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 Parnters:
Frithjof Bense, Fabian Jähne-Klingberg, Heidrun Stück, Björn Zehner† (BGR), Michael Kettermann (IEG)

URS Workshop – Potsdam 05.02.-07.02.2025

Where does the project fit in?

Geoblocks

Sub-areas Interim Report September 28, 2020

Decision on surface exploration (section 15, StandAG)

Decision on underground exploration (section 17, StandAG)

Site decision (section 20, StandAG)

Phase I		**Phase II**	**Phase III**
Step 1 Identification of sub-areas (section 13, StandAG)	**Step 2** Identification of siting regions for surface exploration (section 14, StandAG)	**Surface exploration** Proposal for underground exploration (section 16, StandAG)	**Underground exploration** (section 18, StandAG) Final site comparison and site proposal (section 19, StandAG)

Application of exclusion criteria (AK) (section 22, StandAG)

Application of minimum requirements (MA) (section 23, StandAG)

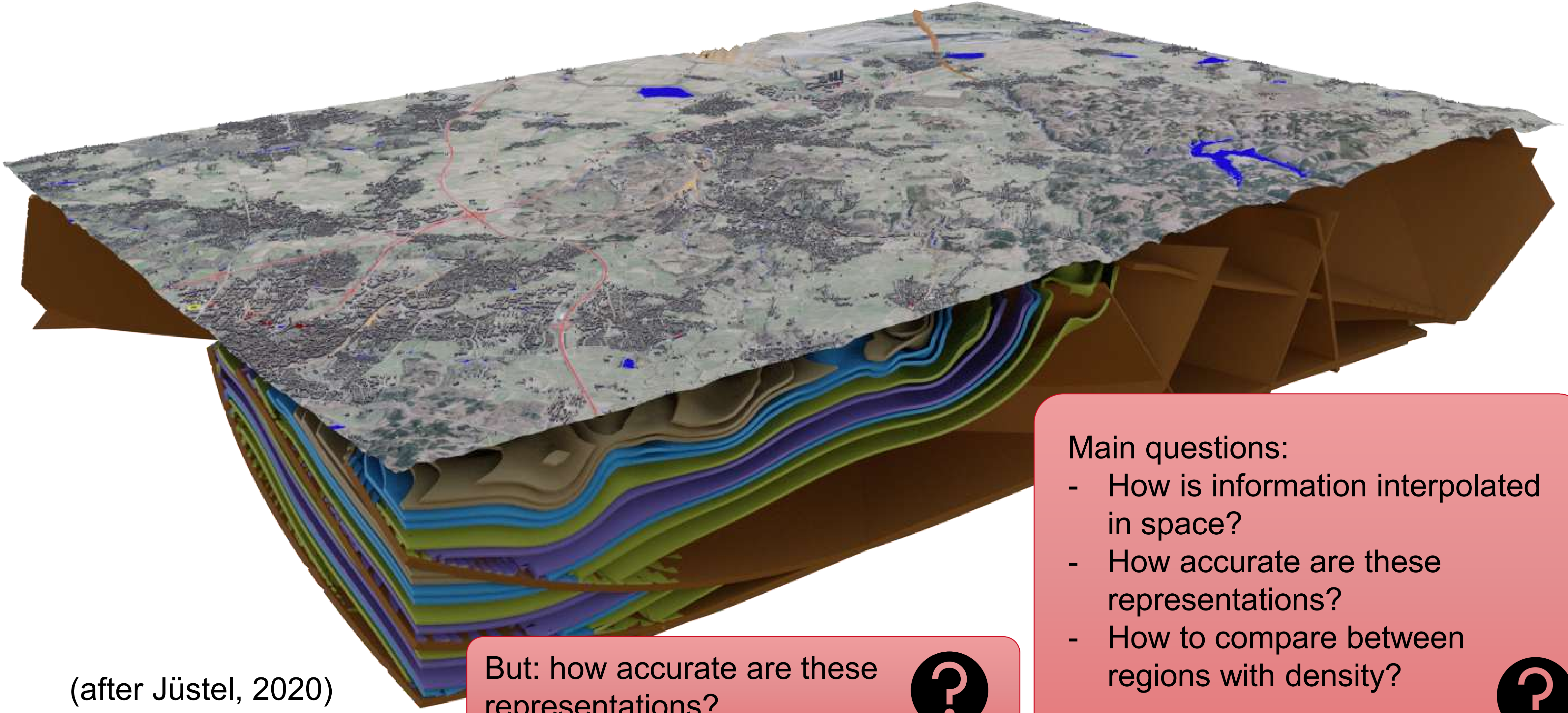
Application of geoscientific weighing criteria (geoWK) (section 24, StandAG)

If necessary, planning-scientific weighing criteria (planWK) (section 25, StandAG)

Preliminary safety analyses (vSU) (section 27, StandAG)

Representative → Further developed → Comprehensive

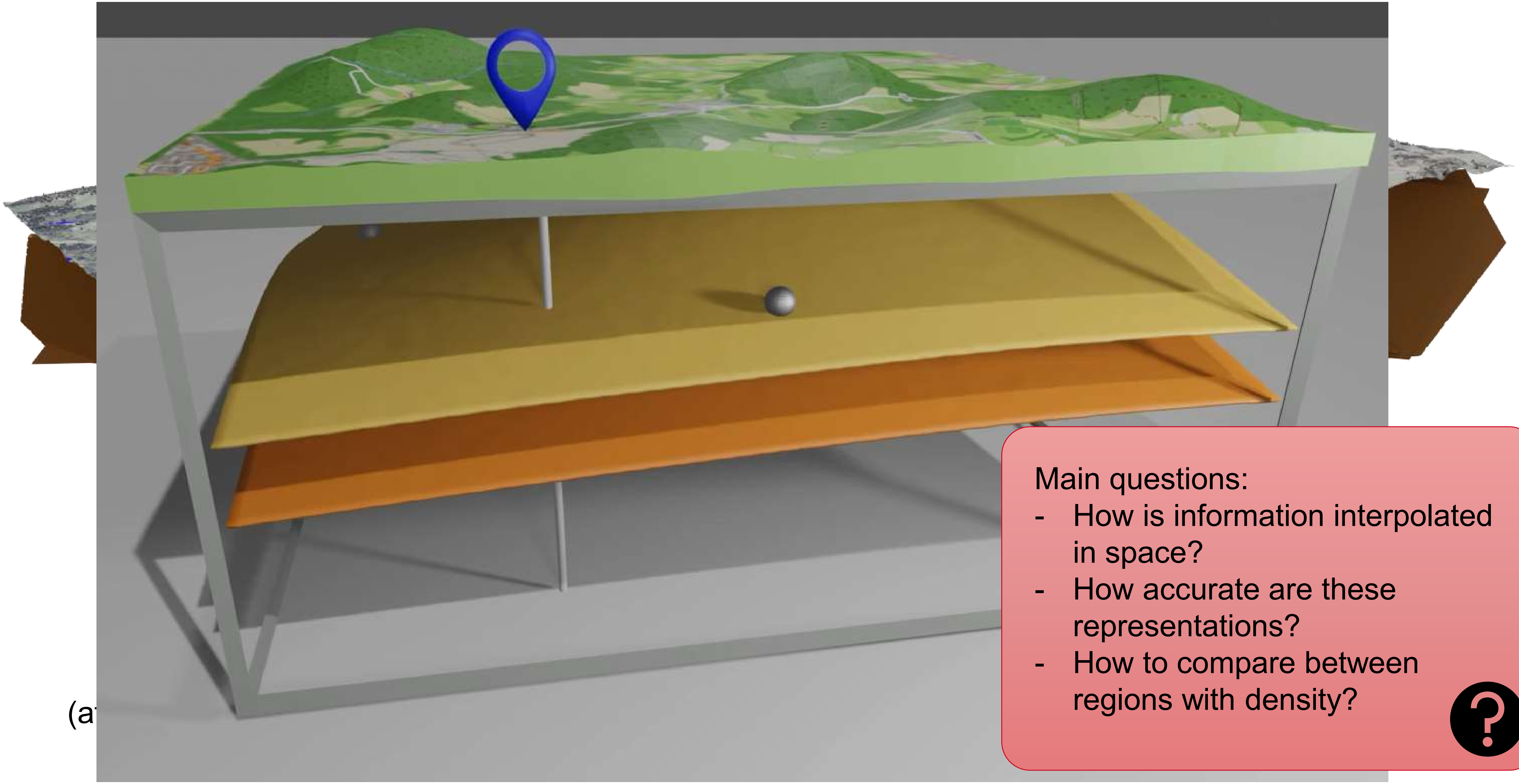
Source: BGE



(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?



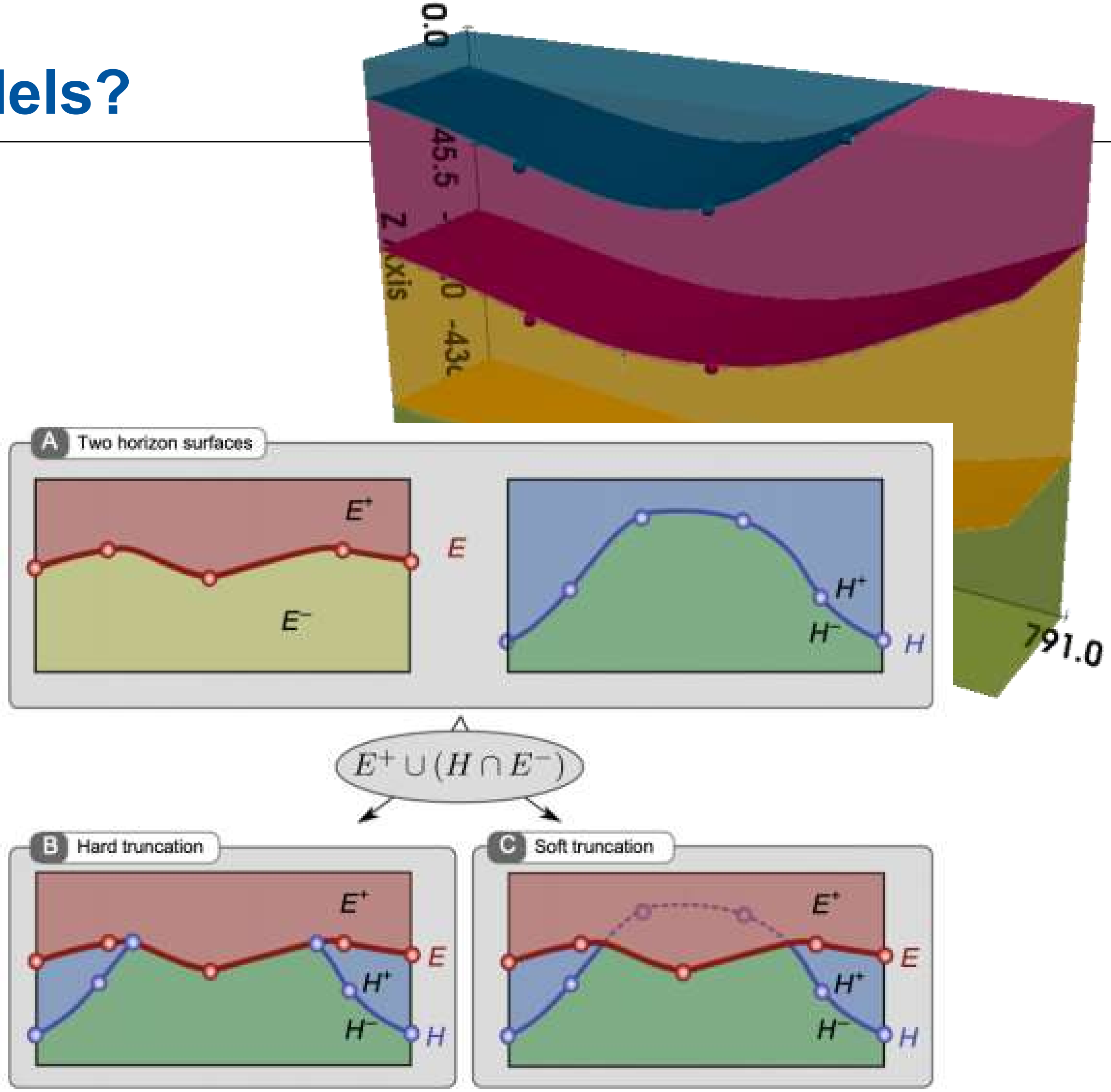
(a

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) paramterizations**
- 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**



Computational Geoscience and Reservoir Engineering
RWTH AACHEN UNIVERSITY

How to construct 3-D geometric models?

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

GemPy

https://www.gempy.org

GemPy: Flexible implicit geological modeling and uncertainty quantification

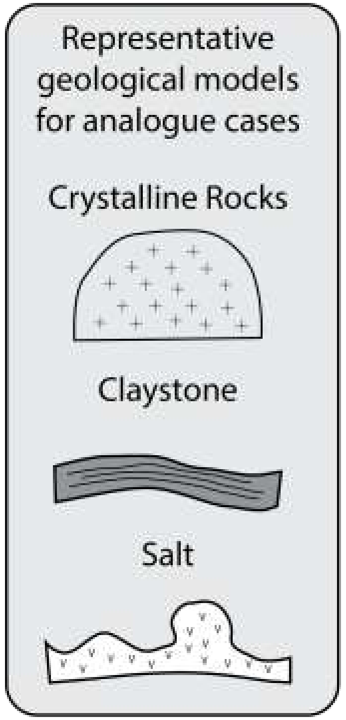
Open-source geomodeling
(www.gempy.org), (de la Varga, Schaaf & Wellmann, 2019)

Sponsor | Edit Pins | Unwatch 52 | Fork 227 | Starred 886

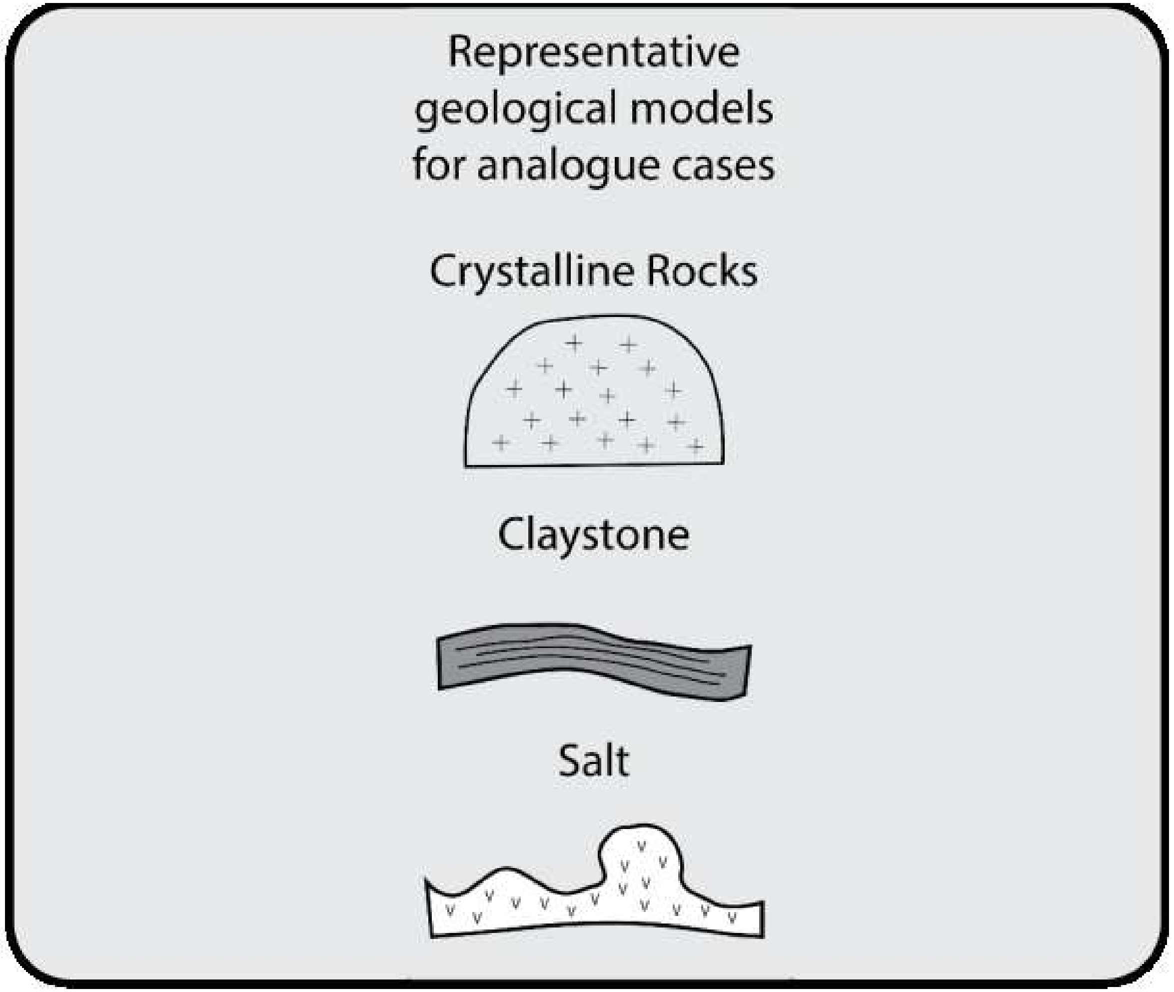
Computational Geoscience and Reservoir Engineering

