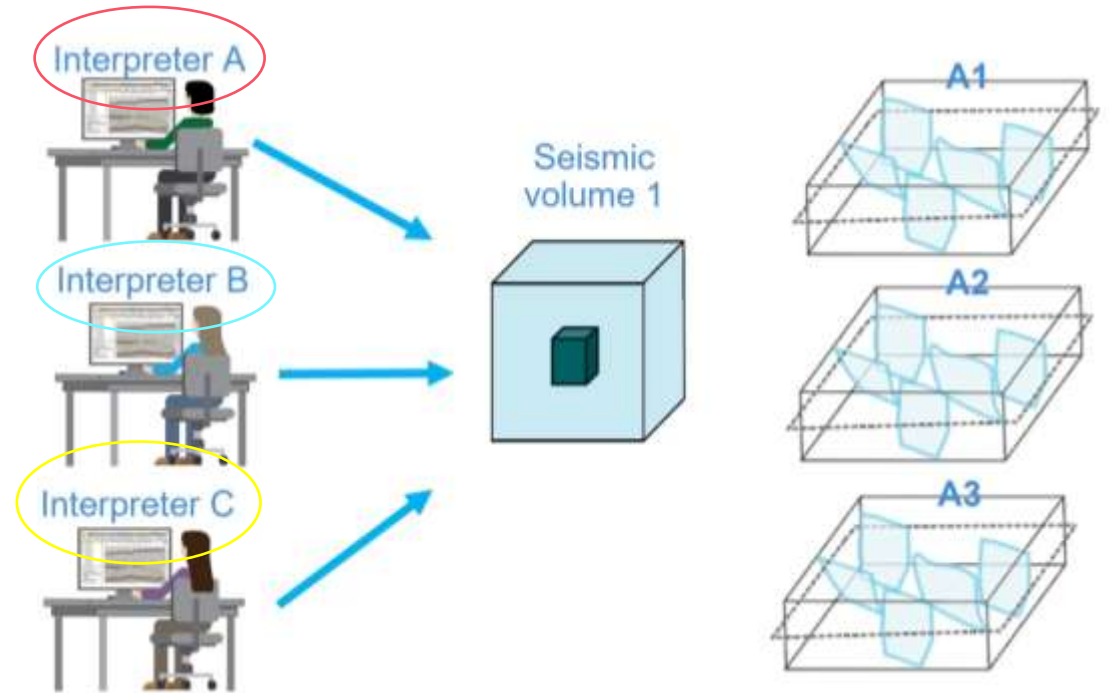
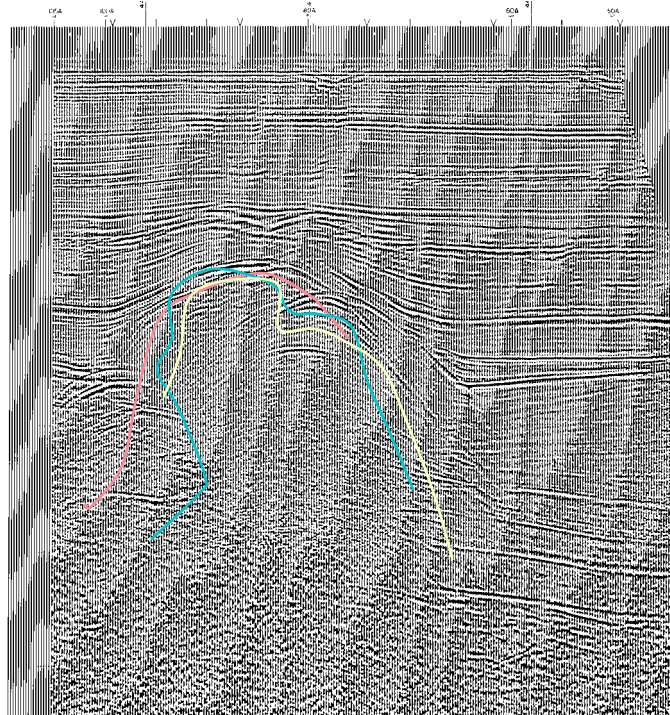
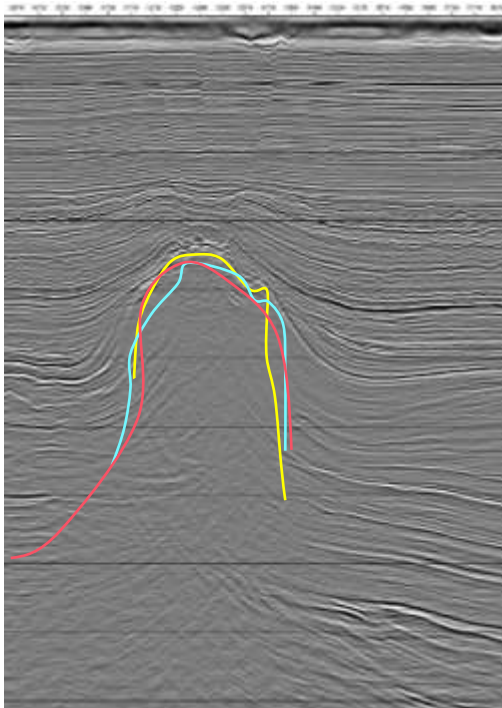
## Seismic Workflow 2 Uncertainties in Interpretation



# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

## Seismic Workflow 2 Uncertainties in Interpretation

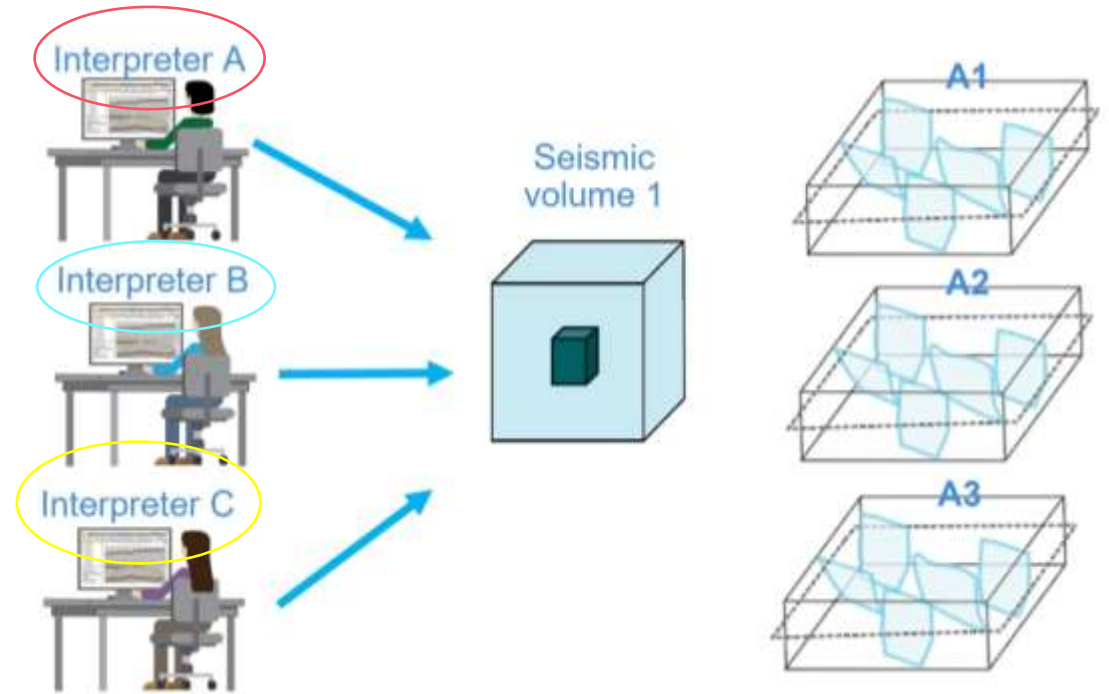
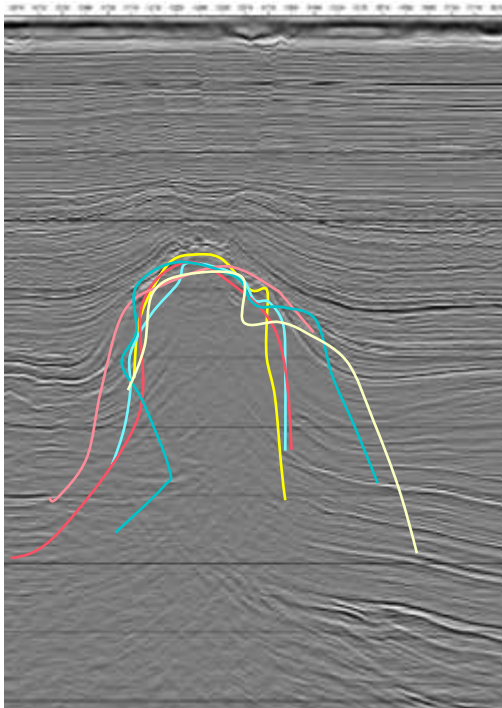




MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

# Uncertainty Quantification from Seismic Data (Current Work)

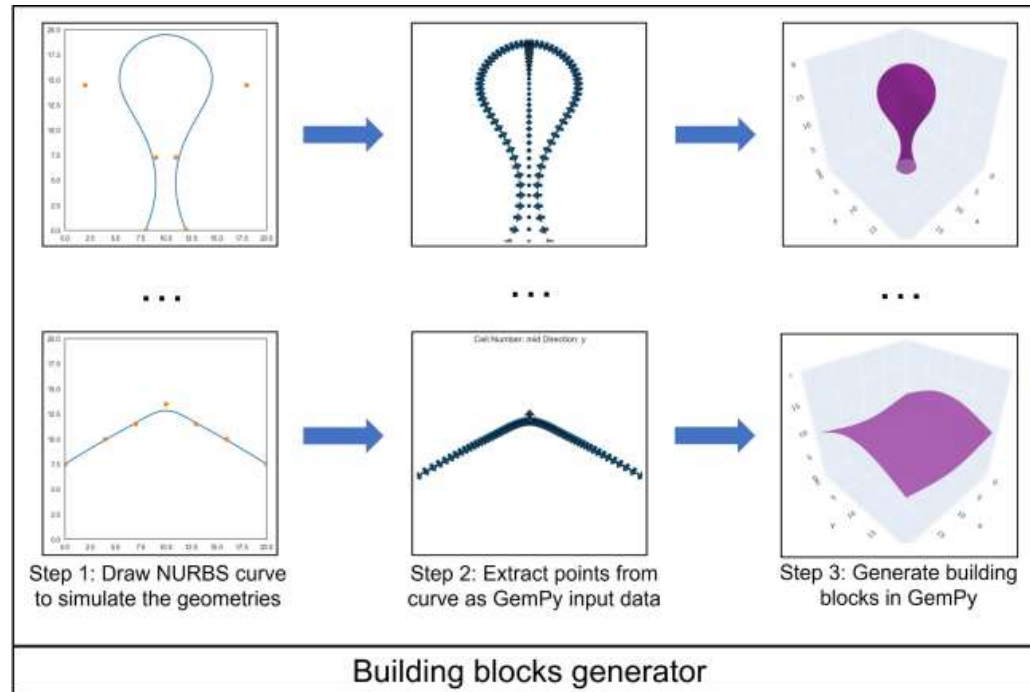
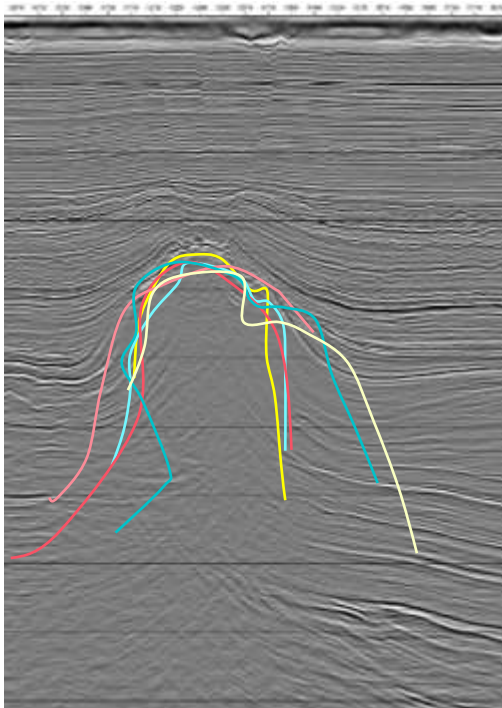
## Seismic Workflow 2 Uncertainties in Interpretation



# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

## Workflow Integration



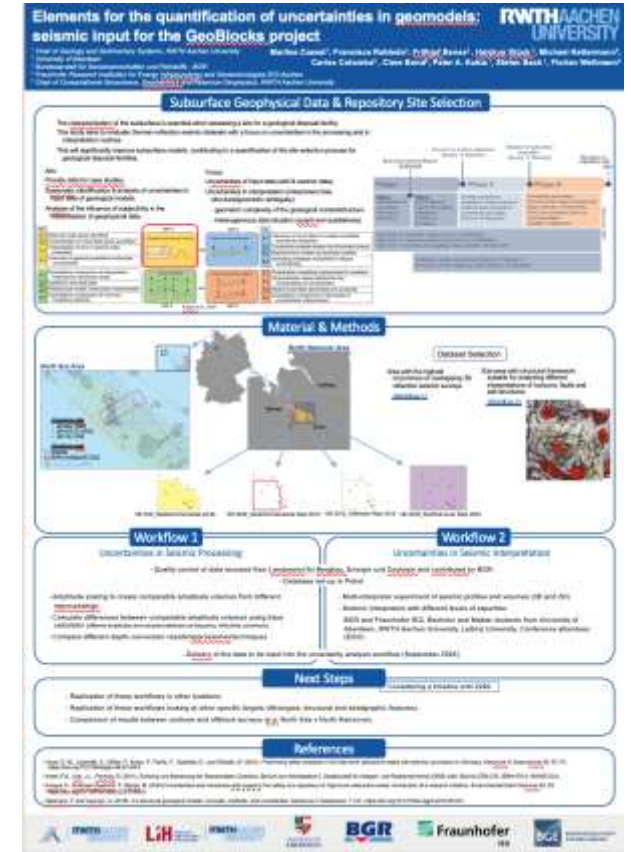
# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

Set of case studies and associated data

Systematic identification & analysis of uncertainties in input data of geological models evidenced by the case studies.

Analysis of the influence of subjectivity in the interpretation of geophysical data - interpretation PDF



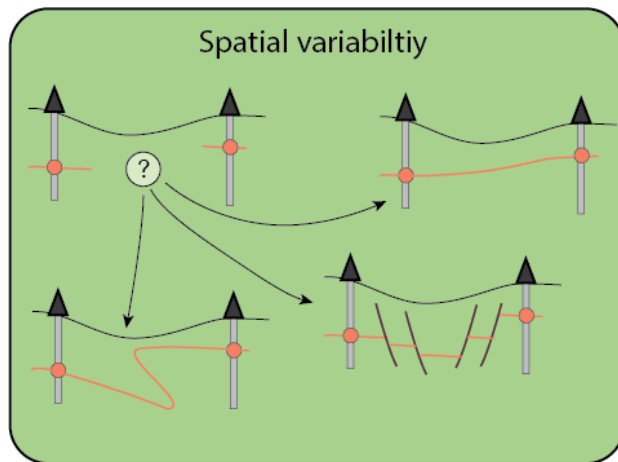
# Uncertainty Quantification from Seismic Data (Current Work)

MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

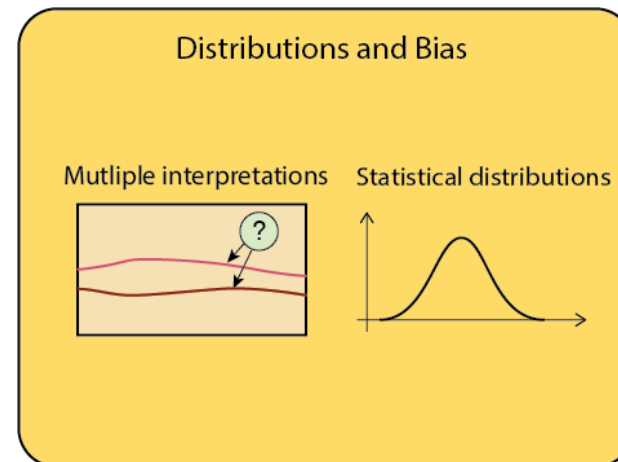
Set of case studies and associated data

Systematic identification & analysis of uncertainties in input data of geological models evidenced by the case studies.