RWTH AACHEN UNIVERSITY

Geoblocks project overview



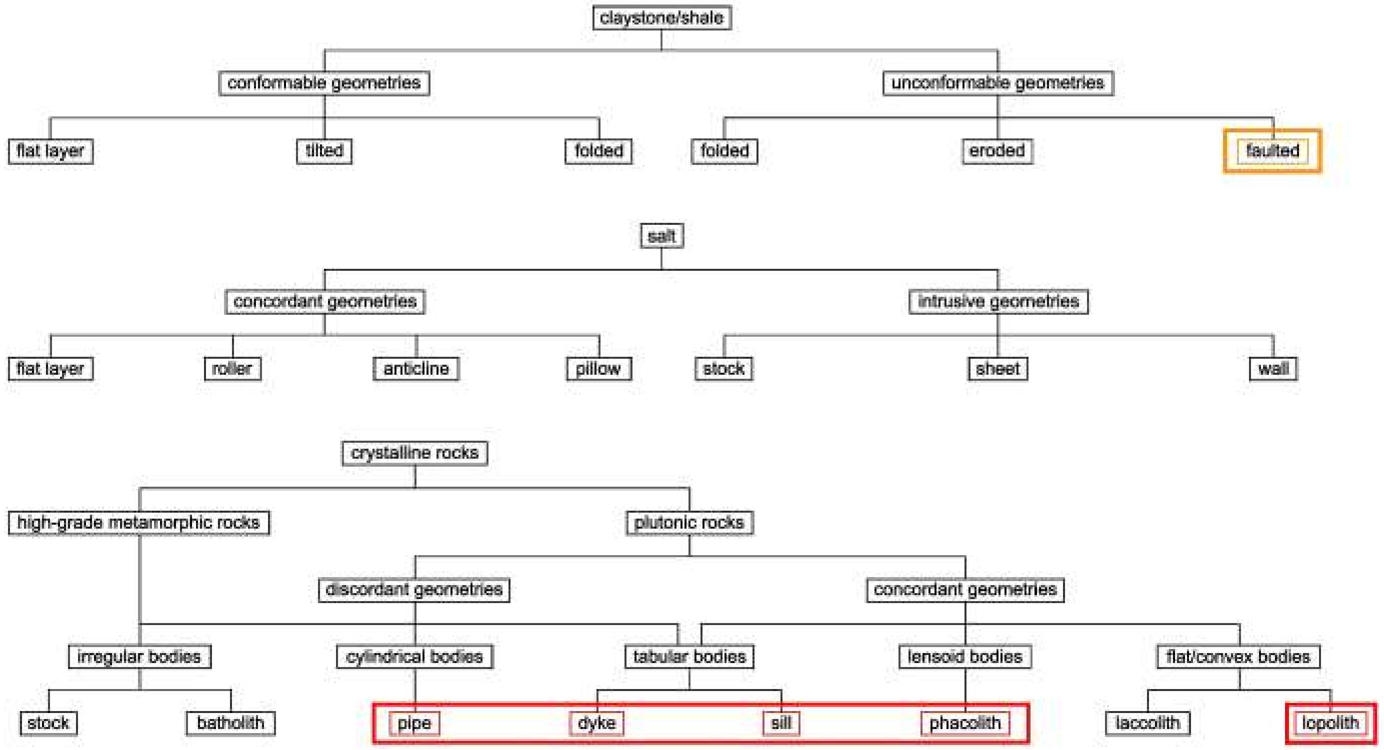
Part 1: Representative models and quantitative comparison of explicit geomodels



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

MS 2.1	Verfügbare Daten gesichtet	M12
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18



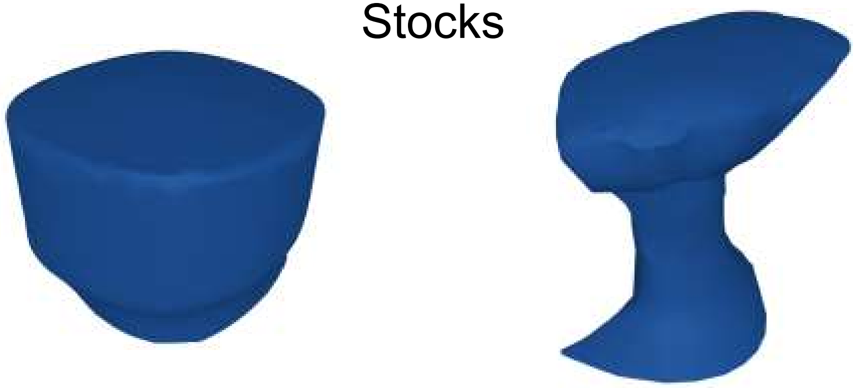
Database of decision-relevant geometries

MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18

- 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



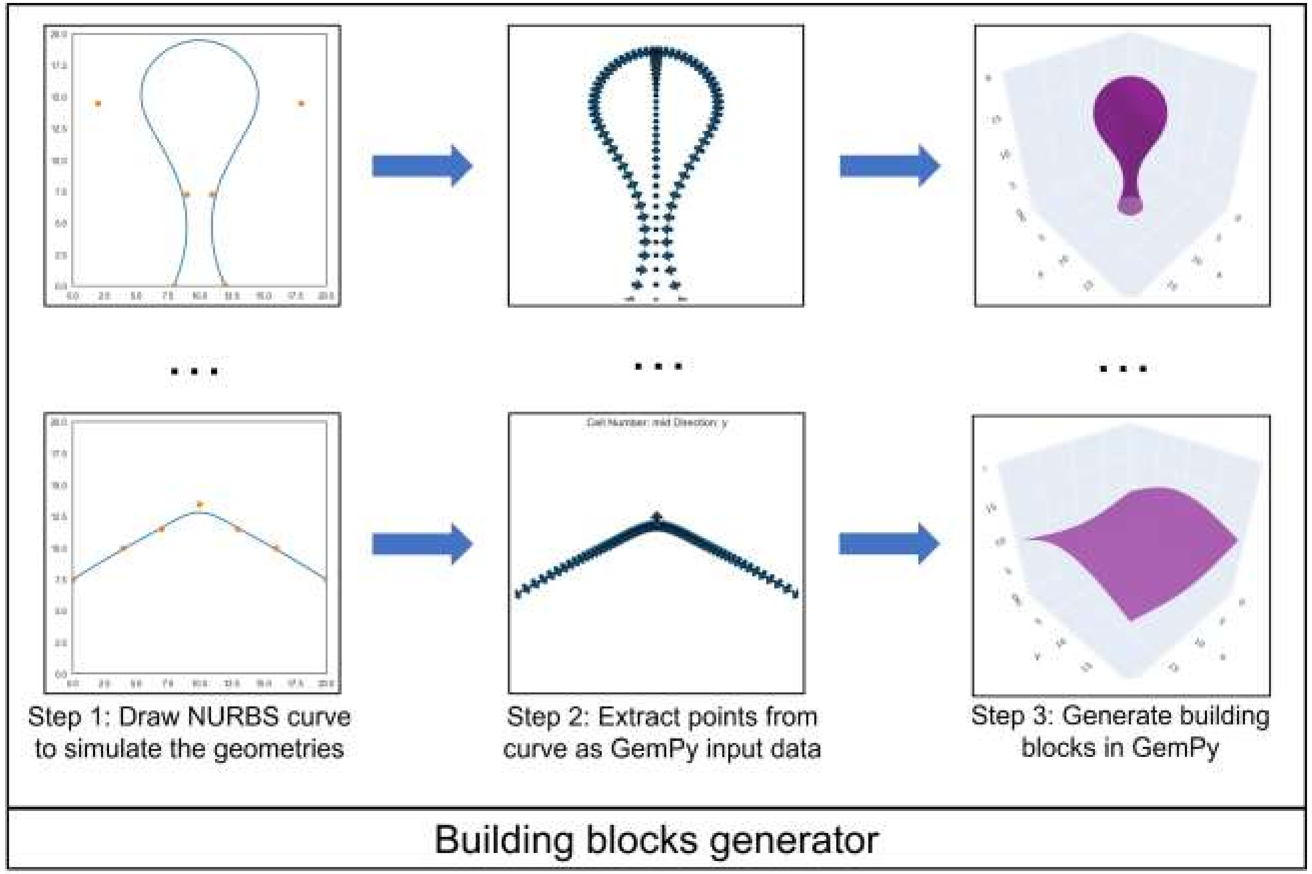
Walls

cylindrical | hourglass-shape | moderately anisotropic | highly anisotropic

Standard geometries and model generator V1.0

MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18

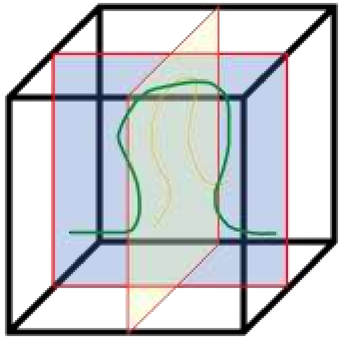
GemPy model generator V1.0



- 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:

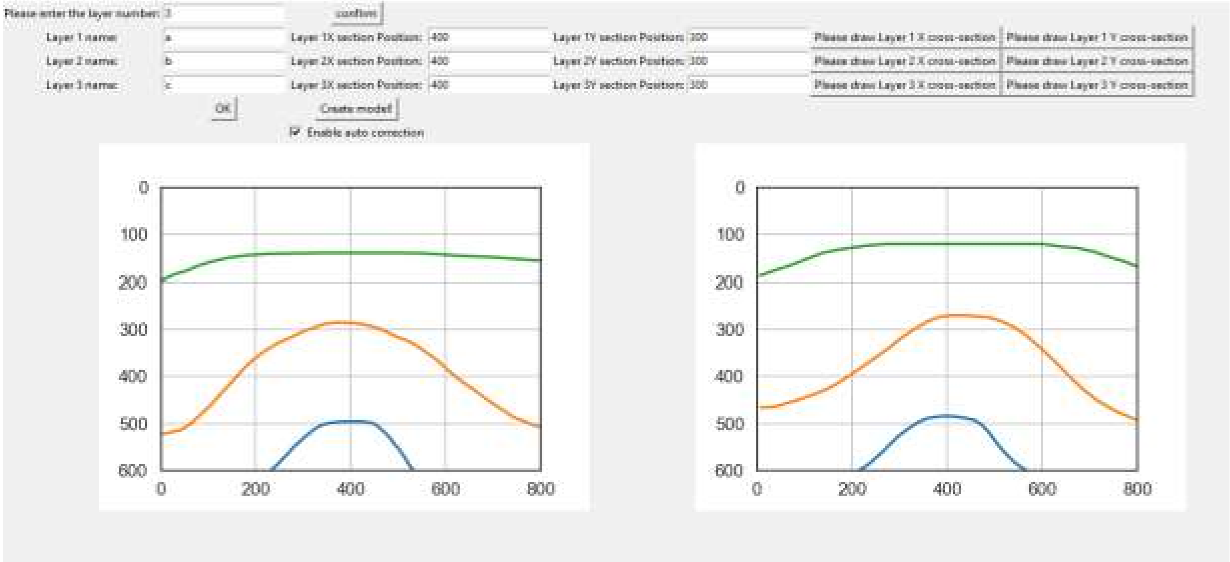
- Only 7 points to manipulate
- Easy to build a regular geometry

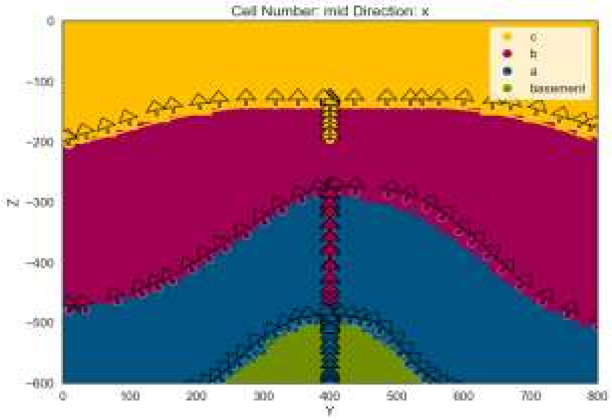
Disadvantages:

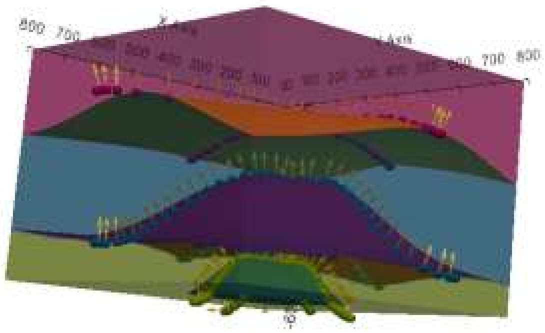
- Only allows one layer and two cross-sections
- Cannot be used to build complex geometries

Model generator v2.0

MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18







GemPy model generator V2.0

Fundamentals:

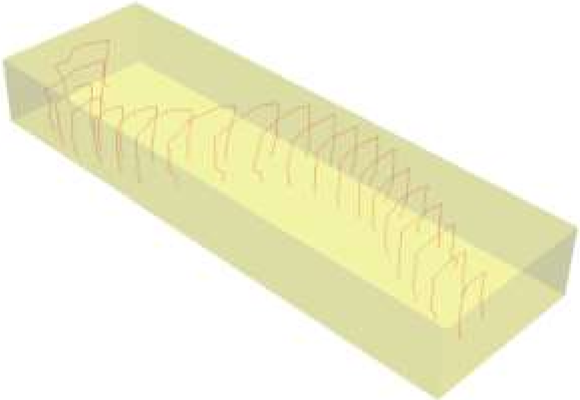
- Use drawing curve to define the geometry
- Sample points from curve as import

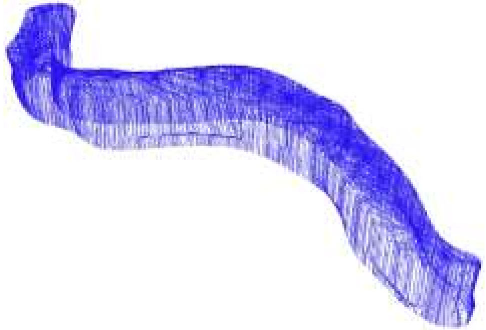
Advantages:

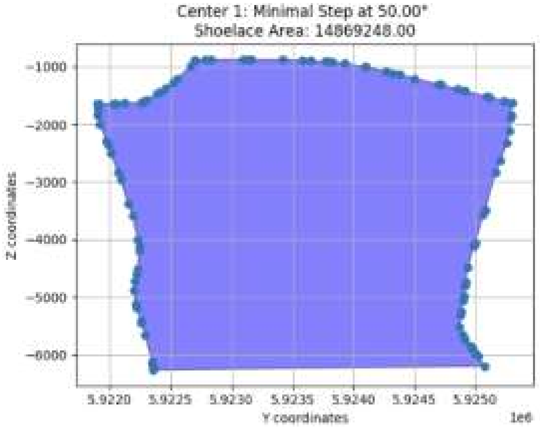
- Allows multiple layers and cross-sections
- Easy to build complex geometry
- A background image can be used as reference for drawing

Quantitative comparison of structures using geometrical/ statistical properties

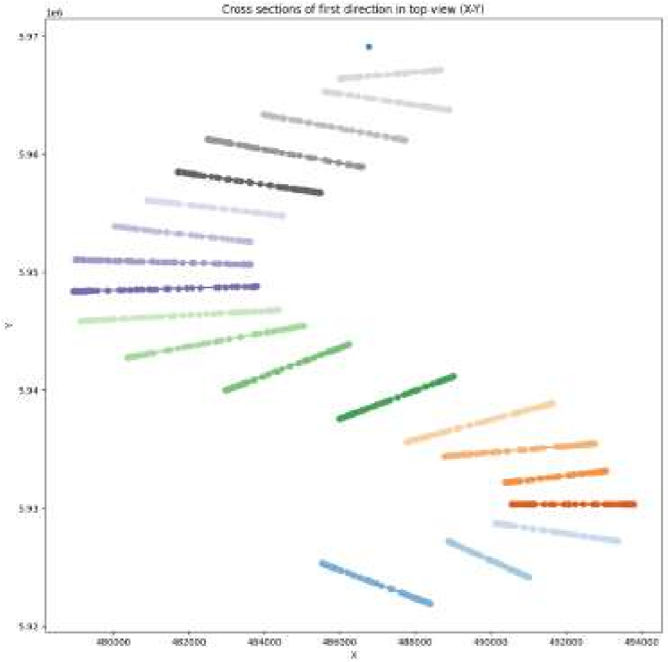
- Question: **how to describe and compare geological structures** quantitatively?
- Measurement of lateral & vertical dimensions and gradient & curvature data on explicit geomodels
- Data analysis can reproduce main geometrical characteristics of input data





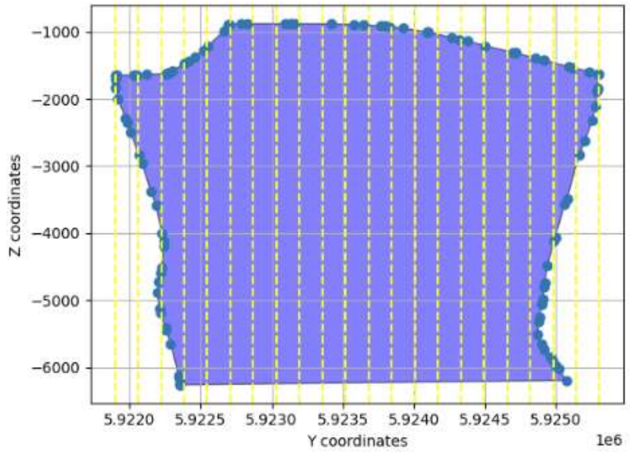


Top view 1st direction



Determination of orthogonal sections

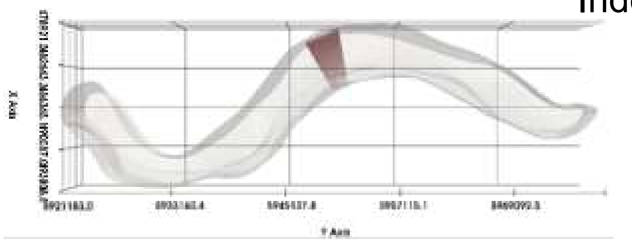
Index n



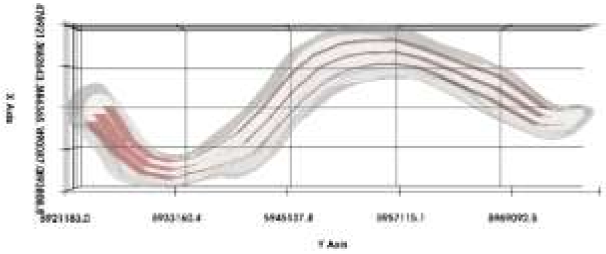
1. Divide the sections of the first direction (index=i) into n=22 vertical lines
Extract X, Y and Zmin +Zmax from each line → 2 points

Index i

Index n



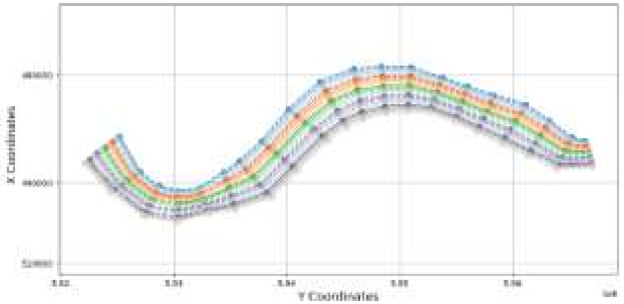
2. Create trapezoidal segments (one from overall 4 points of 2 consecutive sections)