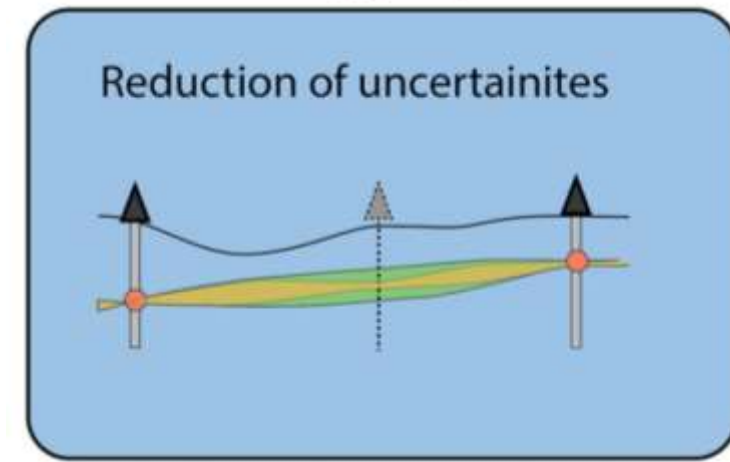
Analysis of the influence of subjectivity in the interpretation of geophysical data - interpretation PDF



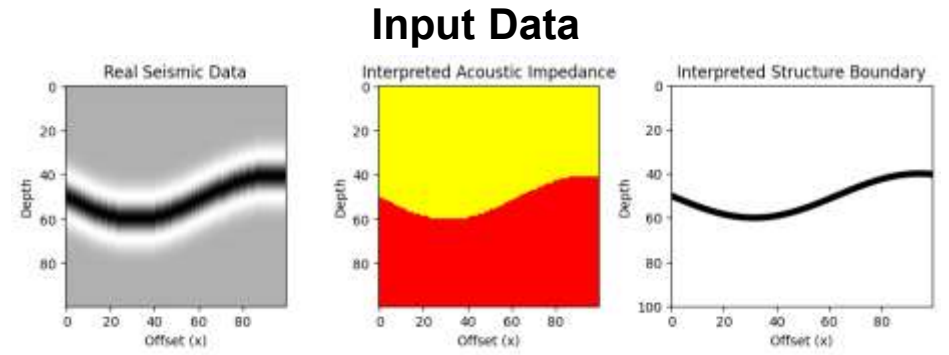
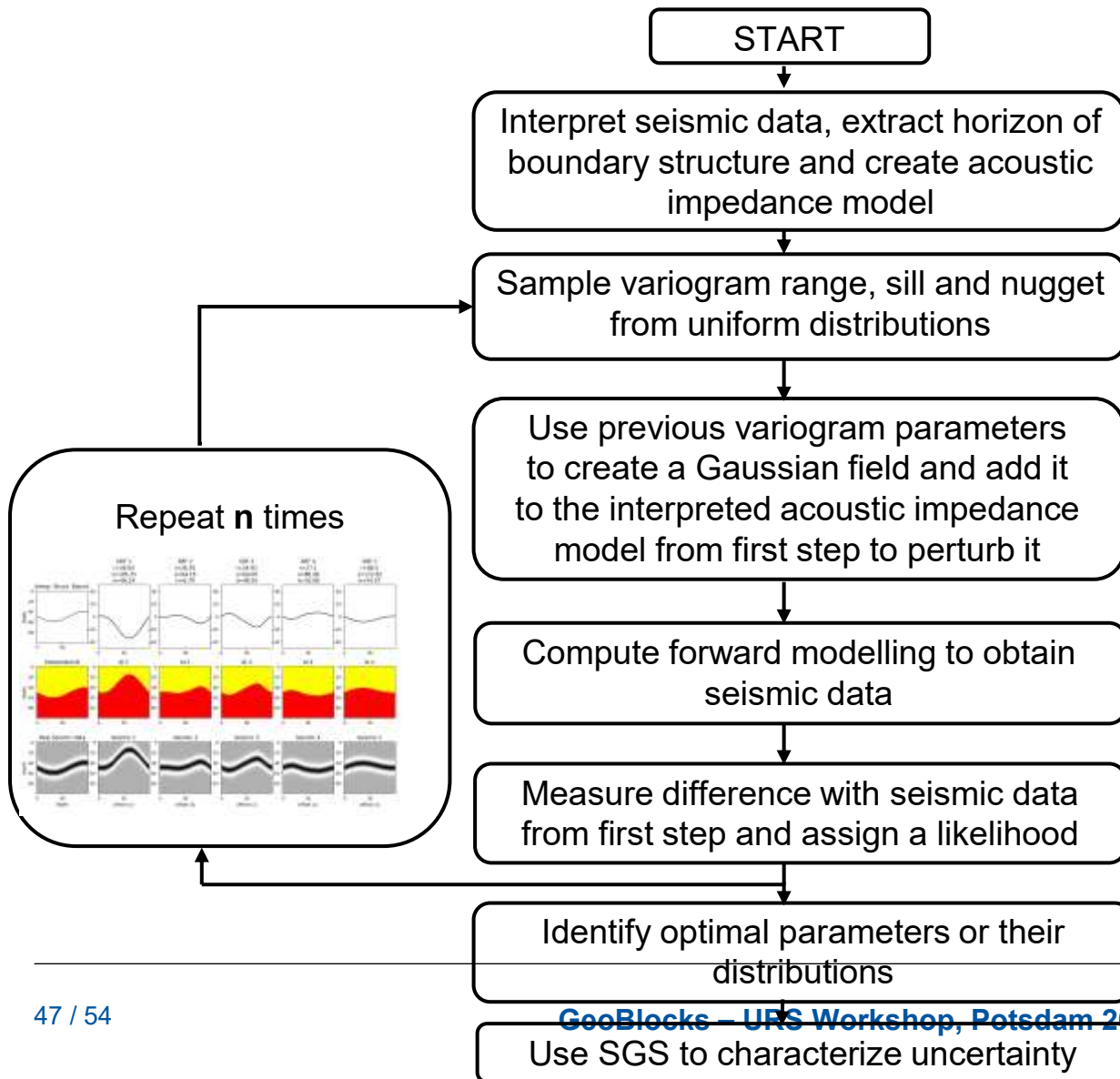
Interpolations –stochastic models



Interpretations – multiple realisations

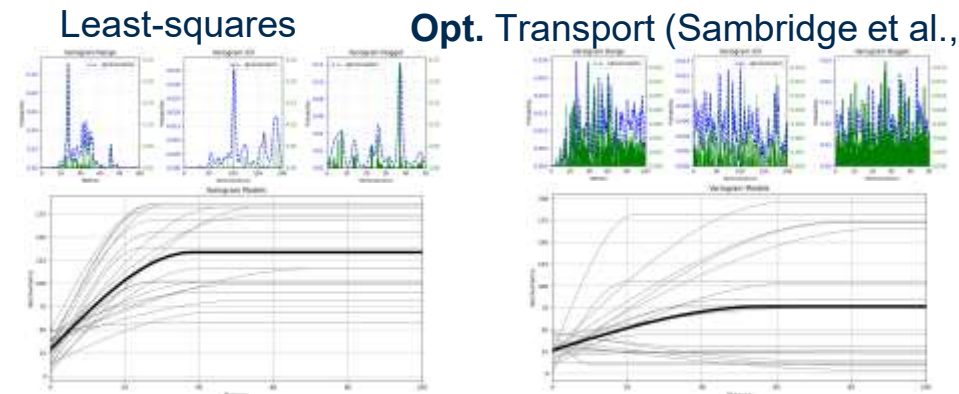


# Uncertainty Quantification from Seismic Data (Current Work)



Forward operator:  $S_i = \mathbf{W}S_i(\mathbf{d}, \theta = 0) = w(\mathbf{d}) * \frac{d \ln \mathbf{M}_i(\mathbf{d})}{dd}$

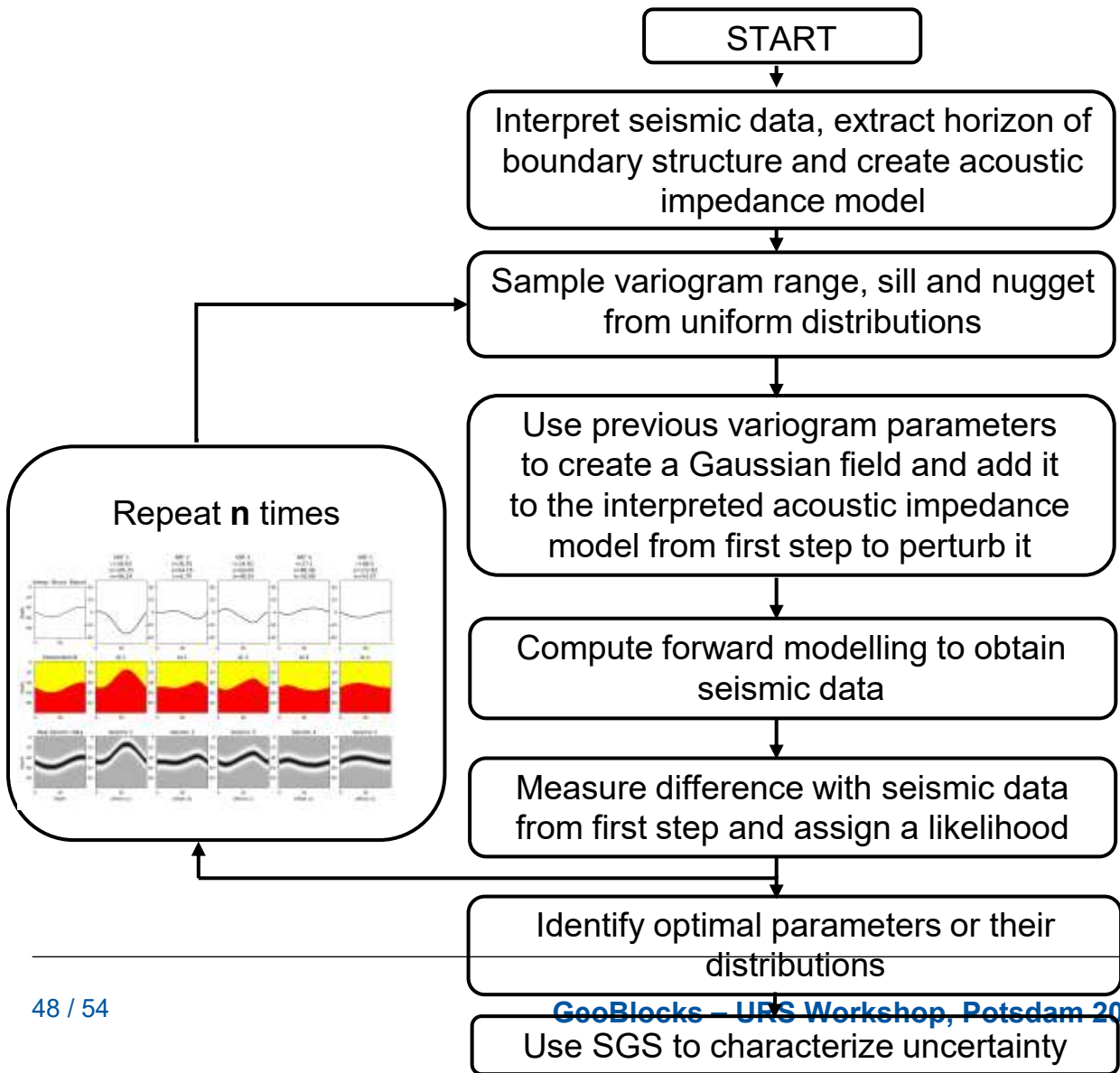
Different error measure methods being tested at the moment



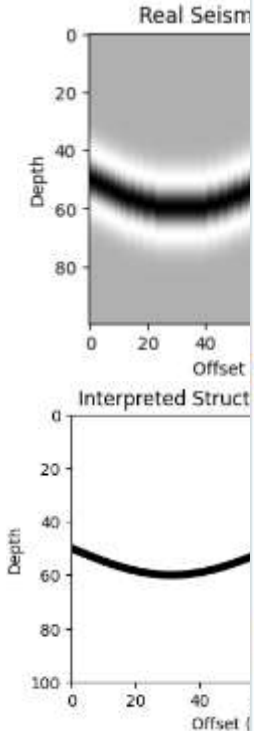
Also contributing to:

MS 3.1	Relevante Datentypen identifiziert	M6	MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 2.1	Verfügbare Daten gesichtet	M12	MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18	MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18
			MS 5.1	Probabilistische Modellierung in Workflow implementiert	M18

# Uncertainty Quantification from Seismic Data (C)



## Spatial un variogram p



### Seismic Forward Modelling to Characterize Spatial Uncertainties of Geological Structures

Carlos Colombo<sup>1</sup>, Peter Aichtziger<sup>2</sup> and Florian Wellmann<sup>1</sup>

<sup>1</sup>Chair of Computational Geosciences, GeoInformatics and Reservoir Geophysics (CGF), RWTH Aachen University, Aachen  
<sup>2</sup>Fraunhofer Research Institution for Energy Infrastructure and Geoelectronics, IEG, Aachen

---

#### Introduction

Spatial uncertainty is critical in the design of underground nuclear waste storage sites as they must comply with specific safety distances to the boundaries of the selected geological structures to hold the site. This uncertainty is larger in areas where data is not available and interpolation among measured data is needed. We propose a method to identify the optimal variogram parameters from available seismic data. Starting from the seismic interpreted model of impedances, we produce perturbations by adding several realizations of a Gaussian field (produced with different range, sill and nugget parameters). We simulate their corresponding seismic data with forward modeling. By comparing the error between the simulated and available seismic data, we are able to obtain the optimal variogram parameters to model the geological structure and use them to characterize its spatial uncertainty with sequential Gaussian simulation. The objective of the method is expected to identify the best locations (if any) to place a storage site and, especially, characterize the spatial uncertainty between hard-data measured points.

#### Method

##### 1 INPUT DATA

The analysis starts with the seismic data of the subsurface region whose uncertainty we expect to characterize. We call these data "real seismic data". At this stage of the study, we consider a 2D synthetic seismic section, in depth domain. We interpret the horizon it represents and create an acoustic impedance model.

Seismic data =  $S_{real}(d)$

Acoustic Impedance =  $M_{real}(d)$

where  $d$  is a point  $(x_d, z_d) \in \mathbb{R}^2$  with  $x_d = x_0, \dots, x_N$

##### 2 FORWARD MODELLING

By selecting  $n_{sim}$  (with  $n_{sim} = 10000$  for least-squares and  $n_{sim} = 500$  for optimal transport errors, Ravasi et al., 2020) uniformly distributed random samples of variogram parameters (range, sill and nugget), we create  $n_{sim}$  realizations of a Gaussian field and add them to the interpreted acoustic impedance model in step 1. We obtain  $n_{sim}$  additional impedance models, whose corresponding seismic data (post-stacked post-stack) are generated with forward modelling (PyLops Development Team, 2023).

$$S_j = WS_j(d, 0) = w(d) \cdot \frac{d \ln M_j(d)}{dt}$$

##### 3 ERROR MEASURE

We measure the error between the available seismic data and the 10000 forward-modelled realizations, and assign a likelihood value to each of them. After testing several methods, we use least-squares and optimal transport (Sambridge et al., 2022) to measure the errors. We obtain the probability distributions for each of the variogram parameters.

$$\text{Least Squares Error} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (S_{real}(d) - S_i(d))^2$$

Optimal Transport Error =  $\int_0^1 |ICDF_{real} - ICDF_i| dd$  where  $ICDF_{real}$  and  $ICDF_i$  are the inverse cumulative distribution functions of the real and modelled seismic data (Sambridge et al., 2022).

##### 4 VARIOGRAM MODEL

We construct the variogram models from the distributions of the parameters found in previous step and use them in the next step for a sequential Gaussian simulation.

$$\gamma(d) = \text{nugget} + (\text{sill} - \text{nugget}) \left( 1 - \exp\left(-\frac{d}{\text{range}}\right) \right)^2$$

##### 5 CHARACTERIZATION OF SPATIAL UNCERTAINTY WITH SEQUENTIAL GAUSSIAN SIMULATION

With the modelled variogram we perform sequential Gaussian simulation to characterize the uncertainty of the area, given the available seismic data.

#### Discussion & Future Work

- The distributions of the variogram parameters show periodicity that needs further investigation.
- Optimal transport provides a more detailed error calculation, however its results are less stable (even with a closer model) and the uncertainty does not follow the interpreted boundary; further investigation is needed.
- The range needs to be of similar magnitude as the width of the section, as larger values can produce similar, non-unique perturbations, resulting in many range values giving the same, indistinguishable results.
- The next step is to apply the method on real seismic data, where an inversion might be required instead of a single horizon interpretation. Further extension to 3D analysis is also planned.

#### References

- Jan F. Claerhout, November 17<sup>th</sup>, 2010. Basic Earth Imaging.
- Malcolm Sambridge, Andrew Jackson, and Andrew P. Valentine, June 21<sup>st</sup>, 2022. Geophysical Inversion and Optimal Transport. Geophysical Journal International 231, n. 1: 172-98. <https://doi.org/10.1093/gji/ggac219>
- Matteo Ravasi and Ivan Vasconcelos, January 2020. PyLops—A Linear-Operator Python Library for Scalable Algebra and Optimization. SoftwareX 11: 100361. <https://doi.org/10.1016/j.softx.2019.100361>
- PyLops Development Team, 2023. Post-stack Linear Modelling (pylops.pyr.poststack.PoststackLinearModelling). <https://github.com/PyLops/pylops/blob/master/pylops/pyr/poststack/PoststackLinearModelling.py> (last accessed January 24<sup>th</sup>, 2025).