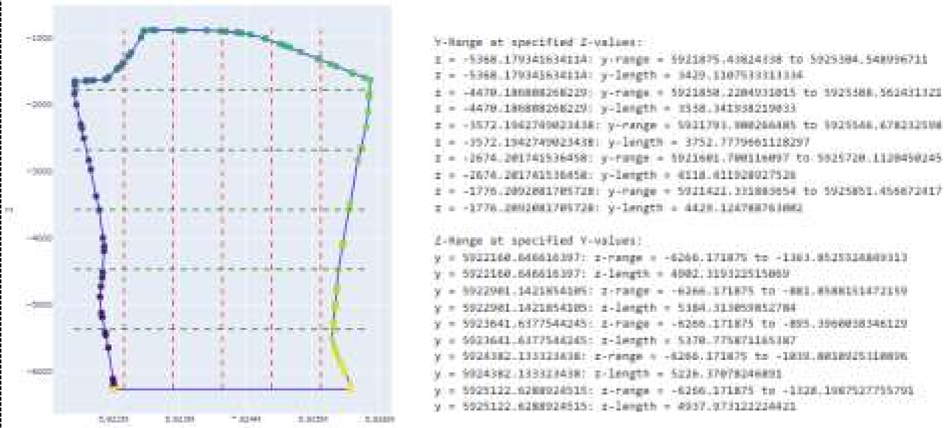


3. Combine trapezoidal segments of given line index=n to form orthogonal section

Dimensional measurements

Position of orthogonal sections & overall coverage of input mesh



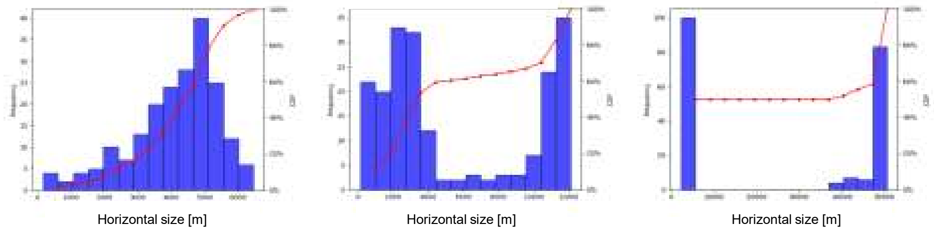


Measurement results: automatic determination of dimensions

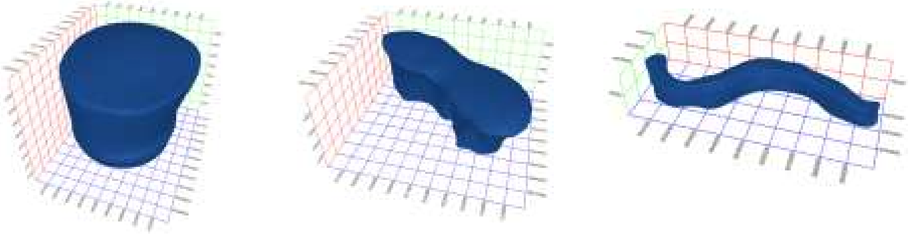
Data analysis: Comparison and systematization of structures and datasets
1) Distributions of combined horizontal dimensional data

Stocks | Moderately anisotropic walls | Highly anisotropic walls

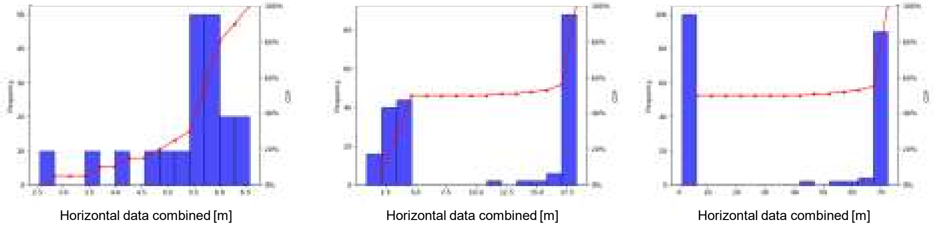
Distribution of horizontal size for selected TUNB models

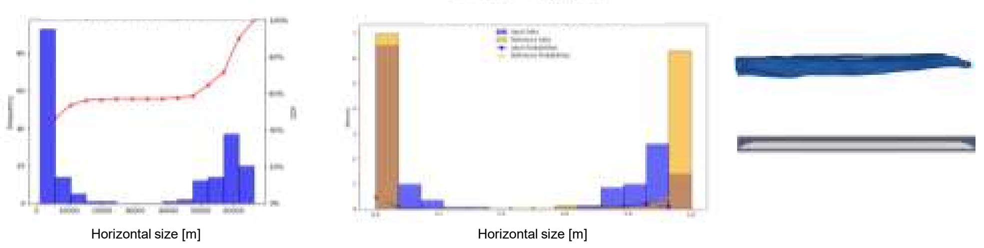


A)



Distribution of horizontal size for standard geometries from the library

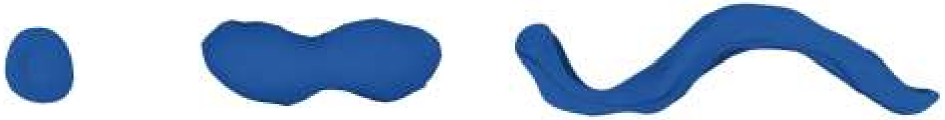




Kullback-Leibler divergence allows to quantify the similarity of a certain body (or slice, e.g. derived from seismics) with the bodies of the library of geometries and select the closest

B)

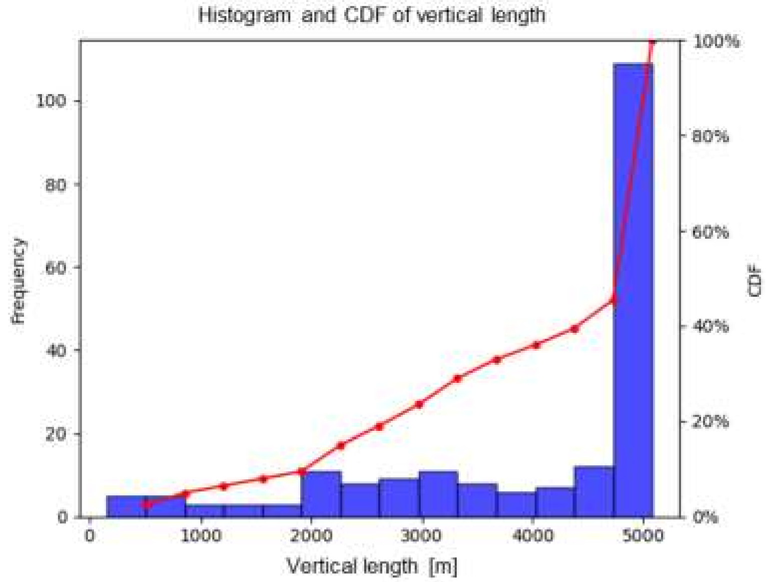
<2 | 2-6 | >6



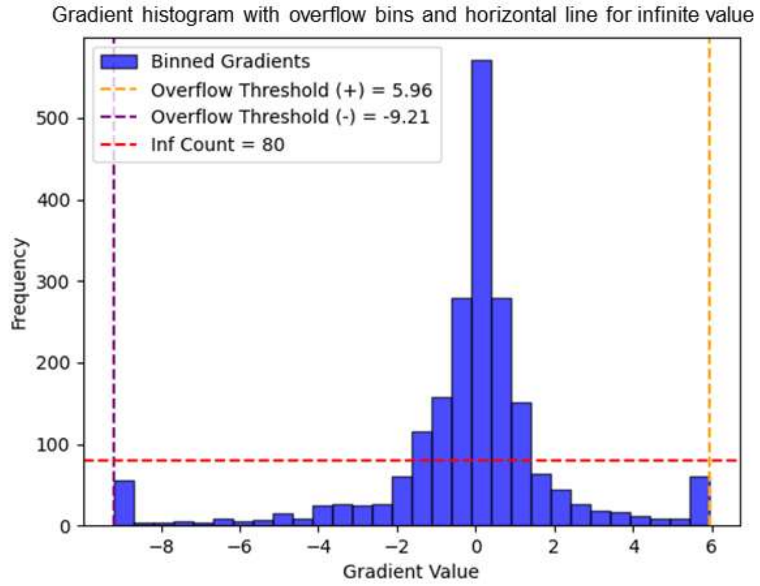
Averaged anisotropy:
ratio between mean of data of orthogonal direction & mean of data of 1st direction

2) Histograms of vertical data and gradient & curvature data

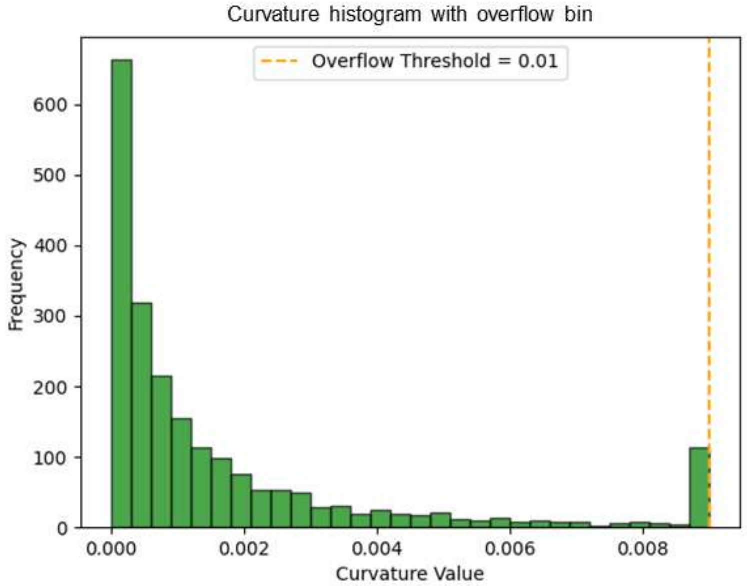
Typical distribution of vertical data



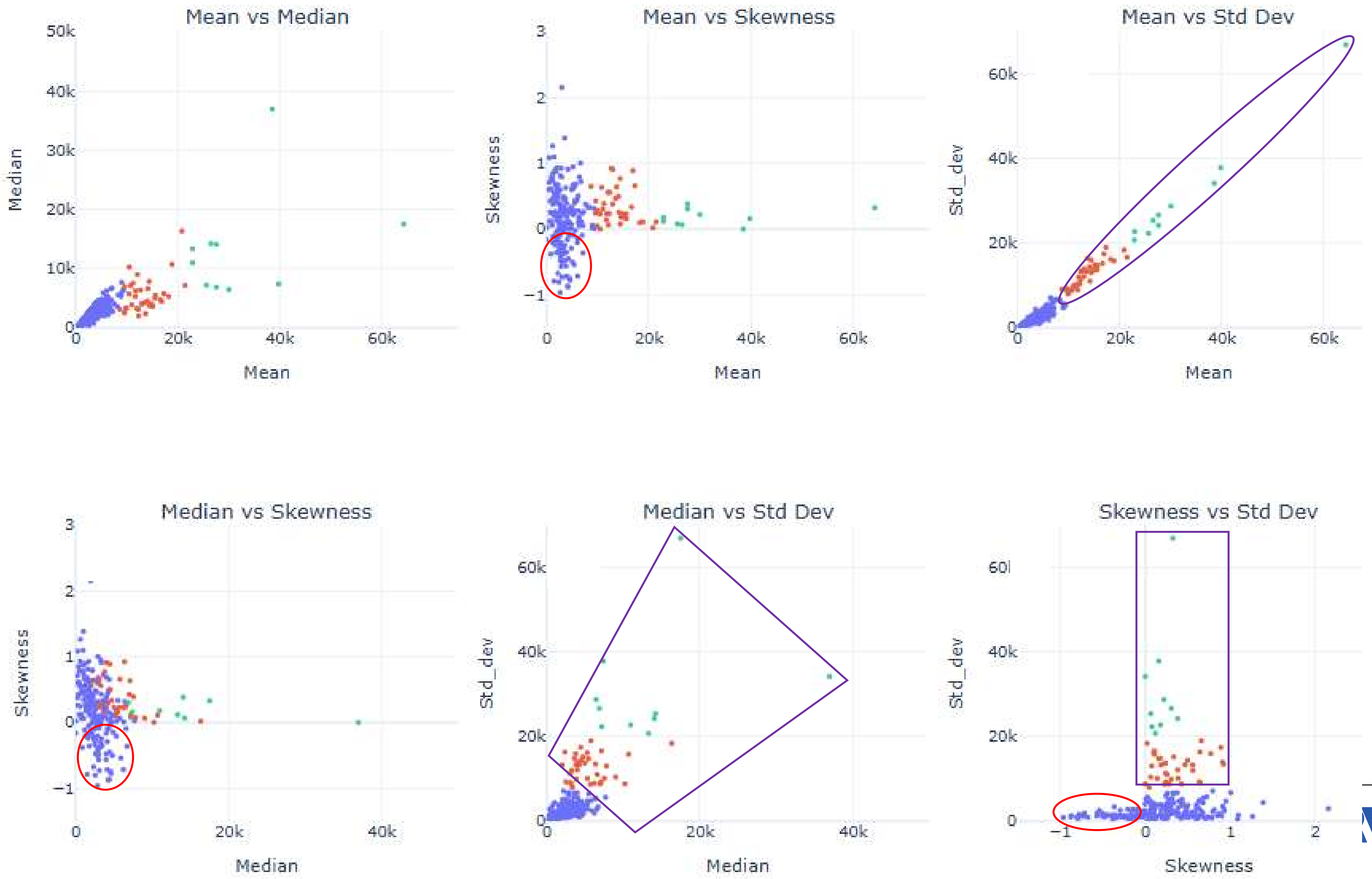
Example of gradient data



Example of curvature data

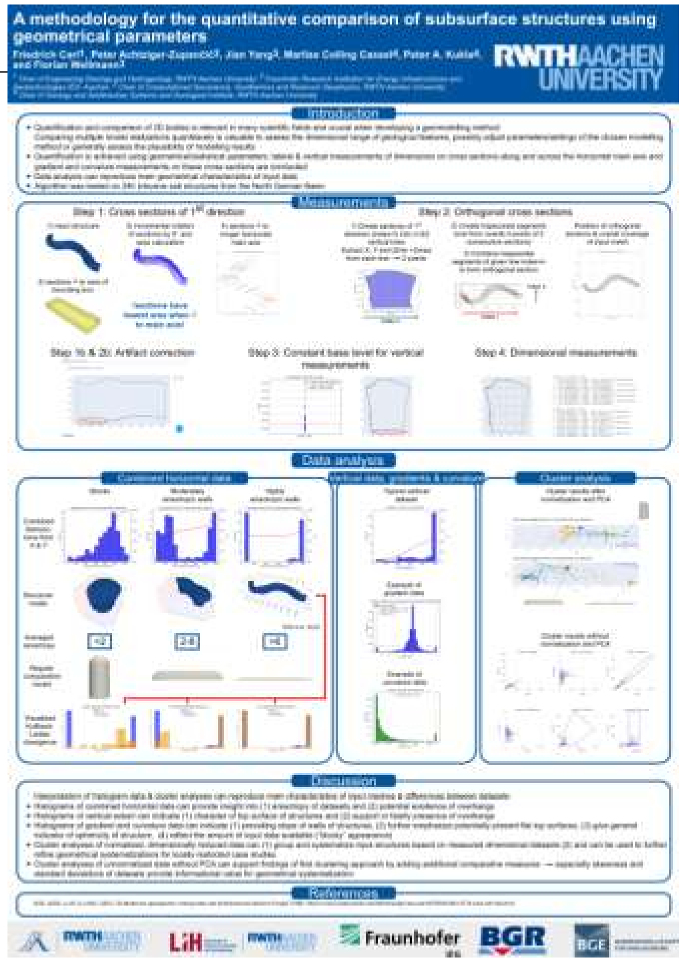


Cluster analysis on statistical moments (Example: horizontal data)



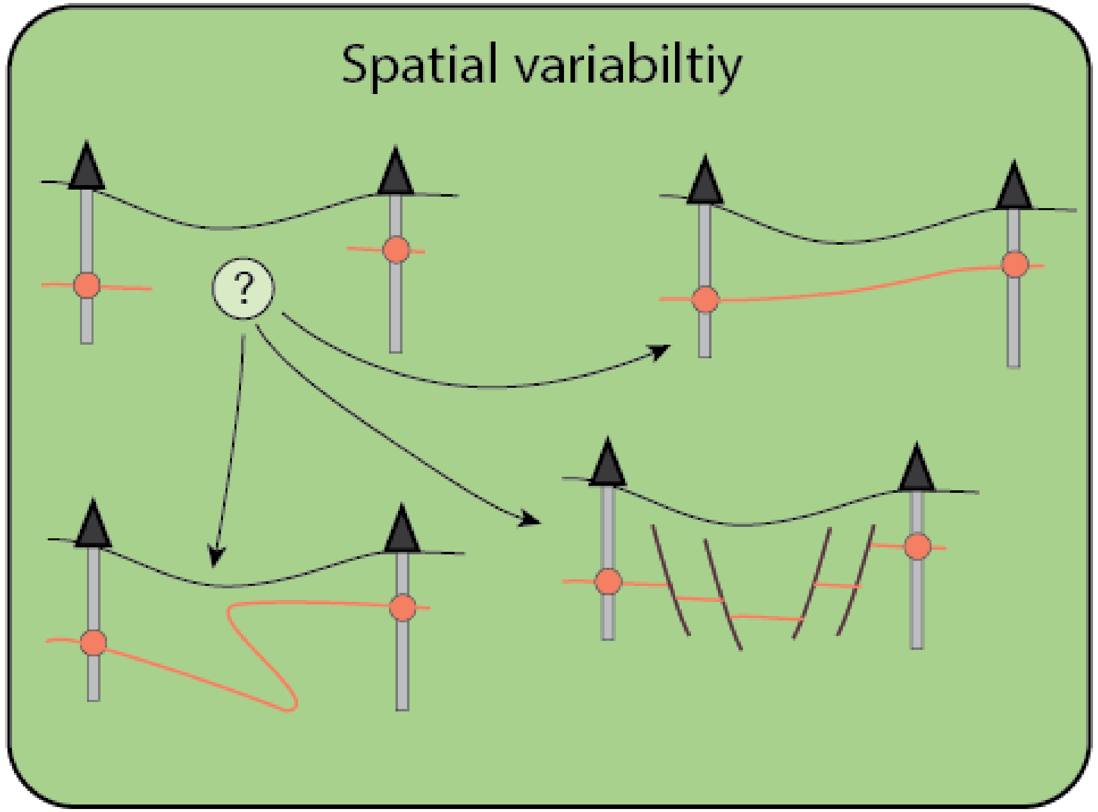
Classification of geometries: intermediate result

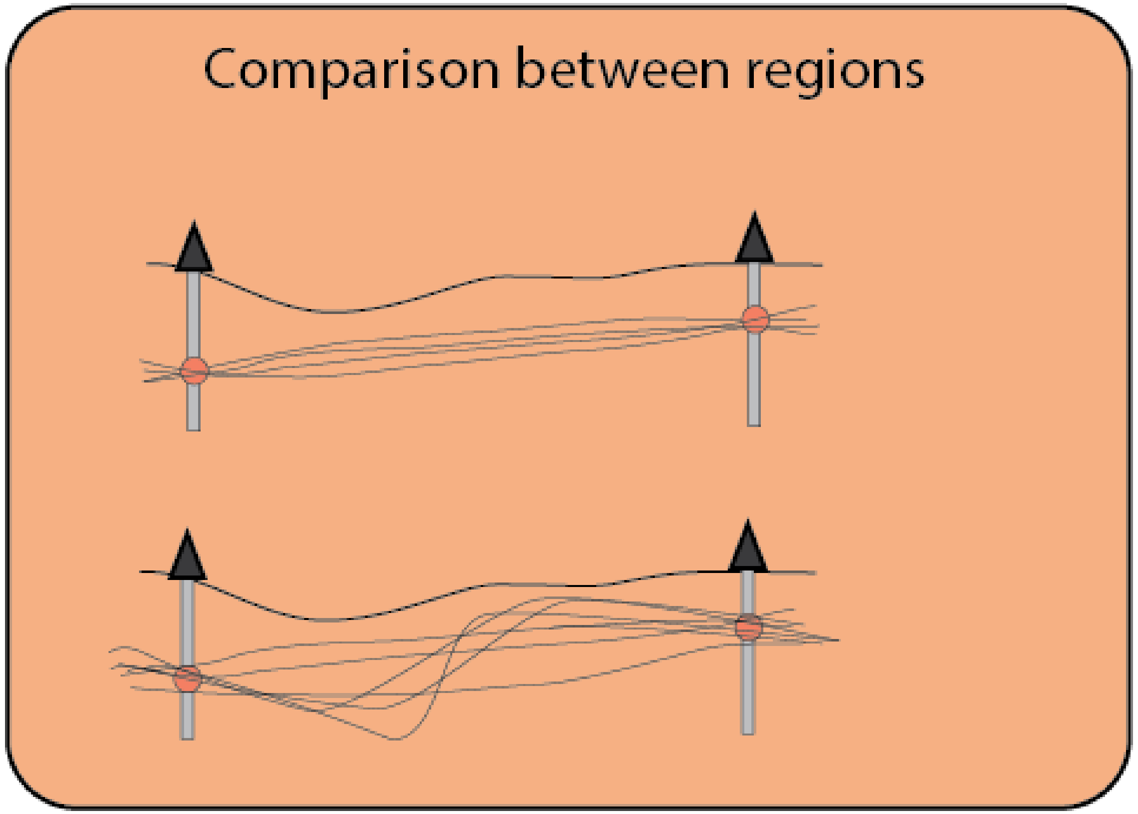
- Interpretation of **histogram data & cluster analyses** can reproduce main characteristics of input meshes & differences between datasets:
  - ➢ Statistics of "simple" geometrical properties **allows to describe each salt body** and carries information on: anisotropy of the model, existence of overhangs, 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?