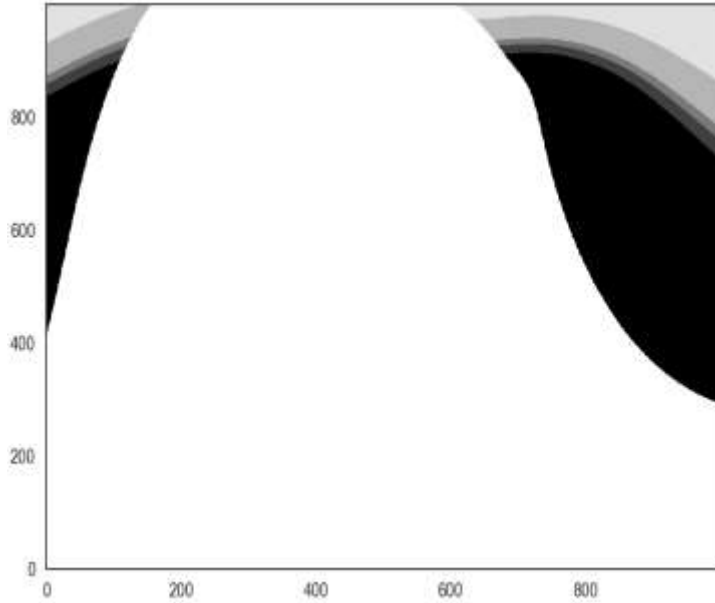
48 / 54

GeoBlocks – URS Workshop, Potsdam 2025

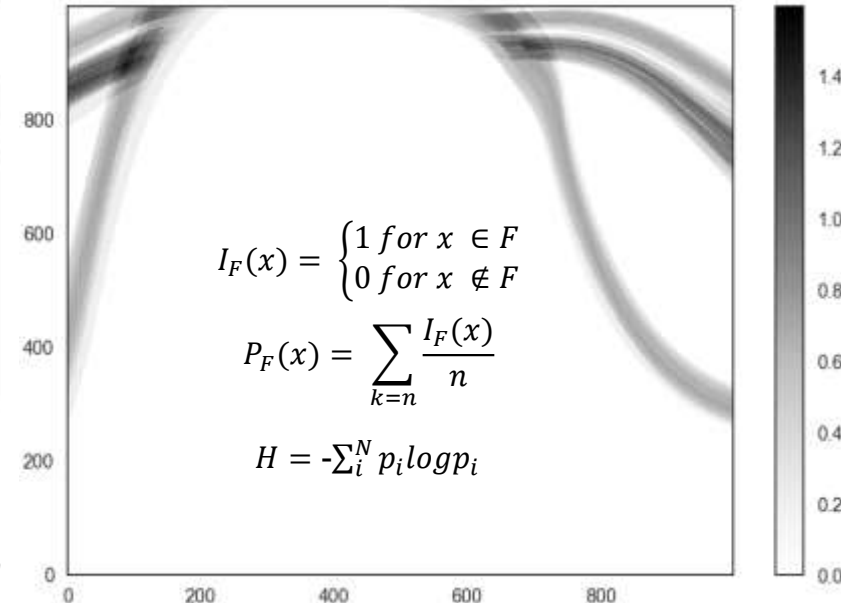
MS 3.1	Rele	M18	MS 5.1	Probabilistische Modellierung in Workflow implementiert	M18
MS 2.1	Verf				
MS 2.2	Repräsentative Modelle ausgewählt/erstellt				

# Uncertainty Quantification with Gaussian Processes and Information Entropy

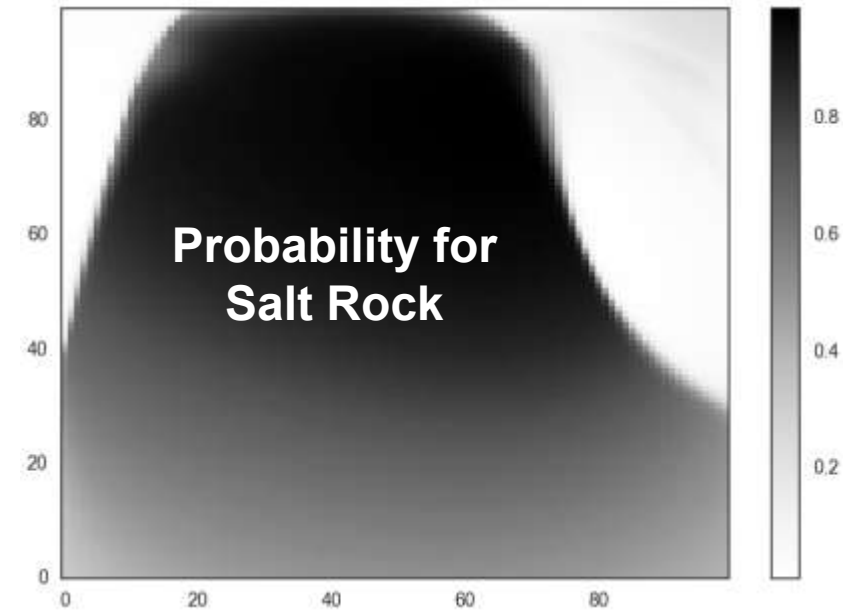
## Salt Dome Structure



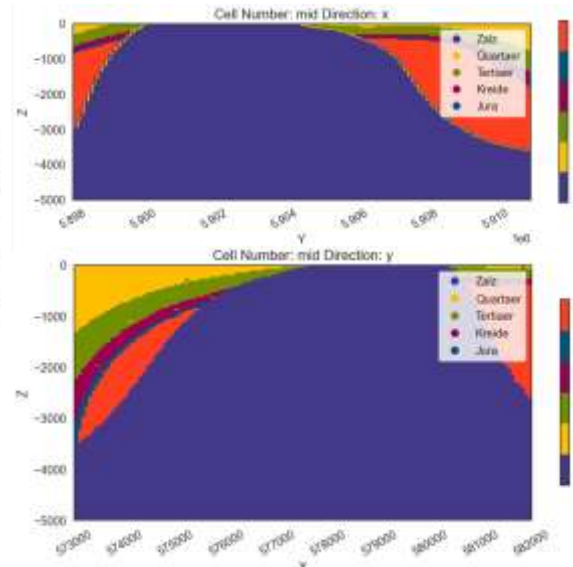
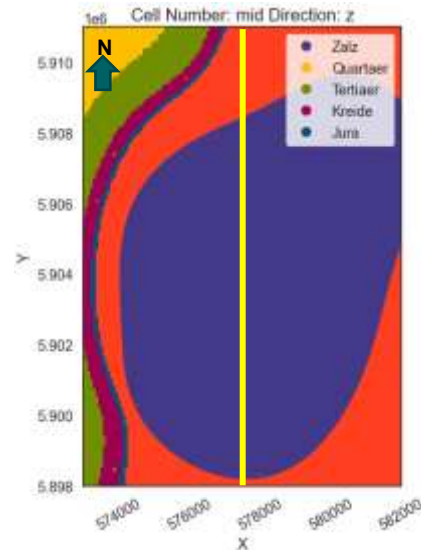
## Information Entropy



## Gaussian Process Classifier



Formation	Perturbation (m) <sub>N</sub>
Quartaer	100
Tertiaer	100
Kreide	100
Jura	100
Salz	200



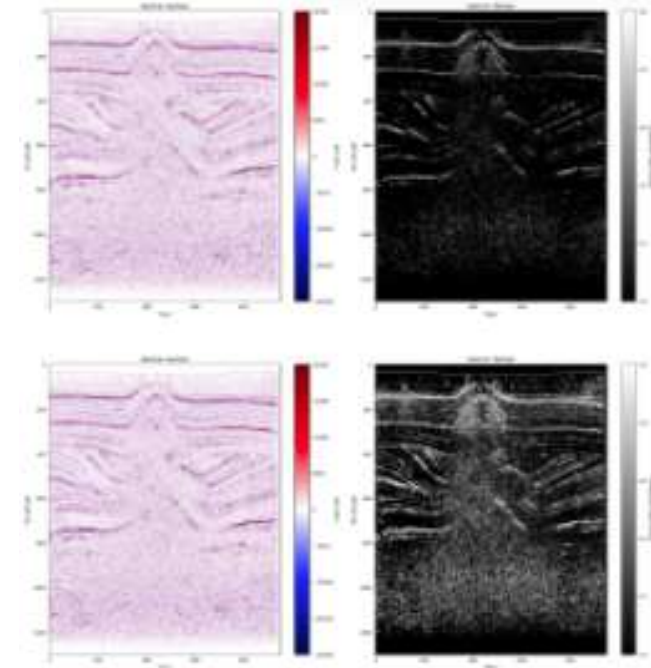
Also contributing to:

MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18
MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18

## Future Plans (considering a timeline until 2026)

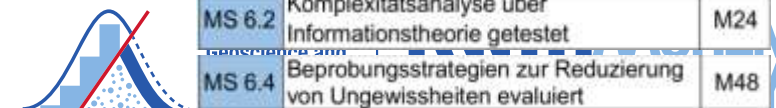
- Analysis of **seismic attributes**: some testing has already been done focused on the detection of salt rock; the goal is to contribute to the comparison of reprocessed seismic data.
- Incorporate **seismic interpretation in uncertainty characterization**.
- Consider **uncertainty quantification from seismic data acquisition, processing and velocity models**.
- **Replication of seismic workflows** in other locations and looking at other specific targets (lithologies, structural and stratigraphic features).
- **Comparison** of results between **onshore and offshore** surveys (e.g. North Sea x North Hannover).