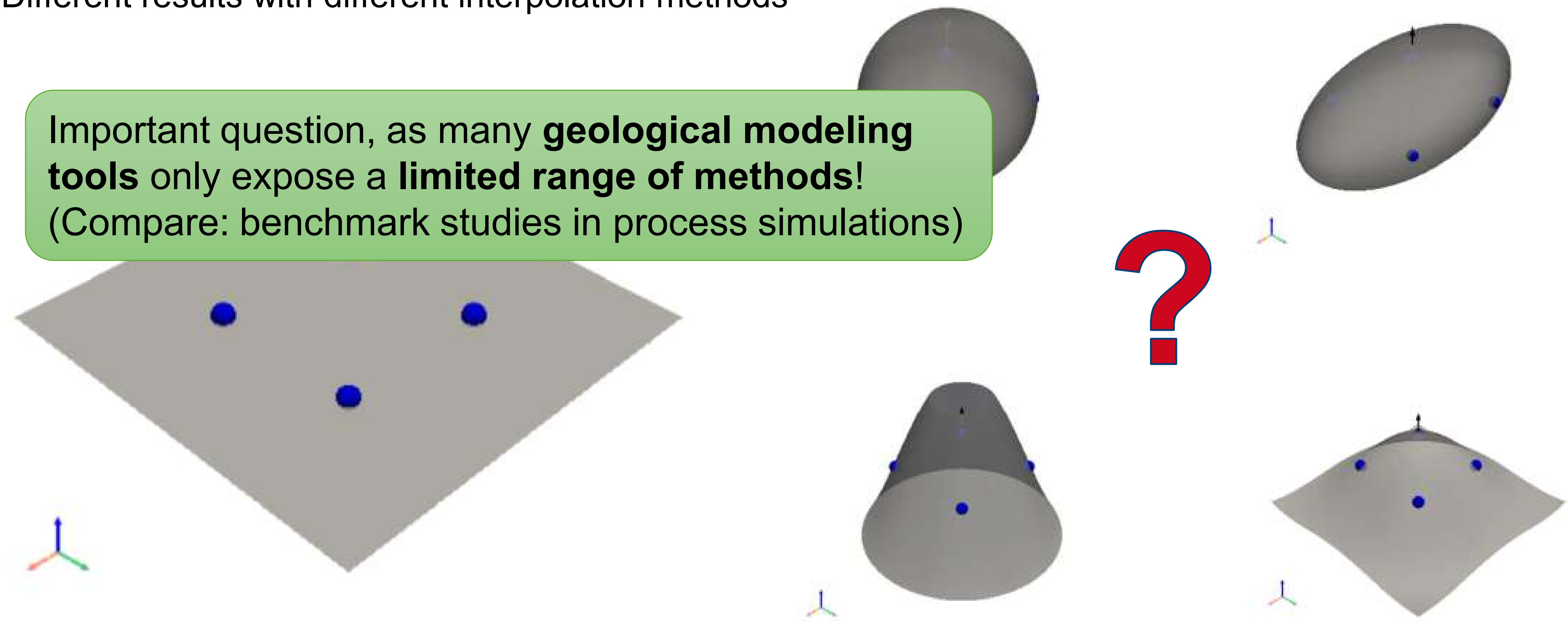




Which model is better?

- 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)



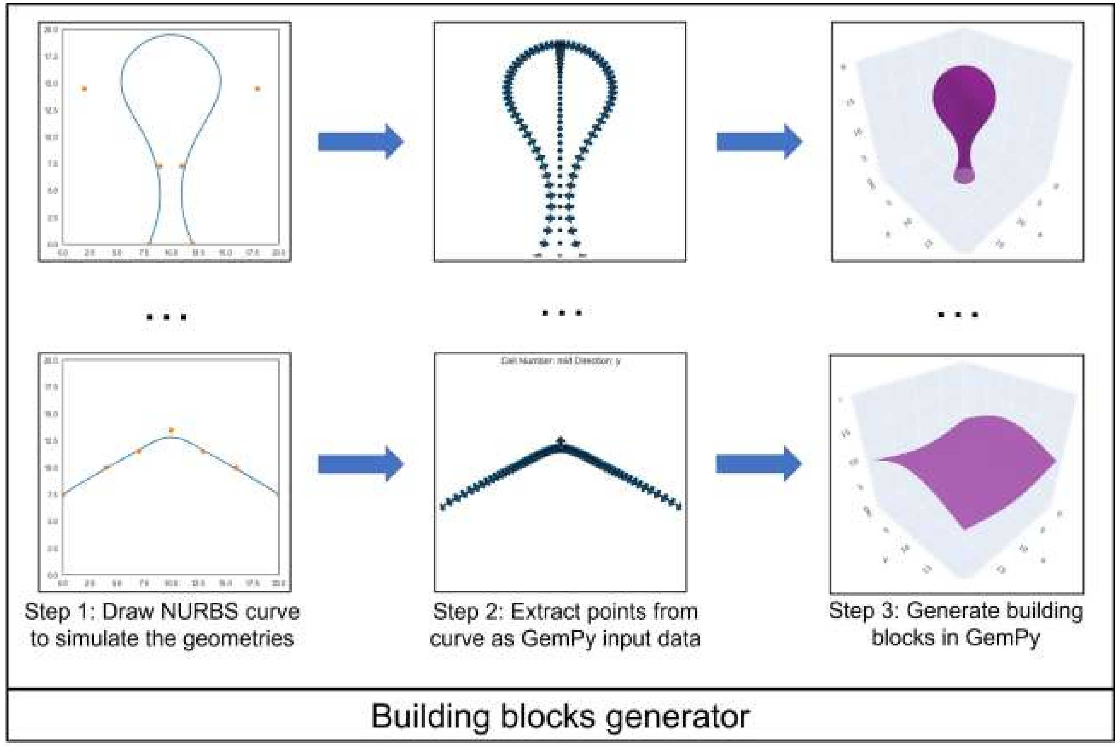
Step1: Comparison between widely used interpolatior

MS 4.1	Quantitativer Vergleich Interpolations-methoden für Strukturmodell	M18

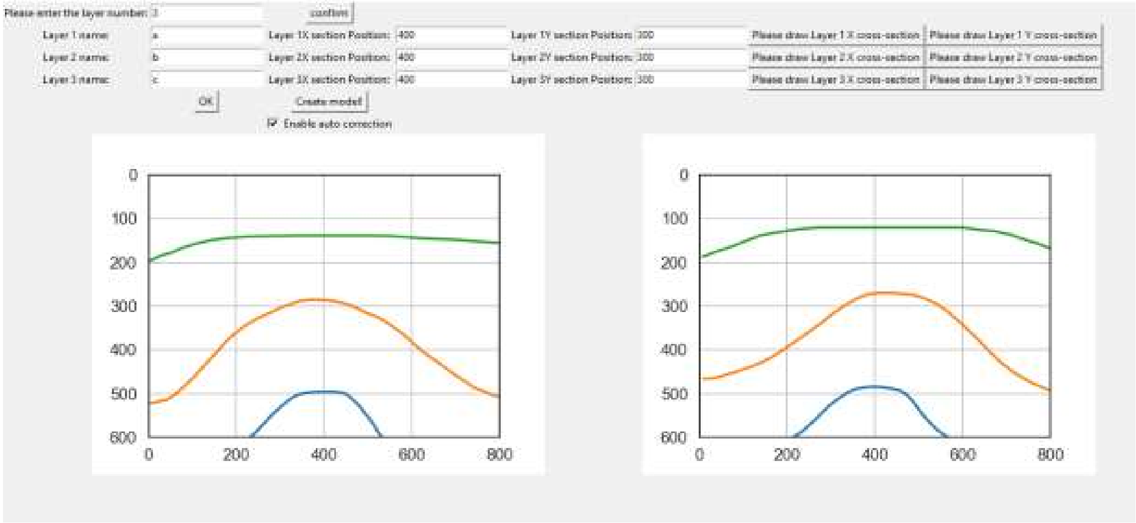
Question: Which interpolation method should be applied to model host rock

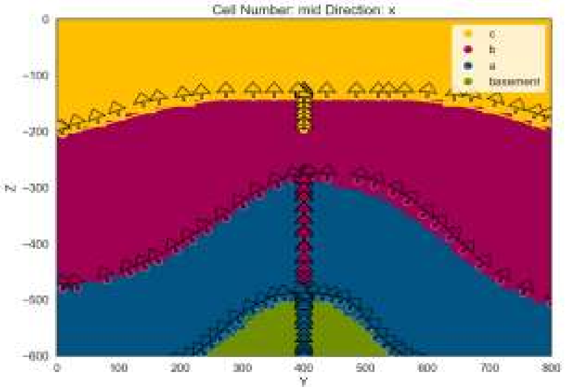
Model generators v1.0 & v2.0

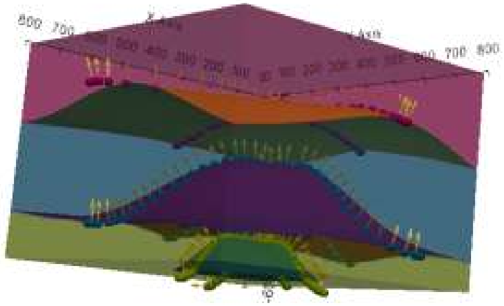
GemPy model generator V1.0



GemPy model generator V2.0



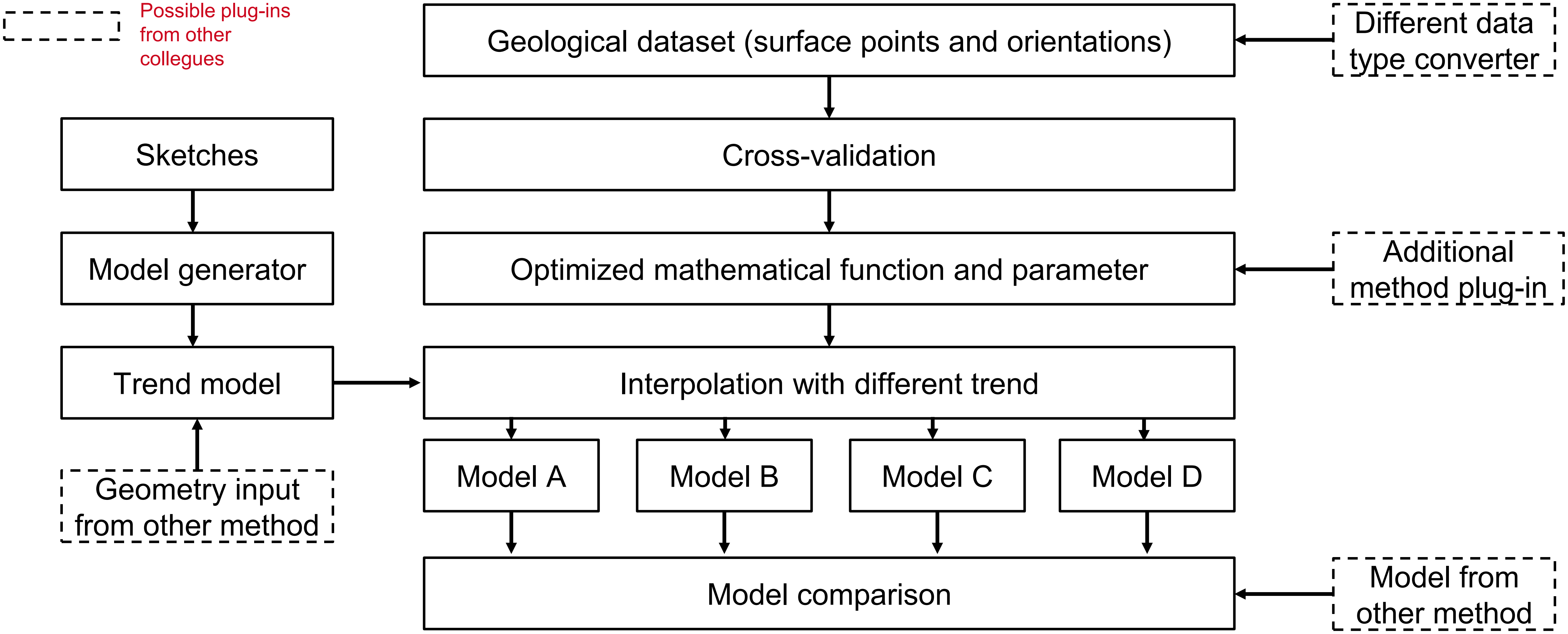




Computational Geoscience and Reservoir Engineering
RWTH AACHEN UNIVERSITY

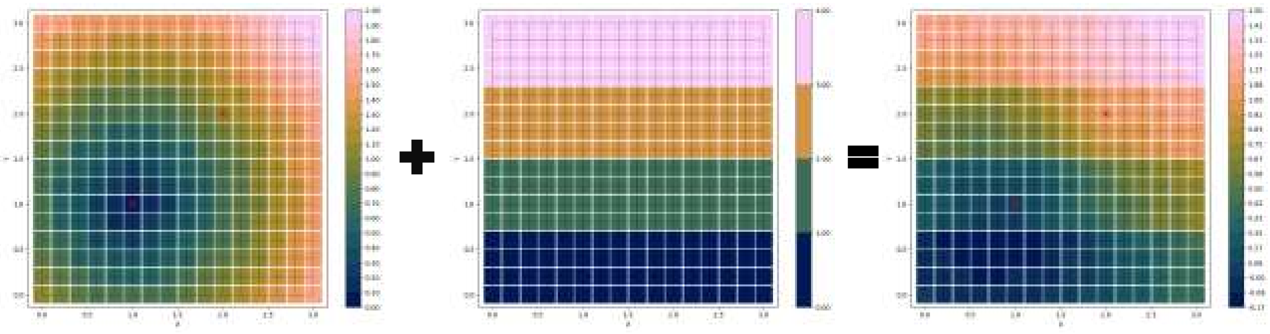
Step2: Comparison between optimized kernel functions and drift

MS 4.1	Quantitativer Vergleich Interpolations-methoden für Strukturmodell	M18
MS 4.3	Hierarchische Modellkonstruktion implementiert	M33



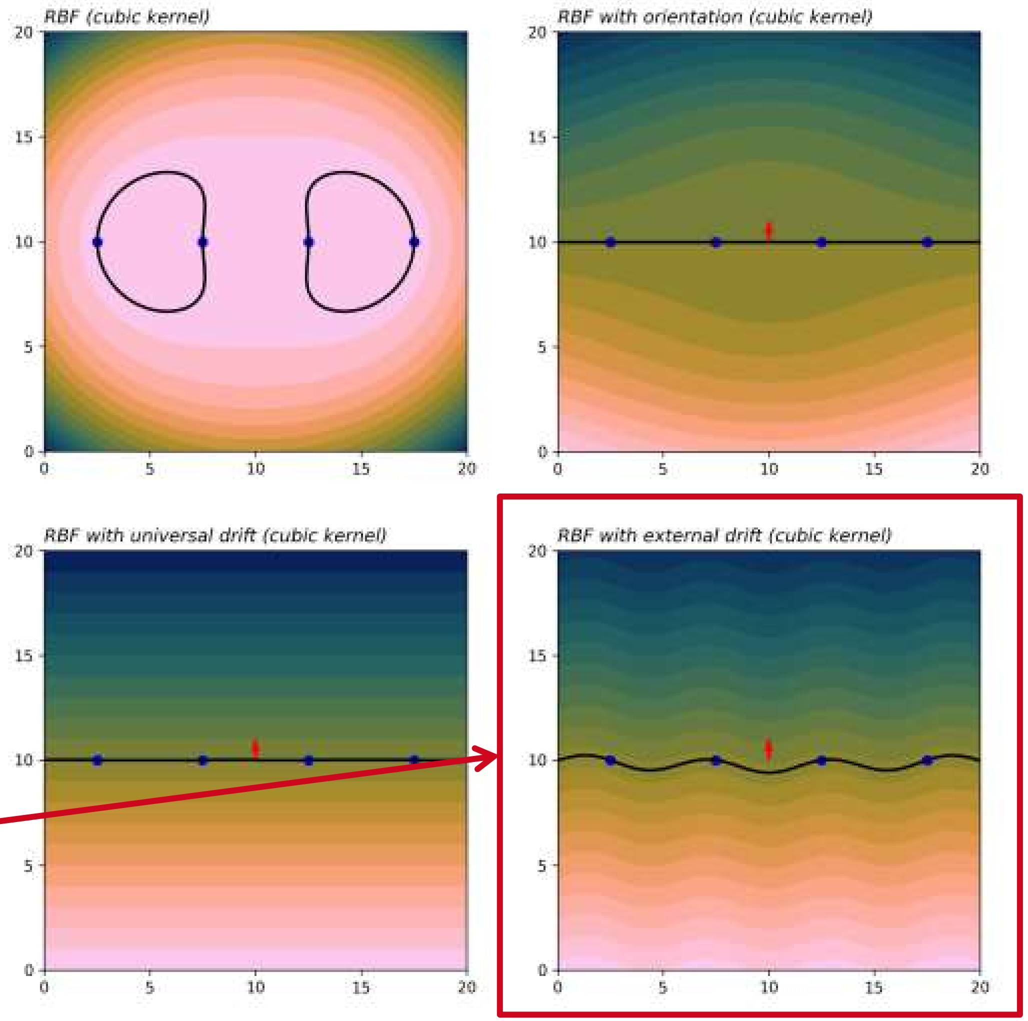
Solution: Suitable Mathematical function + known trend

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



How to compare models quantitatively?

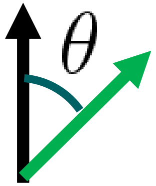
MS 4.1	Quantitativer Vergleich Interpolations-methoden für Strukturmodell	M18

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⟩)

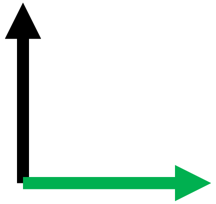


$$d^{\nabla}_{AB}(\mathbf{x}) = 1 - \frac{\nabla f_A \cdot \nabla f_B}{\|\nabla f_A\| \|\nabla f_B\|}$$

$$\frac{\nabla f_A \cdot \nabla f_B}{\|\nabla f_A\| \|\nabla f_B\|} = \cos(\theta)$$

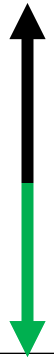
$\theta = 0$

$d^{\nabla}_{AB}(\mathbf{x}) = 0$



$\theta = 90$

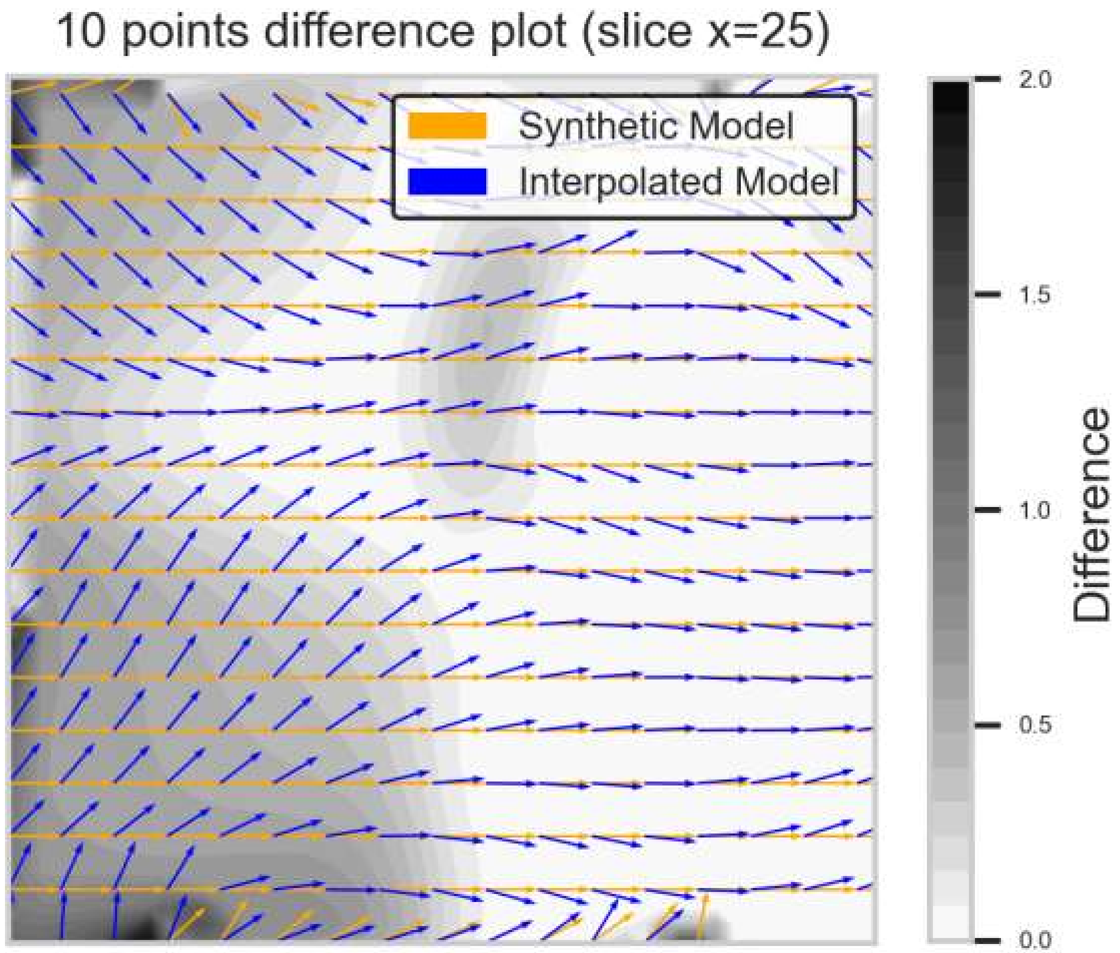
$d^{\nabla}_{AB}(\mathbf{x}) = 1$

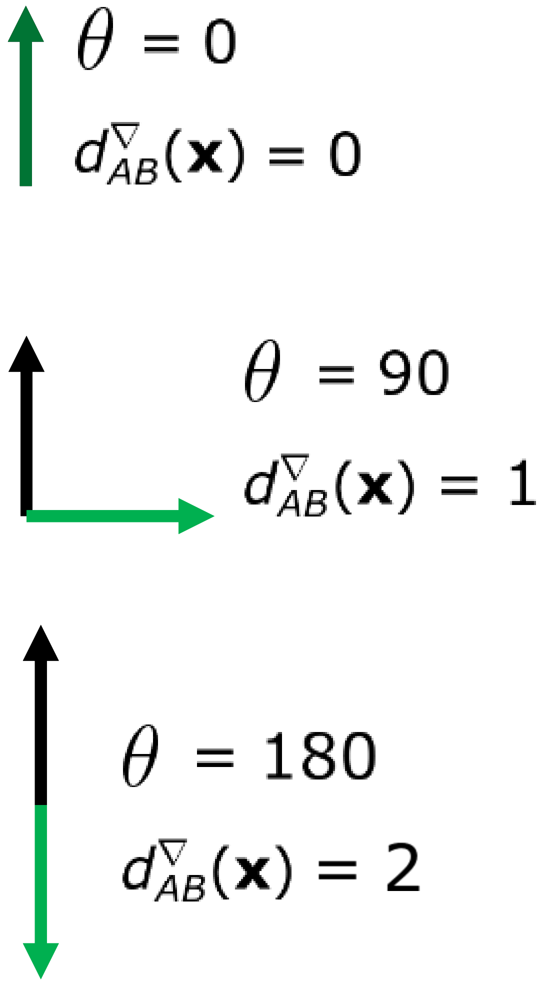


$\theta = 180$

$d^{\nabla}_{AB}(\mathbf{x}) = 2$

Model Comparison method

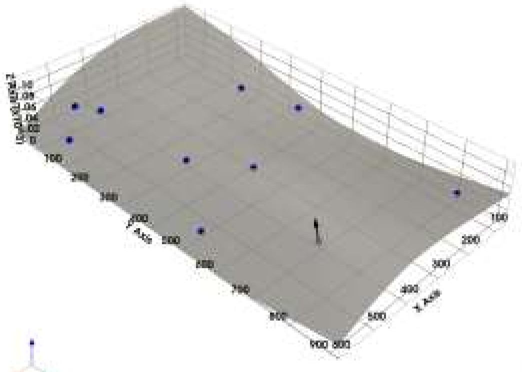




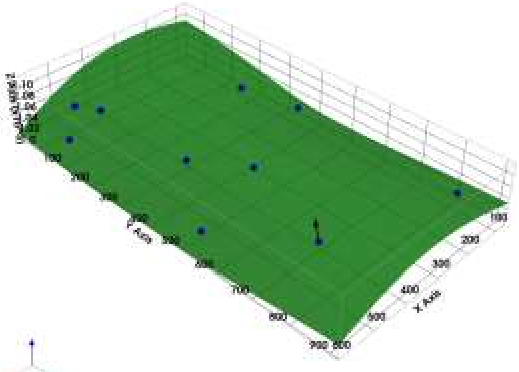
Case study (synthetic fold model)

MS 4.2	Auf Testdatensätze angewendet	M30

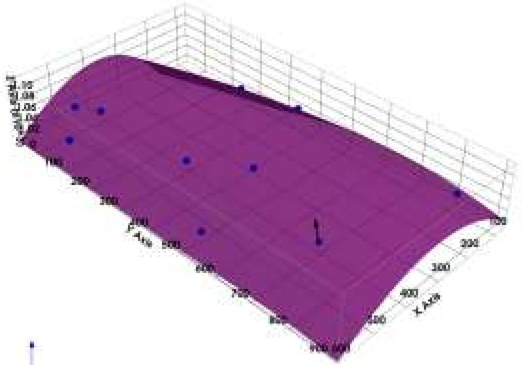
10 surface points and 1 orientation



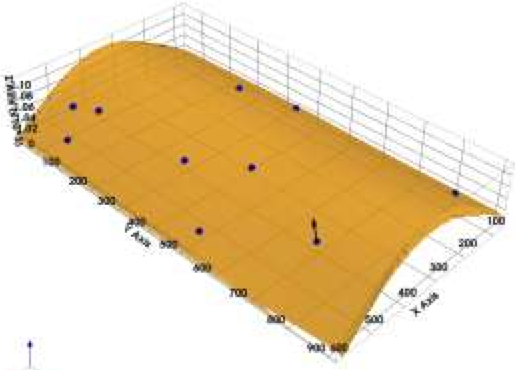
a) without drift



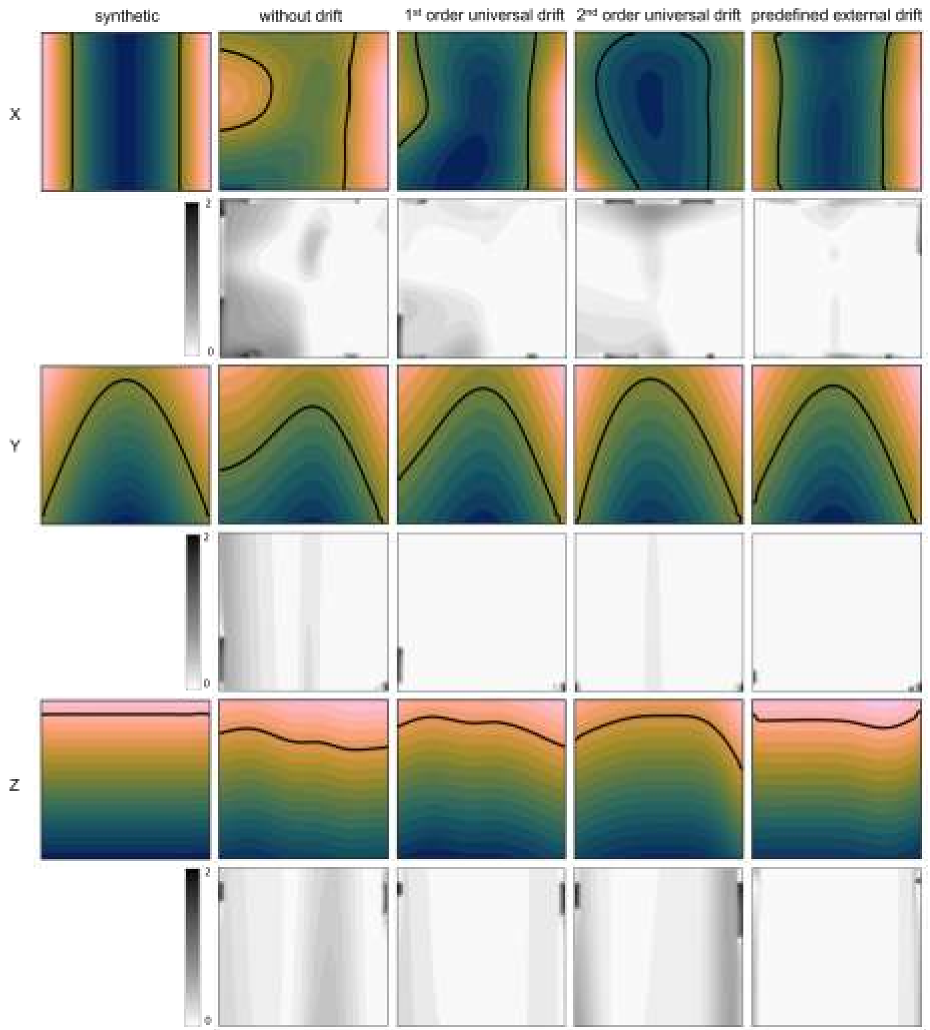
b) 1st order universal drift



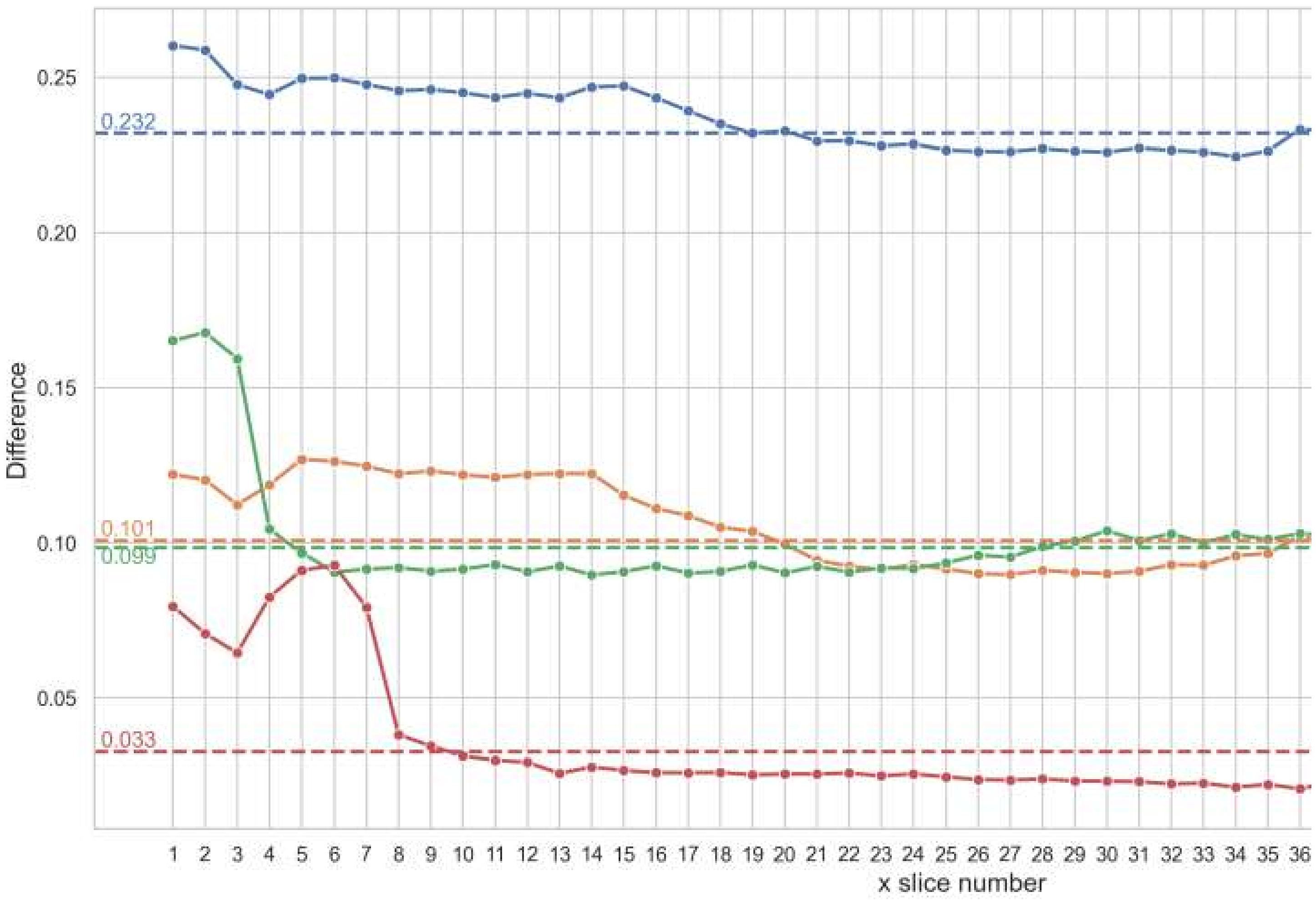
c) 2nd order universal drift



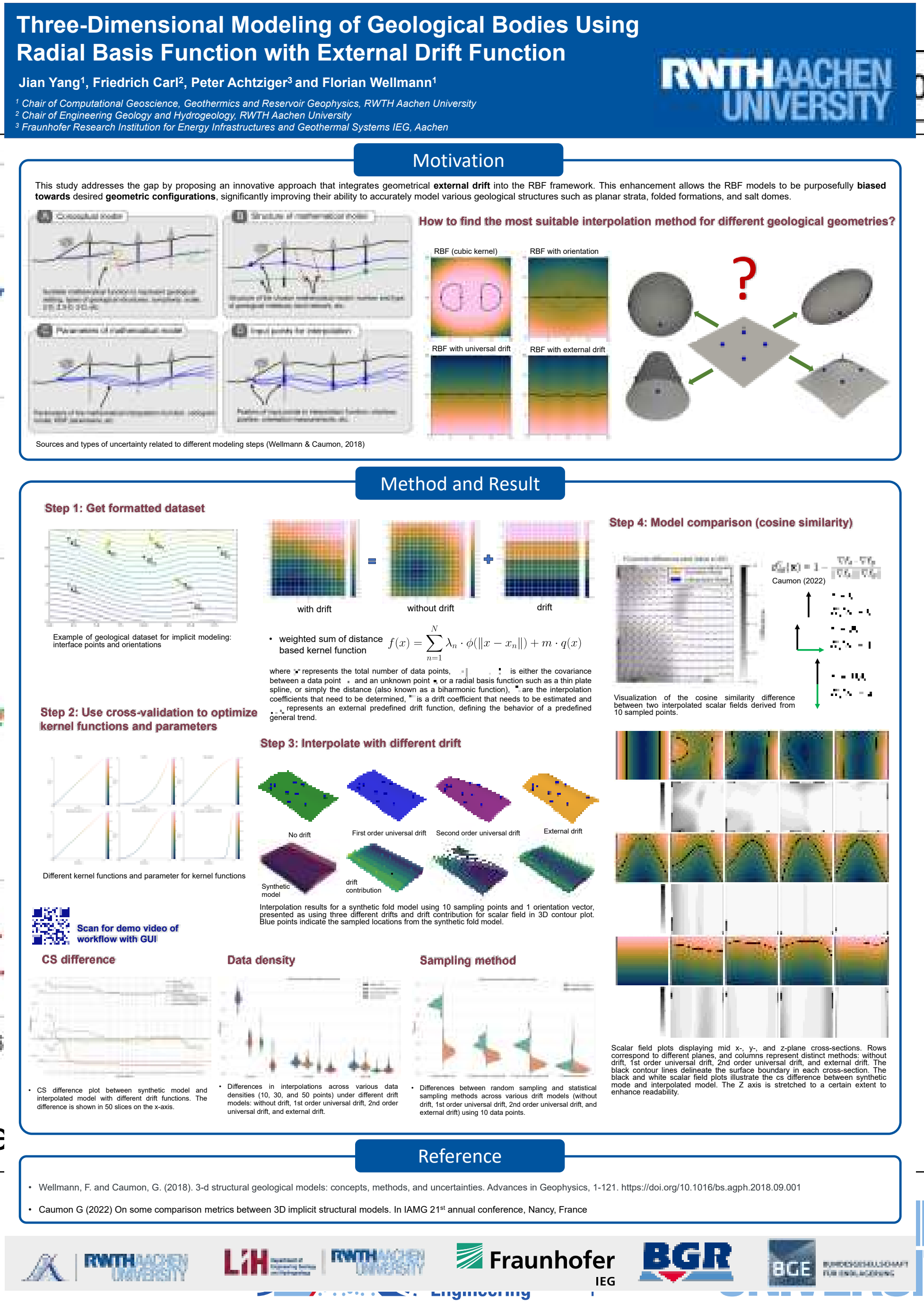
d) predefined external drift



Case study (synthetic fold model)