### Seismic Attribute Testing with PyLops



Input to:

MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30
MS 5.2	Kennwerte zur Vergleichbarkeit von Ungewissheiten definiert	M24
MS 5.3	Modellensembles erzeugt und analysiert	M36
MS 5.4	Quantitativer Vergleich von Prognosen der Ungewissheiten implementiert.	M48
MS 6.1	Einfluss der Eingangsdaten auf Modelle quantifiziert (Sensitivitätsanalysen)	M18
MS 6.2	Komplexitätsanalyse über Informationstheorie getestet	M24
MS 6.4	Beprobungsstrategien zur Reduzierung von Ungewissheiten evaluiert	M48

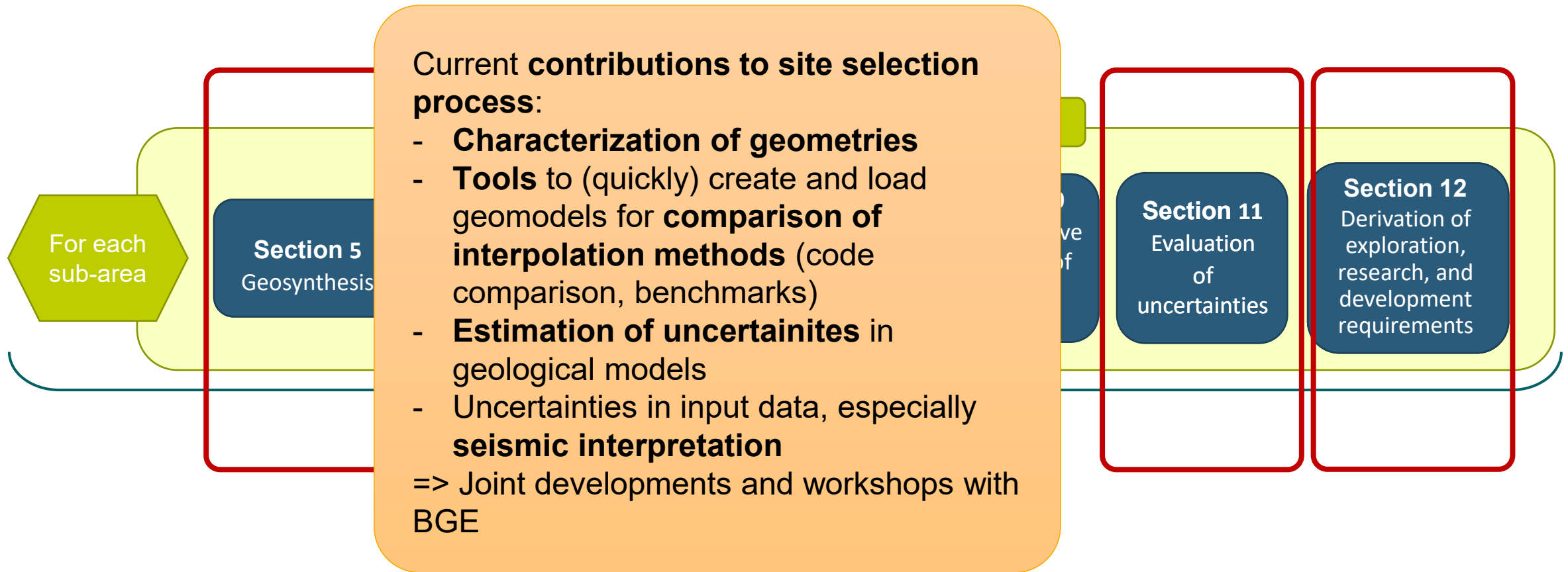




# GeoBlocks project workflow - summary

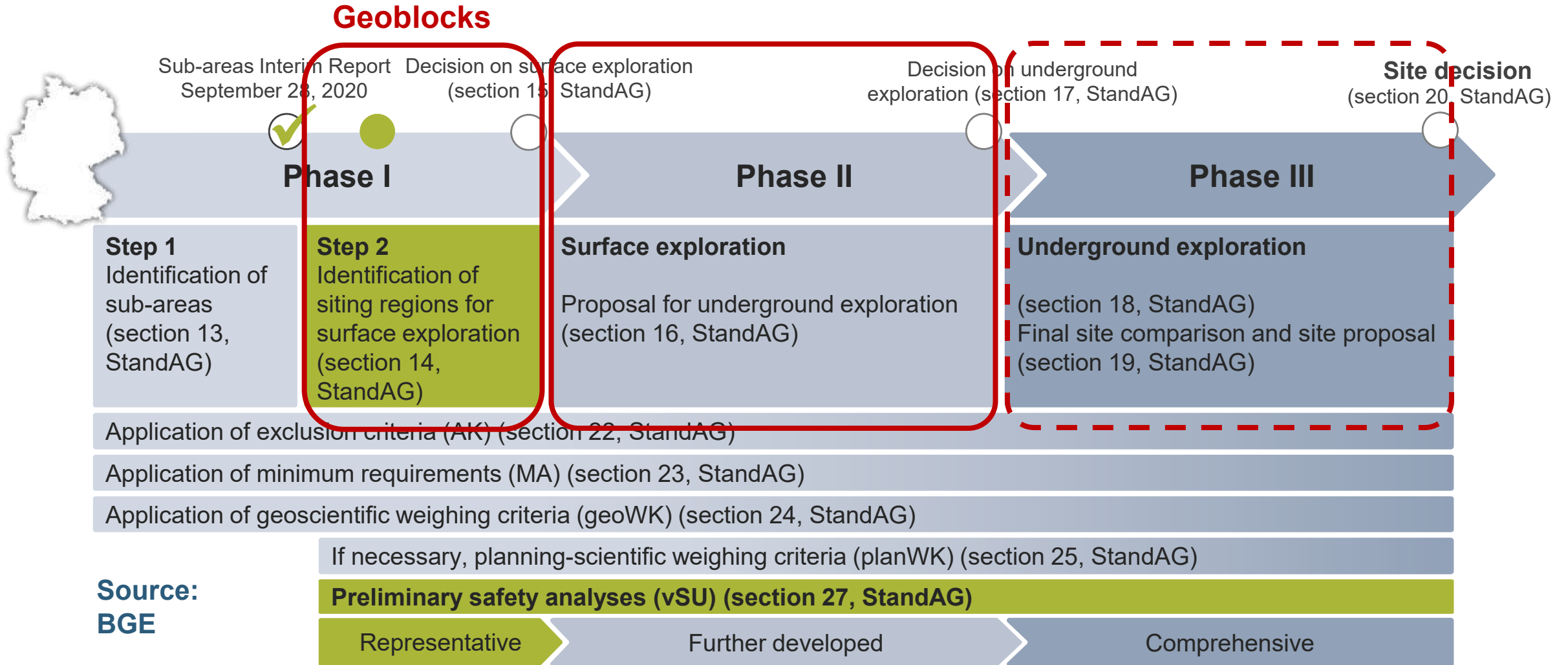


# The representative preliminary safety assessments (rvSU)

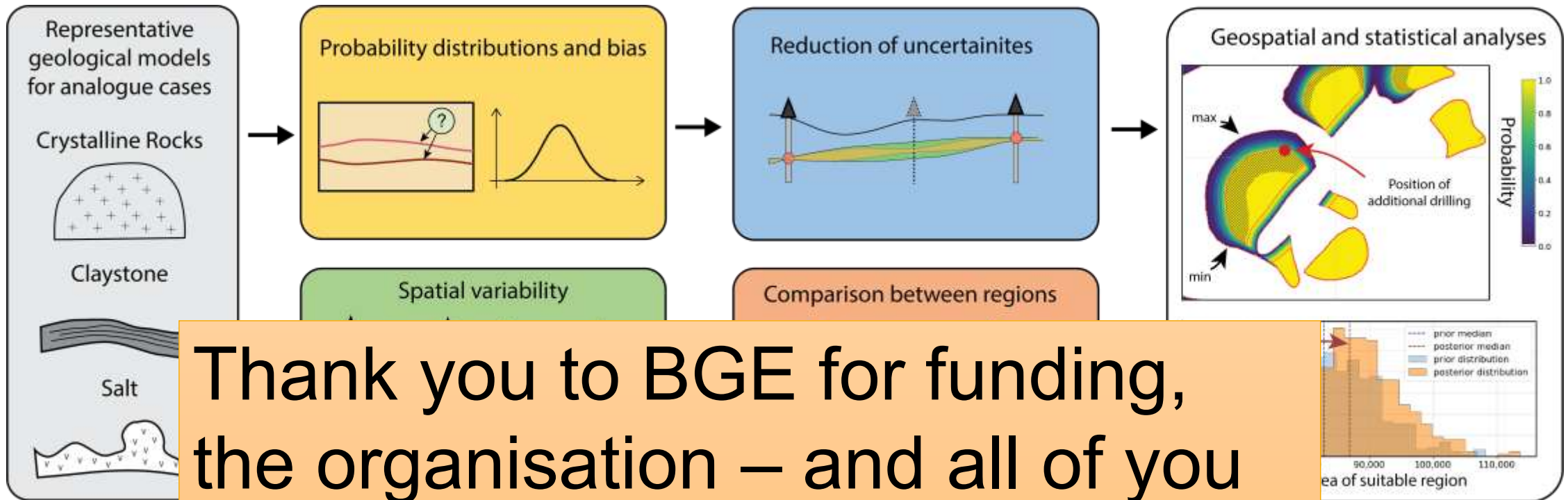


Step-by-step narrowing down of sub-areas to site regions using criteria-based approach

# Where does the project fit in?



# GeoBlocks Project



Thank you to BGE for funding, the organisation – and all of you for your attention (before lunch)!