CS difference plot between synthetic model and interpolate



Part 3. Input data and uncertainties in interpretation

<u>Specific objectives</u>:

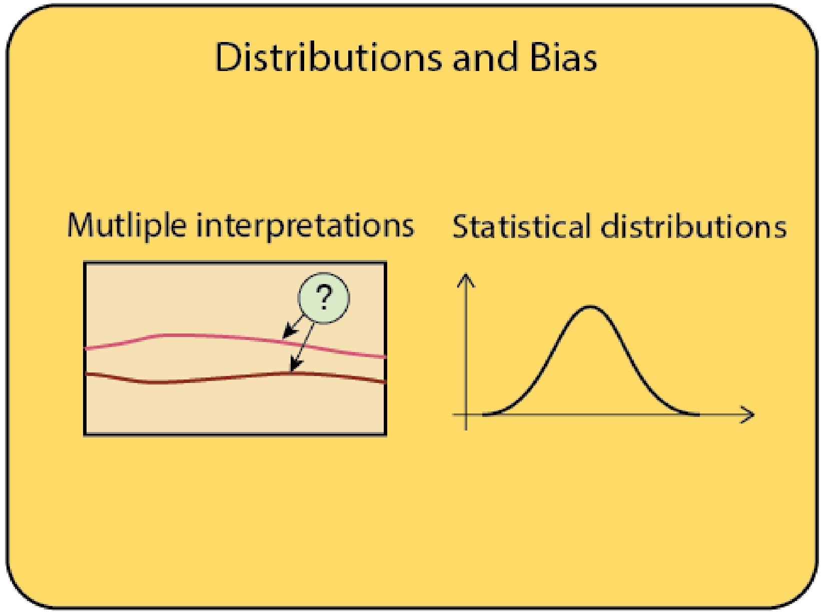
- **Identify data types used in site selection**
- **Identify uncertainties in these input data types**
- **Assess errors in seismic data interpretation**
- **Characterise qualitative irreducible uncertainties.**

- Expected contribution to site selection:

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

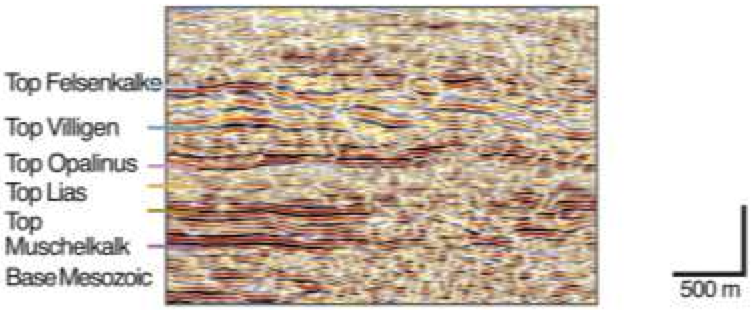
- Input Data Types:

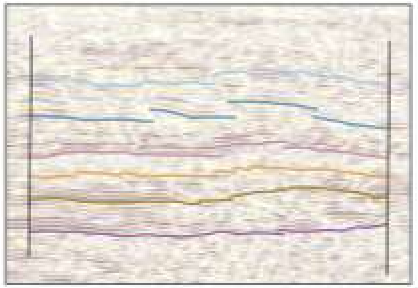
Seismic data
Velocity models –depth conversion
Seismic raw data – acquisition/processing
Borehole data
Geological maps and reports
Other geophysical and well data
Expert prior knowledge



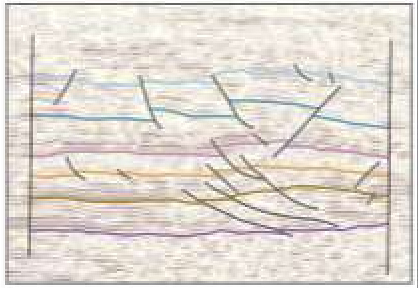
Mutiple interpretations of seismic image data

Interpretational Uncertainty:

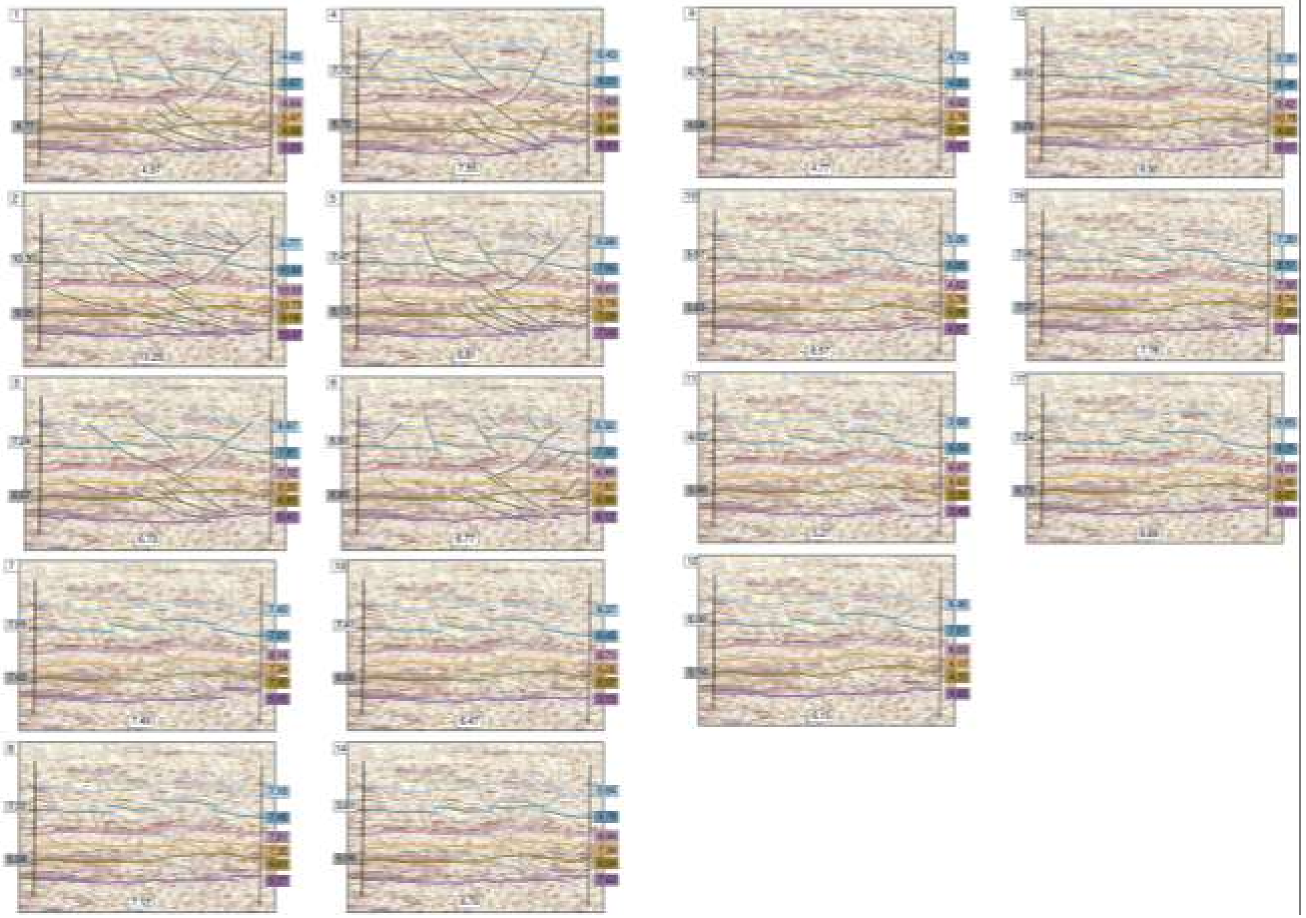




Interpretation Workflow 1: Horizon interpretation leaving fault gaps.



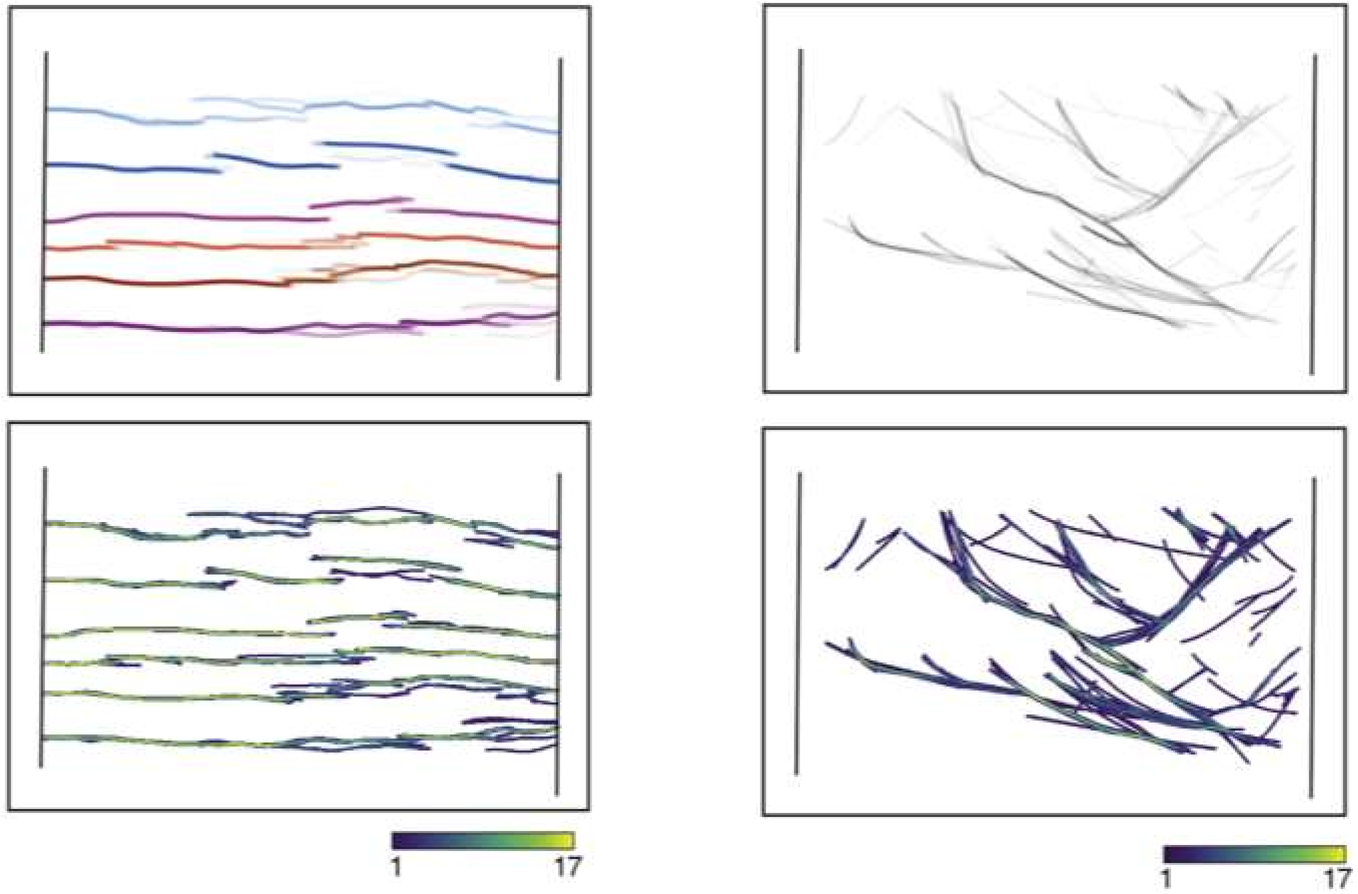
Interpretation Workflow 2: Fault interpretation linking gaps.

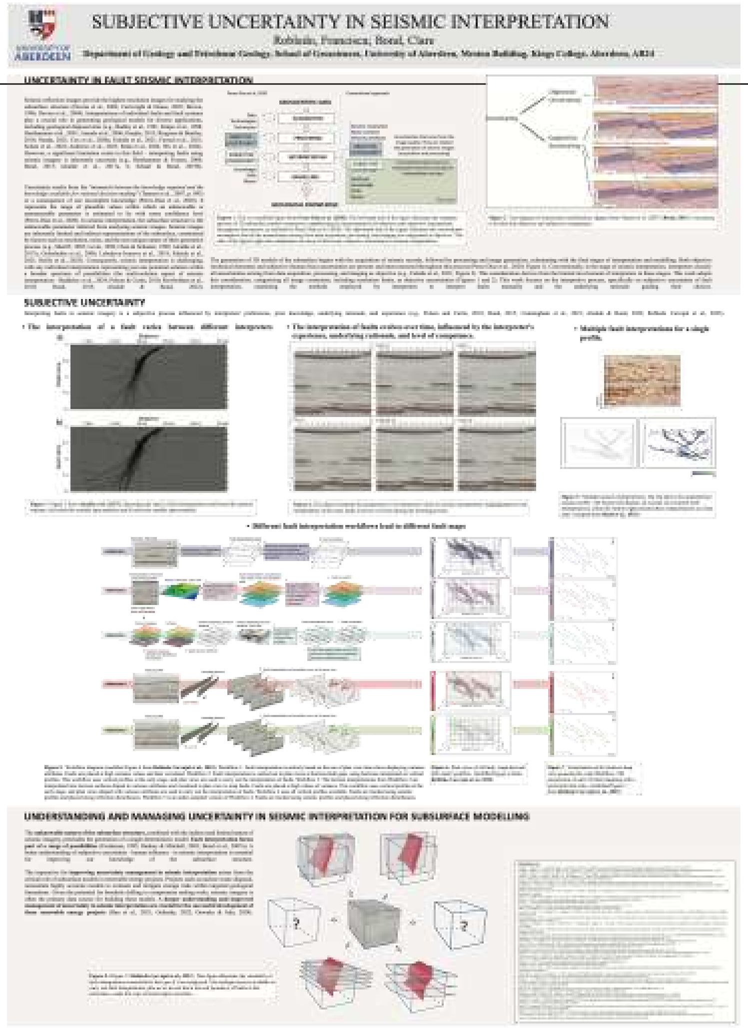


Bond et al. 2025

Mutiple interpretations of seismic image data

Interpretational Uncertainty:





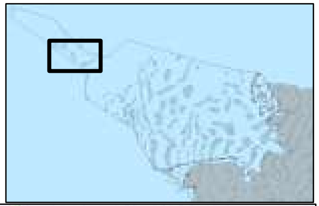
Bond et al. 2025

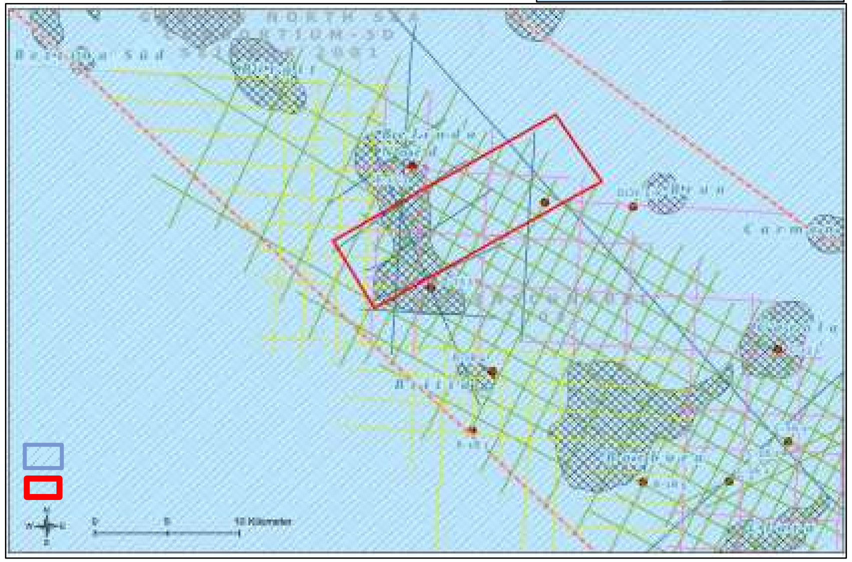
Uncertainty Quantification from Seismic Data (Current Work)

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

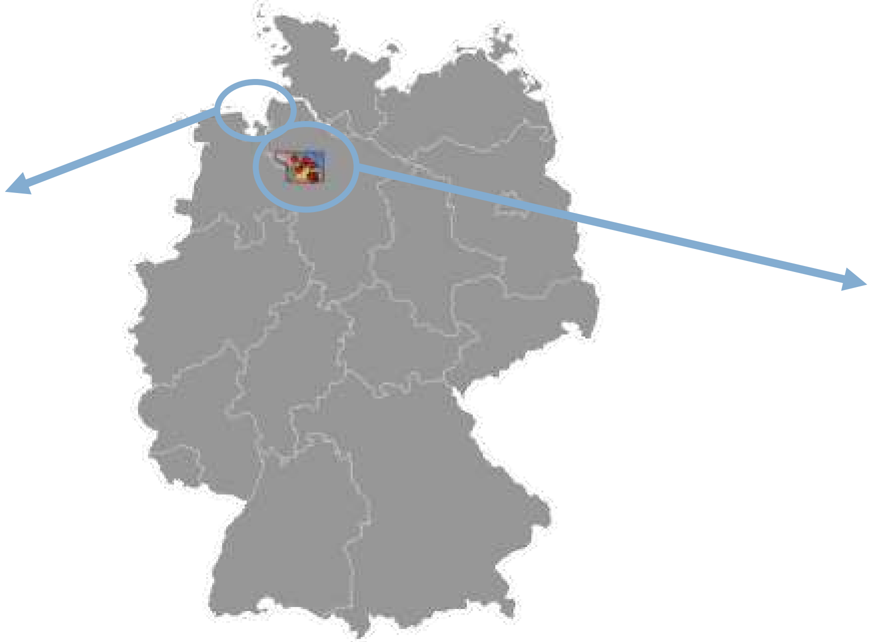
Input Data

North Sea Area

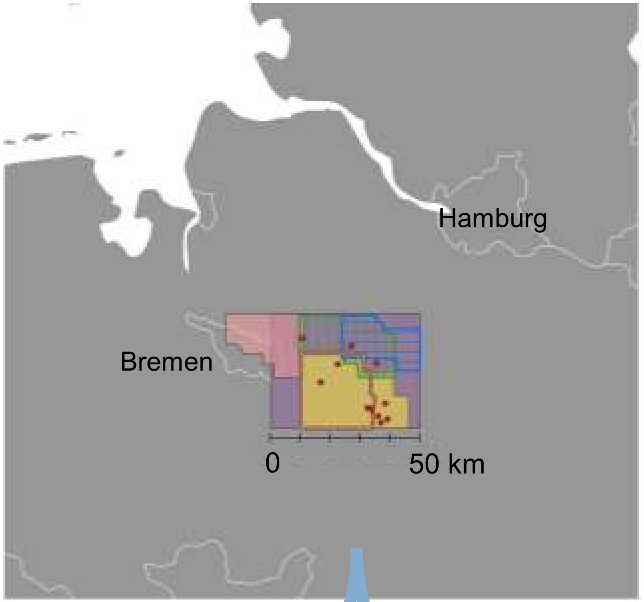


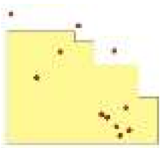


1974-2002 (+ reprocessing)

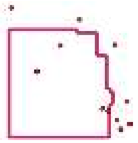


North Hannover Area

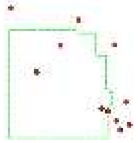




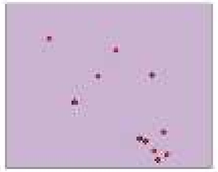
NH 032_Wedehof-Hamwiede 93-94



NH 2040_Wedehof-Hamwiede Repr 2010



NH 2012_Völkersen Repr 2012

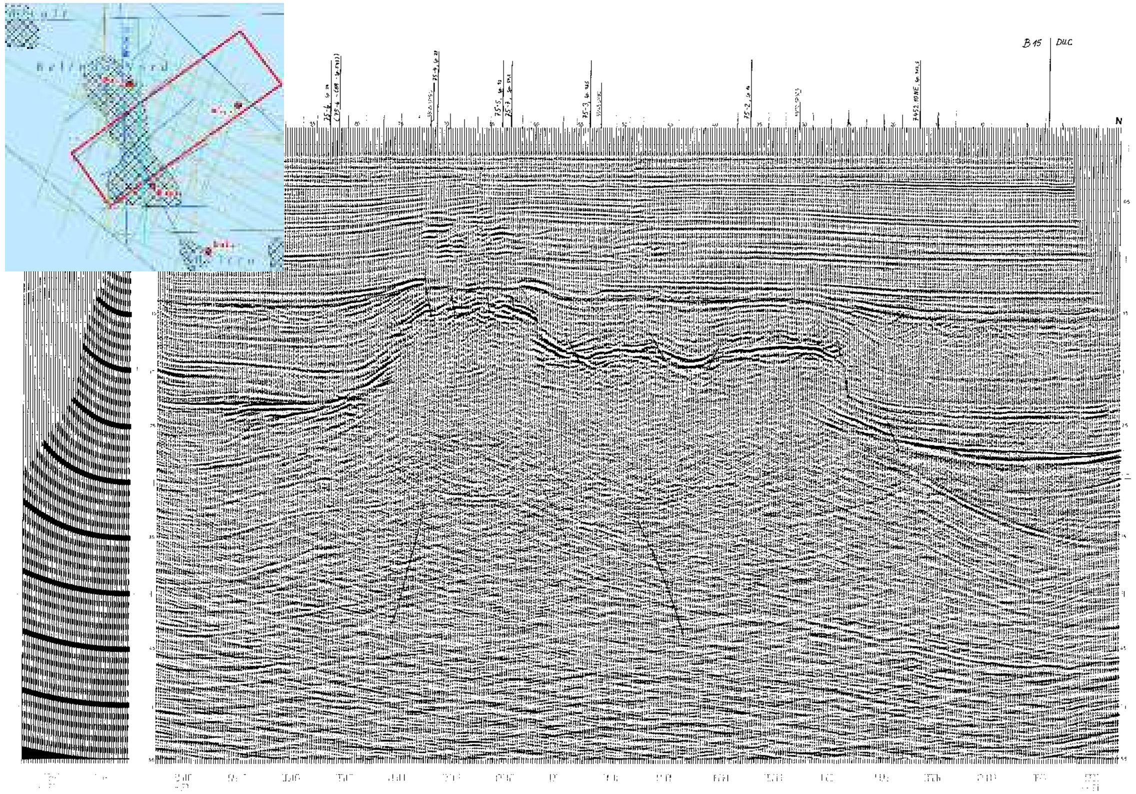


NH 2039_NordHannover Repr 2004

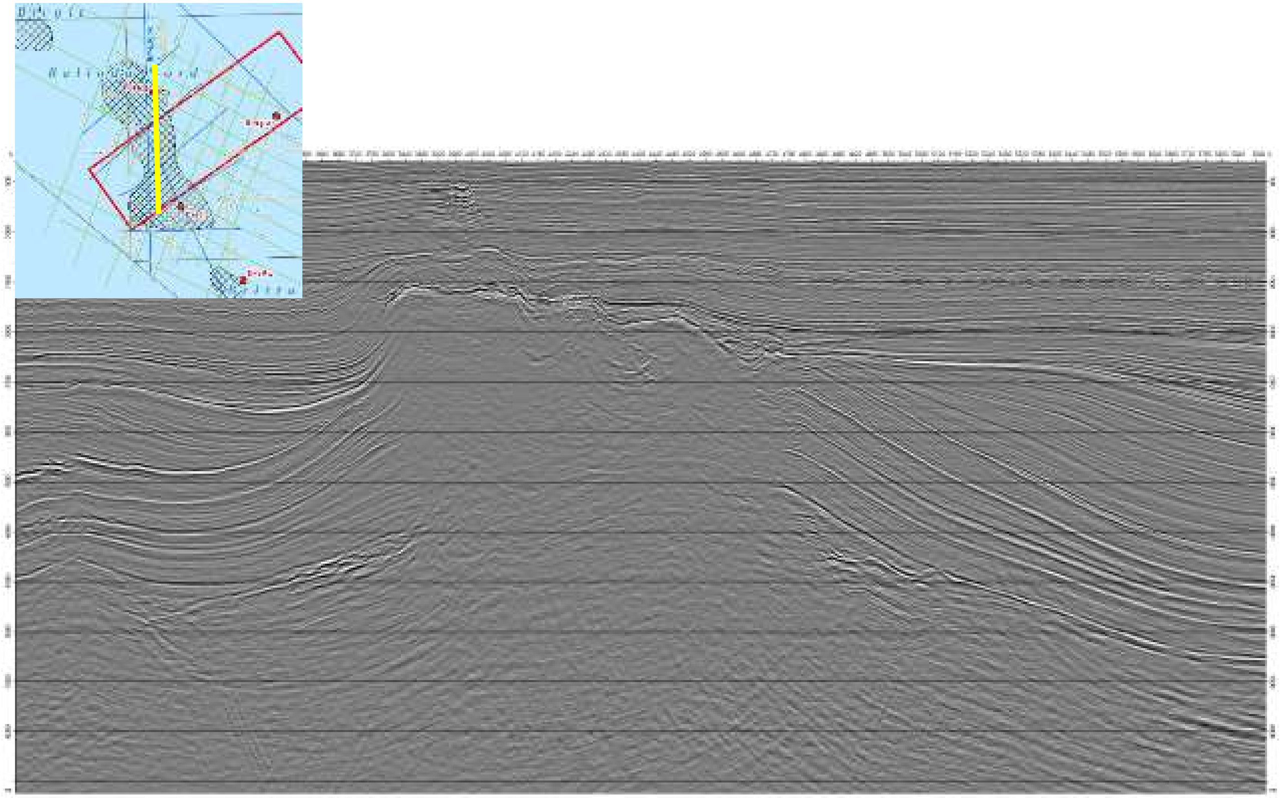
MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

Uncertainty Quantification from Seismic Data (Current Work)

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



1974



2002

Computational Geoscience and Reservoir Engineering

RWTH AACHEN UNIVERSITY

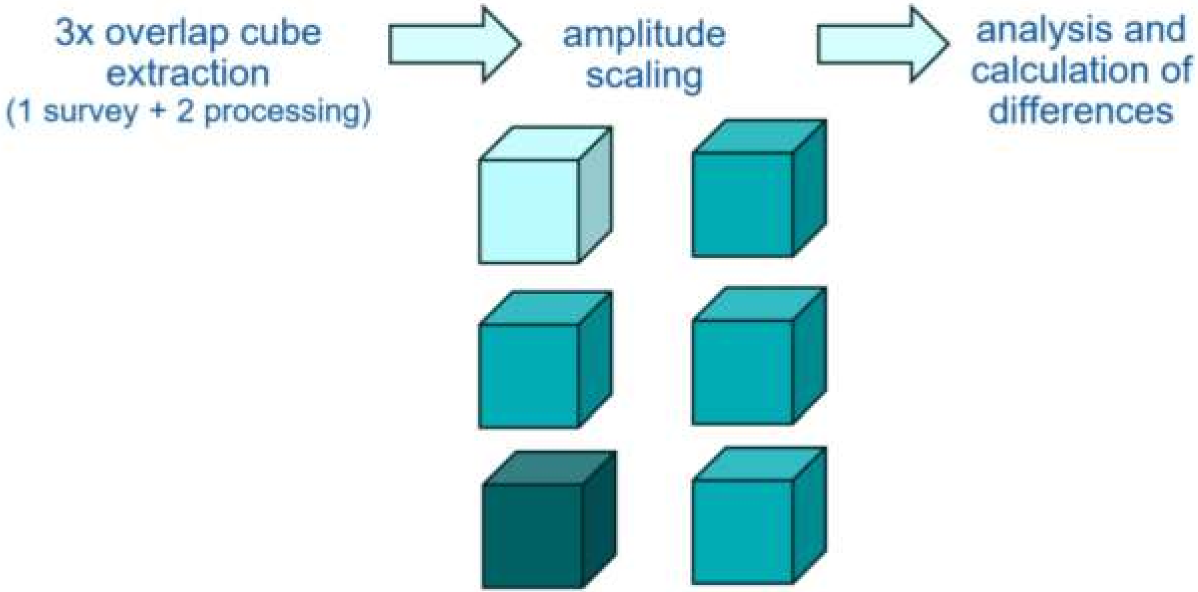
MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

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)

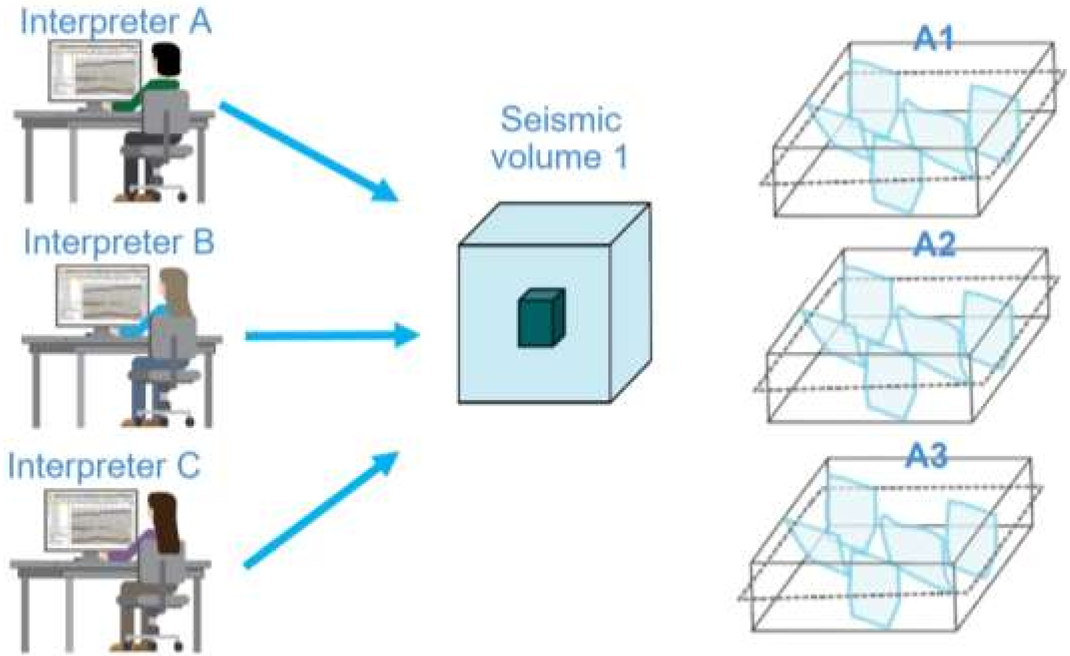
Seismic Workflow 2 Uncertainties in Interpretation

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.

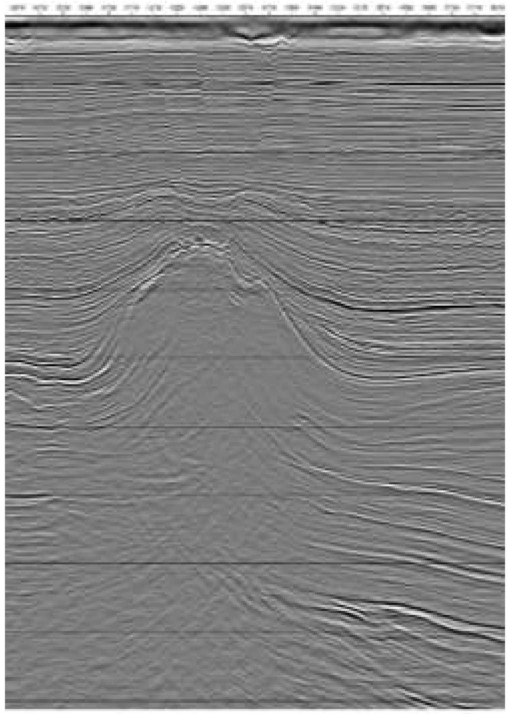


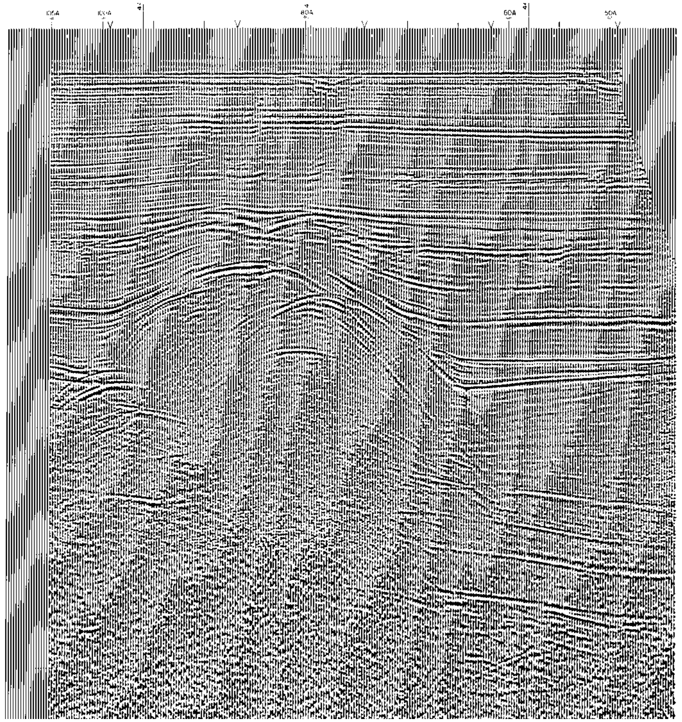
Computational Geoscience and Reservoir Engineering
RWTH AACHEN UNIVERSITY

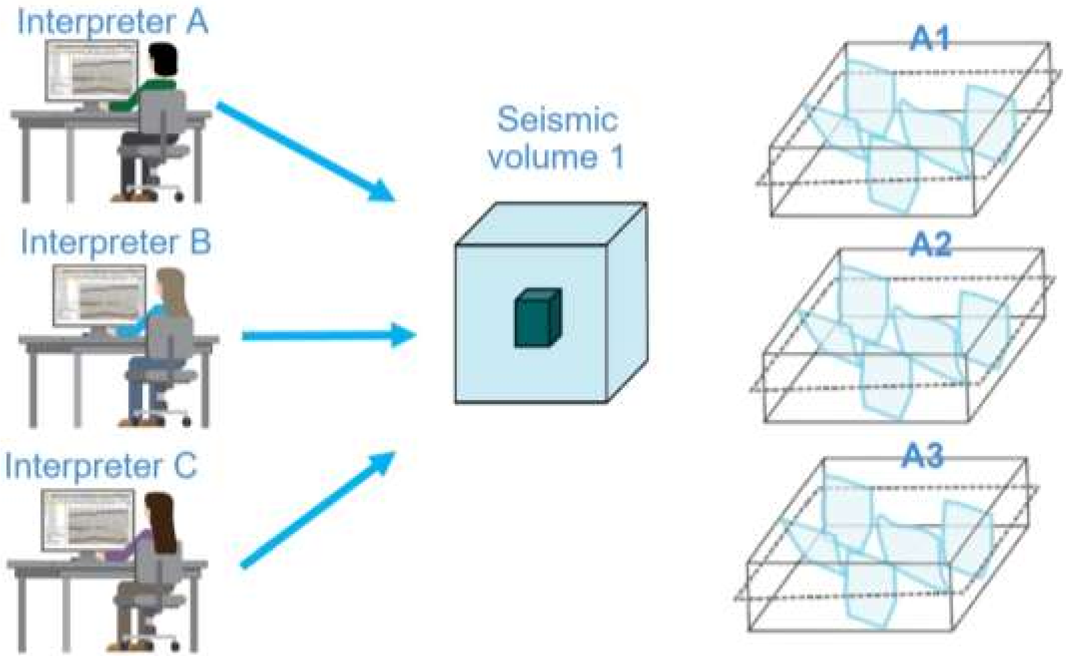
MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