- 
- Practical testing of interpolation workflow (with data and conceptual models from BGE)
  - Extension of catalogue of geometries
  - Quantitative comparison of geometric models
  - Bias in input data: seismic interpretation
  - Link to complexity: how much data required to characterize specific uncertainty?

How robust is interpolation of geometry (with limited data)?

Which geometric settings matter?

Which geometric settings matter?c

Which information to rely on? (and which not?)

Comparison between regions with different data density

# Conclusions

---

- The inclusion of different types of data is key to quantify uncertainties (e.g., seismic, boreholes, gravity, magnetic and/or resistivity) and plays a major role when the approach is probabilistic.
- The reduction of the spatial uncertainty in the spatial definition of the geometrical structures is based on model adaptation to complementary information (in the probabilistic approach) and conditioning to other data (e.g., borehole) in the seismic data analysis.
- The probabilistic approach can be extended to seismic data by considering, for example, probabilistic velocity models.
- Uncertainties arisen during acquisition and processing of seismic data are important sources that could be incorporated eventually.
- The models generated in the uncertainty characterization with variogram parameters could eventually be used to train neural networks and use them to characterize the uncertainty.

# Work Objectives

---

- Main objective of the whole work:

Characterize spatial uncertainty, especially between known points, of geological models for site selection process.

- **Specific objective for this milestone group: to build probabilistic models to be applied in uncertainty characterization.**

- Expected contribution to site selection:

Provide a method for spatial uncertainty in geological model that shows locations where the uncertainty is minimum, or where a specific host rock is most probable.

- Input Data:

Seismic data

Velocity models

Seismic raw data

Borehole data

Geological maps and reports

Other geophysical and well data

# Ongoing Project Work

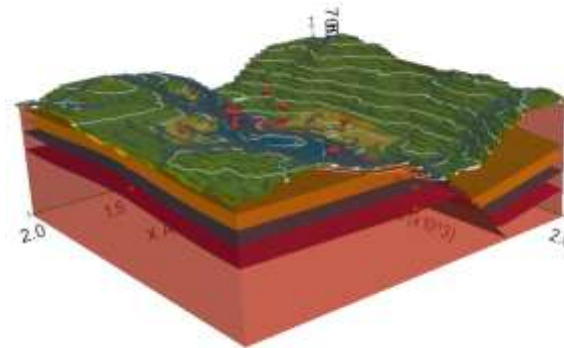
- Revision of geomodelling concepts and interpolation methods
- Review Bayesian inversion
- Started building general GemPy models
- Write draft of project data management plan

Covering so far project description, requirements, FAIR data, research outputs, allocation of resources, storage, backup, naming, data security, ethics, metadata

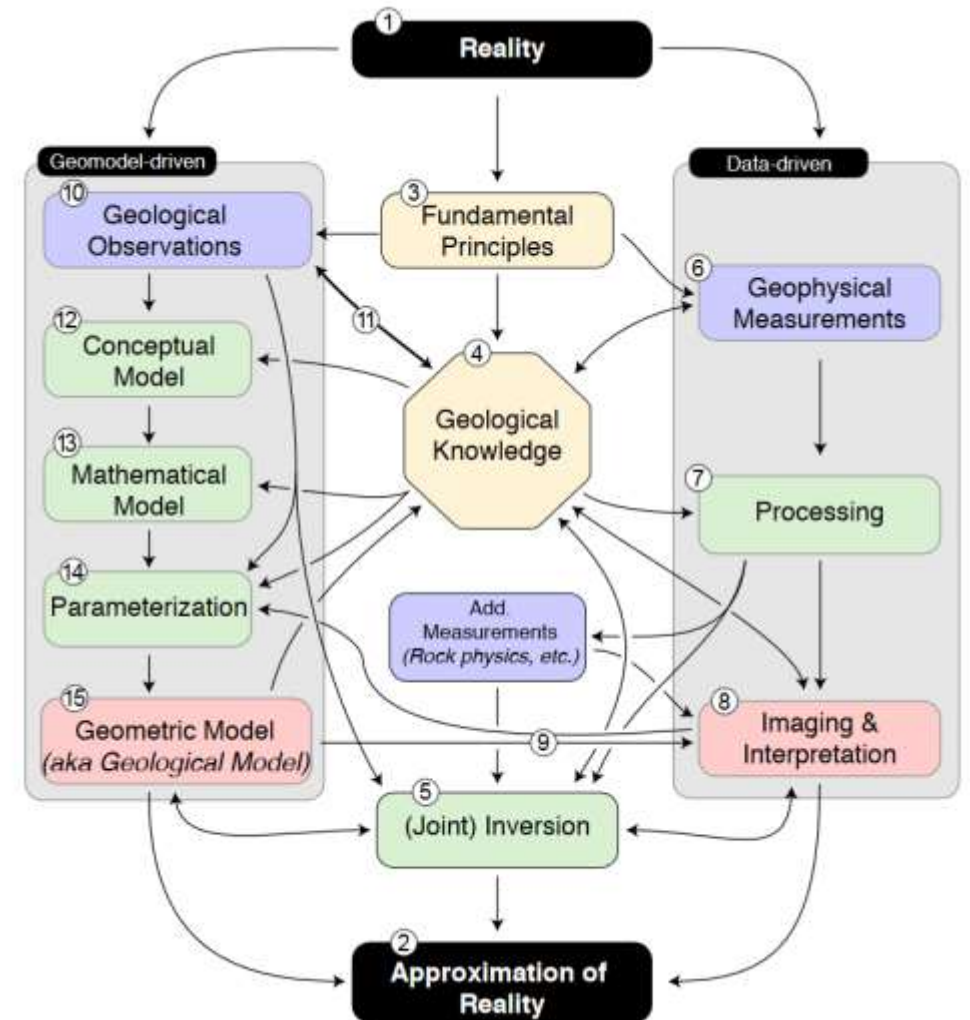
- Built GIS Database

Also contributing to:

MS 1.1	Plan für Forschungsdatenmanagement erstellt	M6
MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18
MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18



<https://docs.gempy.org/index.html>





# Uncertainty Quantification and Machine Learning Literature Review

## Epistemic Uncertainty

State/lack of knowledge, can be reduced with observations

## Aleatory Variability or Uncertainty

Unpredictability due to inherent randomness, cannot be reduced

Witter et al. 2018, Wellmann and Caumon 2018

Classification of Wellmann and Caumon, 2018, after Mann 1993 and Cox 1982:

### Type 1 - Error, bias, and imprecision

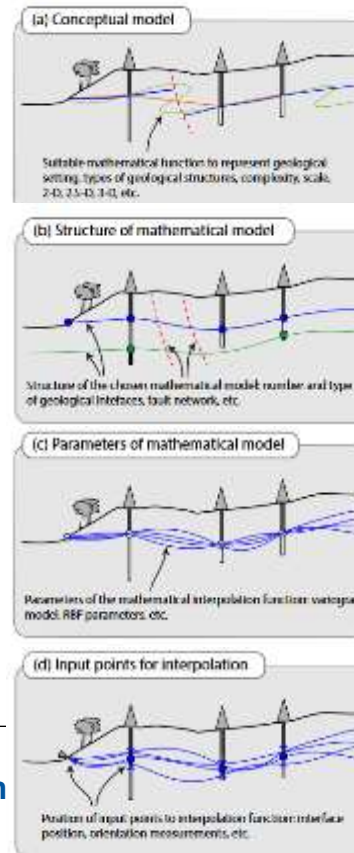
Ambiguity of structure based on uncertainties in raw data

### Type 2 - Stochasticity, and inherent randomness

Uncertainty of interpolation and extrapolation away from know points

### Type 3 - Imprecise knowledge

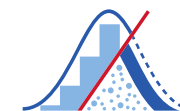
Problem of incomplete knowledge of structures in subsurface



## Reviewed Concepts and Methods:

- Bayesian Inference
- Approximate Bayesian Computation
- Geological Model-based Uncertainties
- Information Entropy (Shannon Entropy)
- Gaussian Processes
- Sequential Gaussian Simulation
- Monte Carlo (Model Based)
- Markov chain Monte Carlo
- Uncertainty from Seismic Data
- Spatial Diffusion
- Neural-Network-based Multi-step Prediction
- Intervals Construction Methods
- Neural-Network-based Direct Prediction Intervals Construction Methods
- Variational inference
- Bayesian Active Learning (BAL)
- Bayes By Backprop (BBB)
- Variational autoencoders
- Laplacian approximations
- Uncertainty quantification in reinforcement learning
- Ensemble methods
- Bayesian Neural Networks

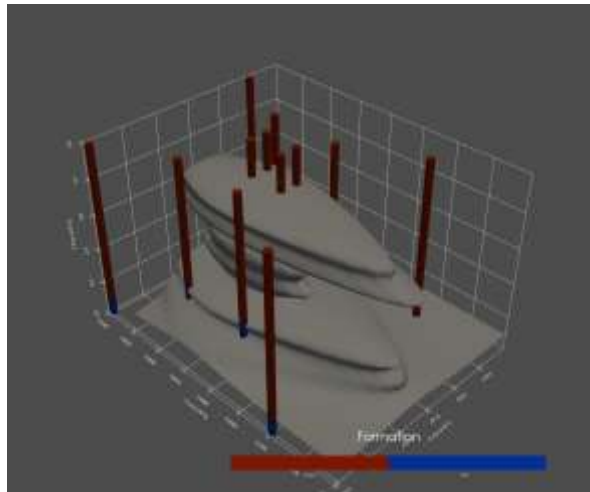
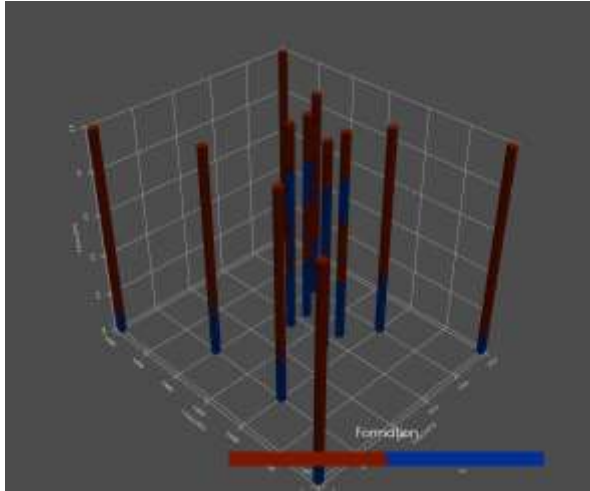
Also contributing to:



Co	MS 3.1	Relevante Datentypen identifiziert	M6
Ge			
Res			
Ent	MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18

# Variational Gaussian Processes for Uncertainty Quantification

## Synthetic salt dome



$$z(x) = f(x) + \xi$$

$$p(z^* | \mathbf{z}) = N(z^*; \mathbf{m}^*, \sigma^{*2})$$

$$\mathbf{m}^* = \mathbf{K}^* \mathbf{f} (\mathbf{K}^* \mathbf{f} \mathbf{f}^T + \sigma^2 \mathbf{I}N) - 1 \quad \mathbf{z}$$

$$\sigma^{*2} = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}^* \mathbf{f} (\mathbf{K}^* \mathbf{f} \mathbf{f}^T + \sigma^2 \mathbf{I}N) - 1 \quad \mathbf{k}^* \mathbf{f}^T + \sigma^2$$

$$\log p(\mathbf{z}) = -\frac{1}{2} \mathbf{z}^T (\mathbf{K}^* \mathbf{f} \mathbf{f}^T + \sigma^2 \mathbf{I}N) - 1 \quad \mathbf{z} - \frac{1}{2} \log |\mathbf{K}^* \mathbf{f} \mathbf{f}^T + \sigma^2 \mathbf{I}N| - \frac{N}{2} \log 2\pi$$

$$p(\mathbf{f}, \mathbf{u}) = N \left( \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \end{bmatrix}; \mathbf{0}, \begin{bmatrix} \mathbf{K}^* \mathbf{f} \mathbf{f}^T & \mathbf{K}^* \mathbf{f} \mathbf{u}^T \\ \mathbf{K}^* \mathbf{u} \mathbf{f}^T & \mathbf{K}^* \mathbf{u} \mathbf{u}^T \end{bmatrix} \right)$$

$$p(\mathbf{z}, \mathbf{f}, \mathbf{u}) \propto p(\mathbf{z} | \mathbf{f}, \mathbf{u}) p(\mathbf{f}, \mathbf{u}) = p(\mathbf{z} | \mathbf{f}) p(\mathbf{f} | \mathbf{u}) p(\mathbf{u})$$

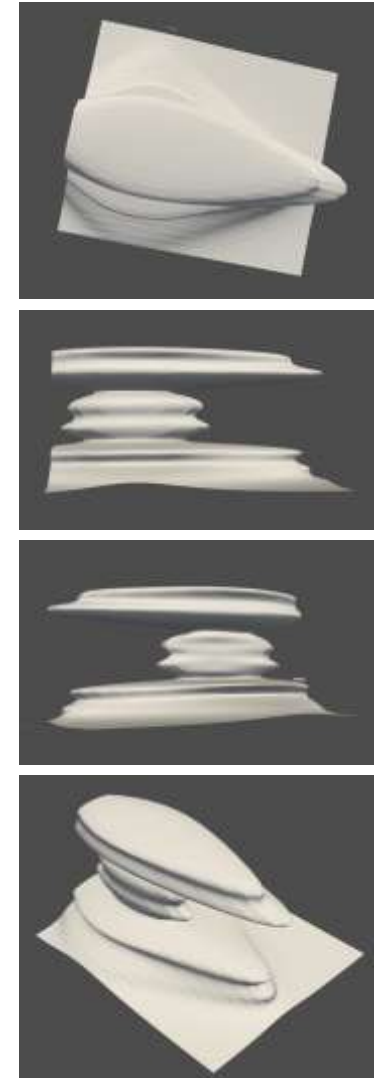
to maximize  $p(\mathbf{z})$ ,  $p(\mathbf{f} | \mathbf{u})$  is replaced by  $q(\mathbf{f} | \mathbf{u})$

Variational inference approximates the posterior as a Gaussian  $q(\mathbf{f}, \mathbf{u}) = p(\mathbf{f} | \mathbf{u}) q(\mathbf{u})$  by introducing the variational distribution

$q(\mathbf{u}) = N(\mathbf{u}; \mathbf{m}, \mathbf{S})$ , whose mean  $\mathbf{m}$  and covariance matrix  $\mathbf{S}$  are variational parameters to optimize

$$\log p(\mathbf{z}) \geq \mathbb{E} q(\mathbf{f}) [ \log p(\mathbf{z} | \mathbf{f}) ] - \text{KL} [ q(\mathbf{u}) || p(\mathbf{u}) ] = \text{ELBO}$$

Gonçalves et al. 2022



Also contributing to:

MS 2.1	Verfügbare Daten gesichtet	M12
--------	----------------------------	-----

MS 3.1	Relevante Datentypen identifiziert	M6
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18

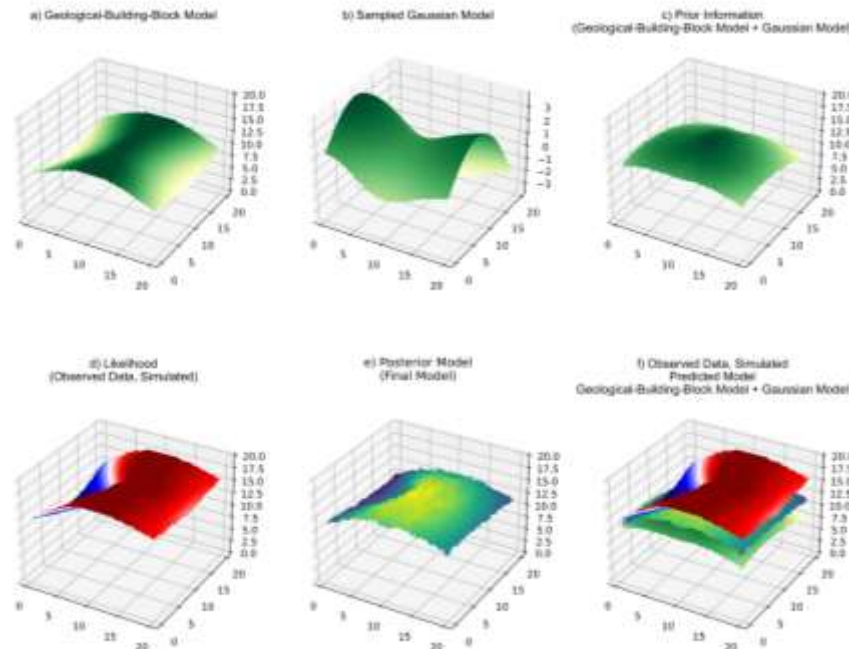
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18
MS 4.1	Quantitativer Vergleich Interpolationsmethoden für Strukturmodell	M18

# Probabilistic Modelling for Uncertainty Quantification



**GemPy:** Miguel de la Varga and J. Florian Wellmann. 2016. Structural geologic modeling as an inference problem: A Bayesian perspective. *Interpretation* 2016; 4 (3): SM1-SM16.  
**ODSIM:** Vincent Henrion, Guillaume Caumon and Nicolas Cherpeau. 2010. ODSIM: An Object-Distance Simulation Method for Conditioning Complex Natural Structures. *International Association for Mathematical Geosciences*.

- Incorporate synthetic geological model from catalogue
- Create Gaussian surface within the uncertain volume
- Add previous surfaces to get the mean of the prior model
- Incorporate observations as likelihood
- Adaptation, posterior model



Also contributing to:

MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18
MS 5.1	Probabilistische Modellierung in Workflow implementiert	M18