Uncertainty Quantification from Seismic Data (Current Work)

Seismic Workflow 2 Uncertainties in Interpretation



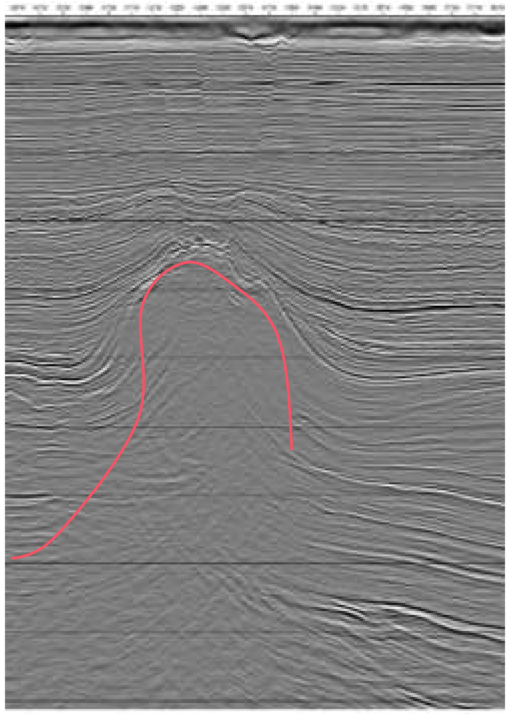


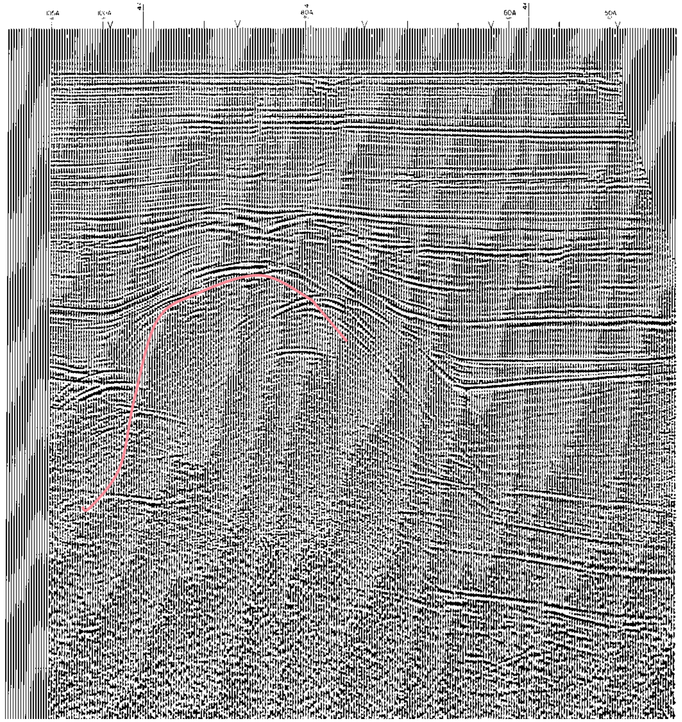


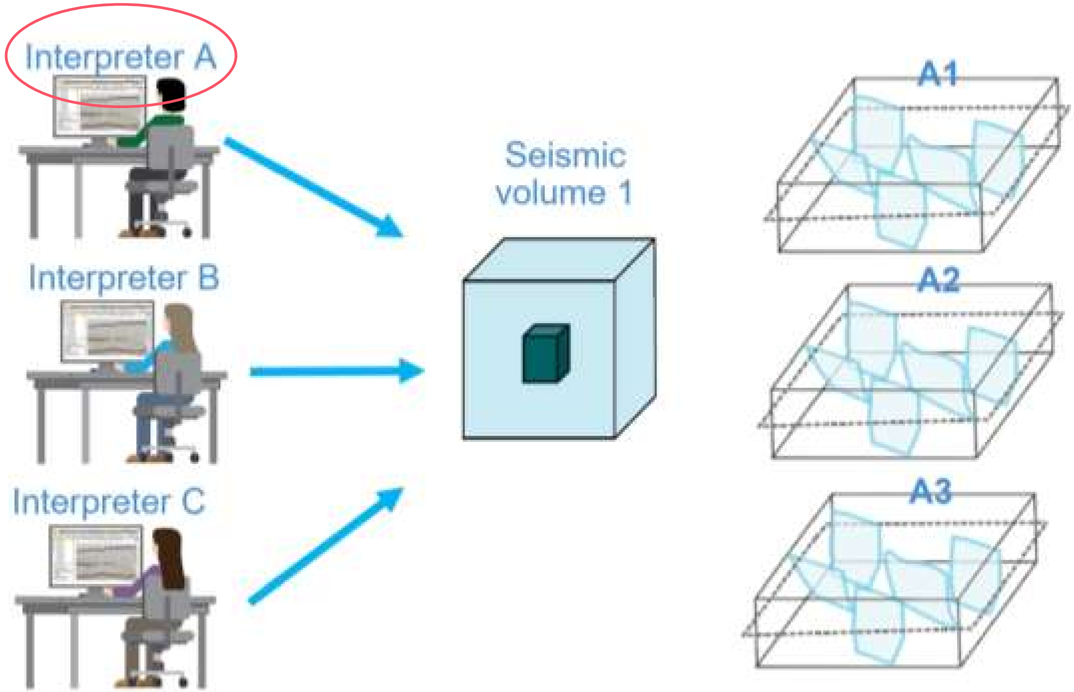
Uncertainty Quantification from Seismic Data (Current Work)

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



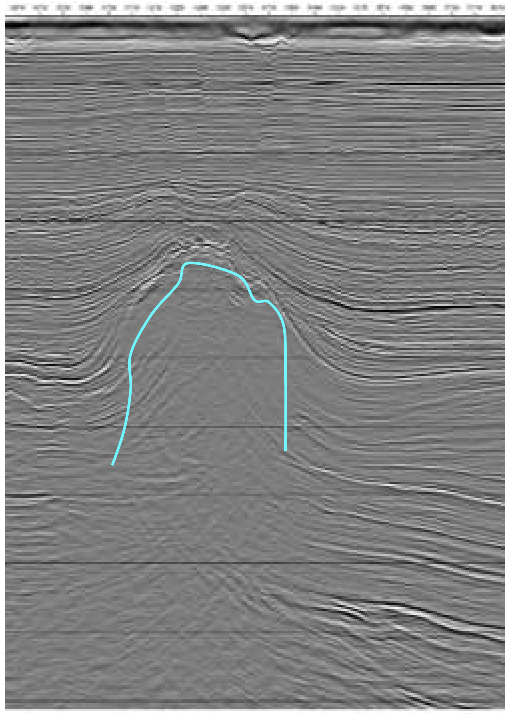


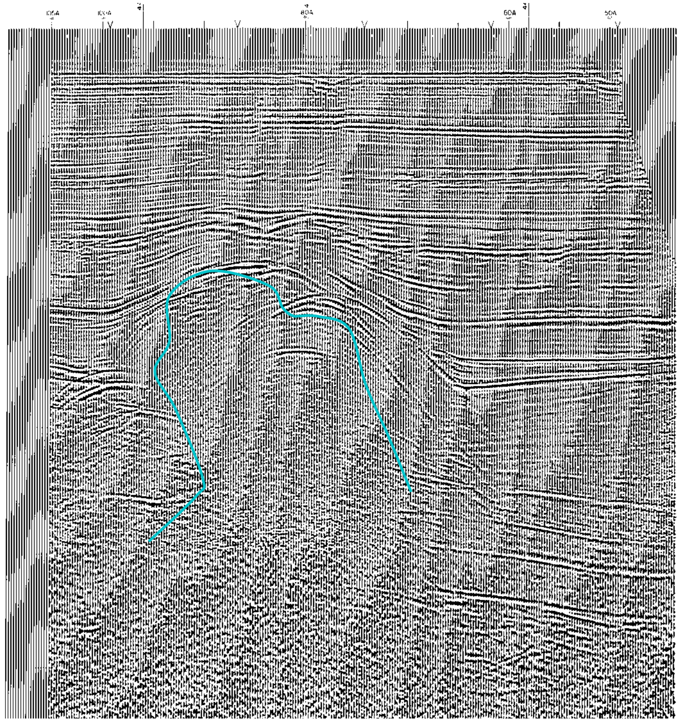


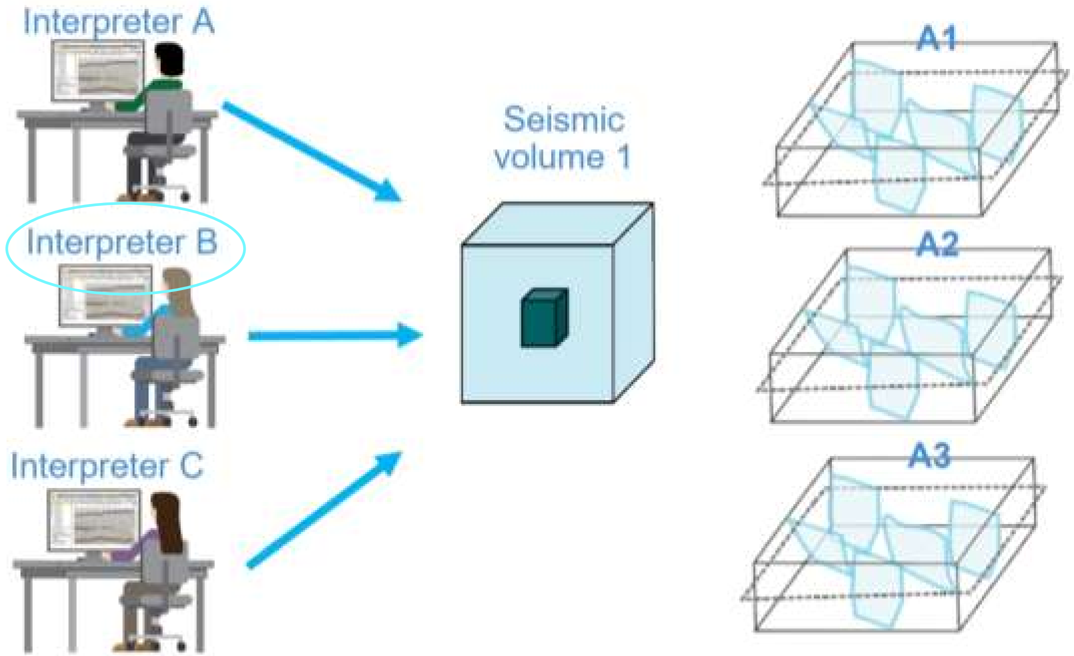
MS 2.3	Testdatensätze extrahiert	M18
MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30

Uncertainty Quantification from Seismic Data (Current Work)

Seismic Workflow 2 Uncertainties in Interpretation



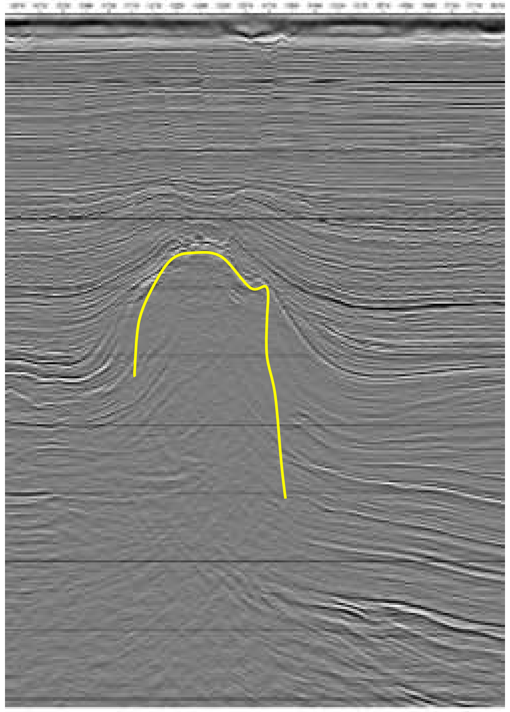


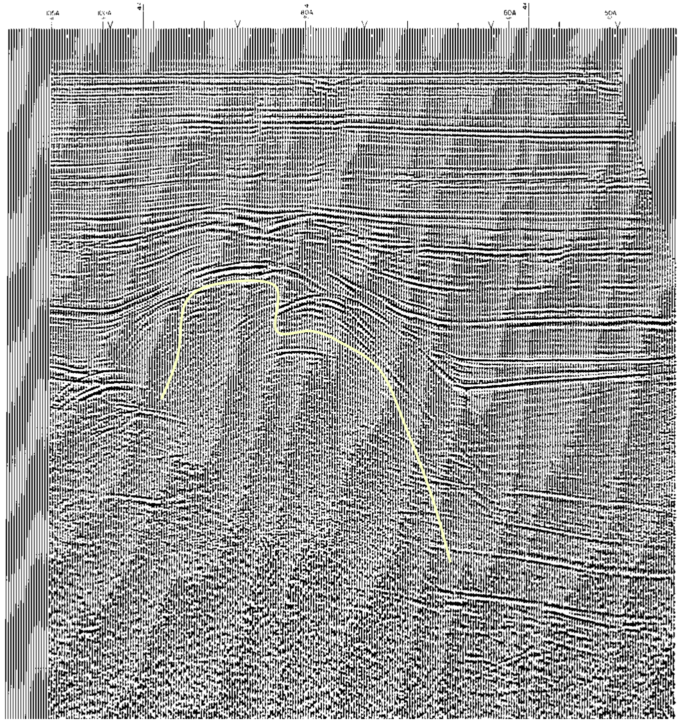


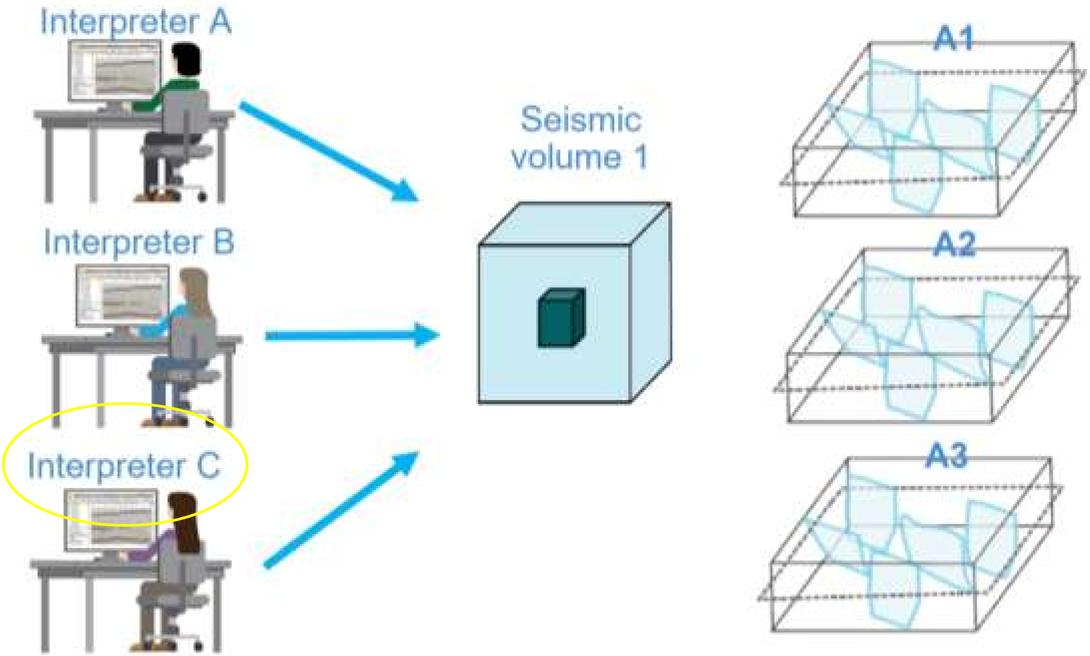
Uncertainty Quantification from Seismic Data (Current Work)

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

Seismic Workflow 2 Uncertainties in Interpretation



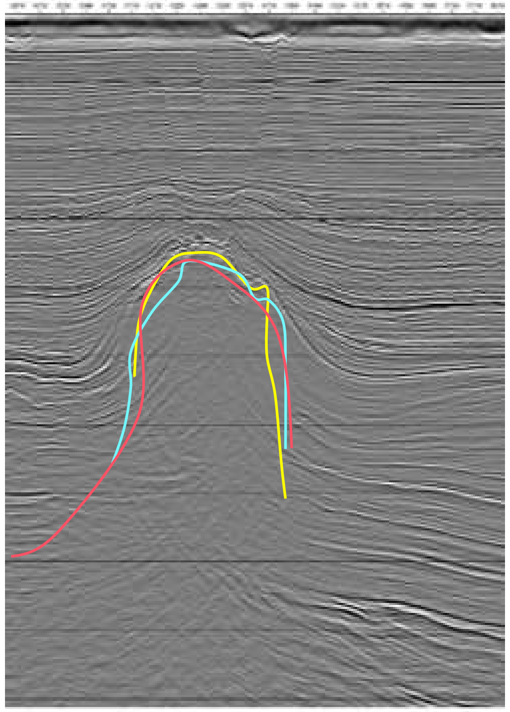


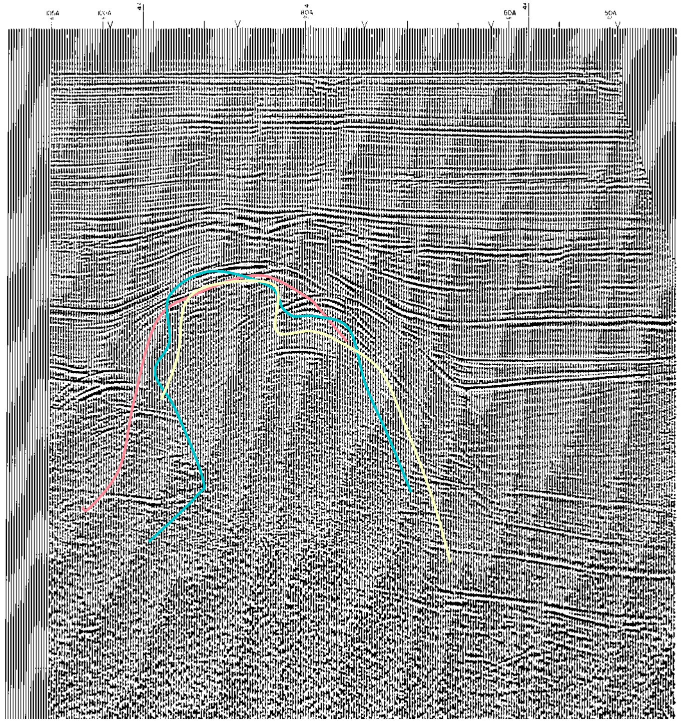


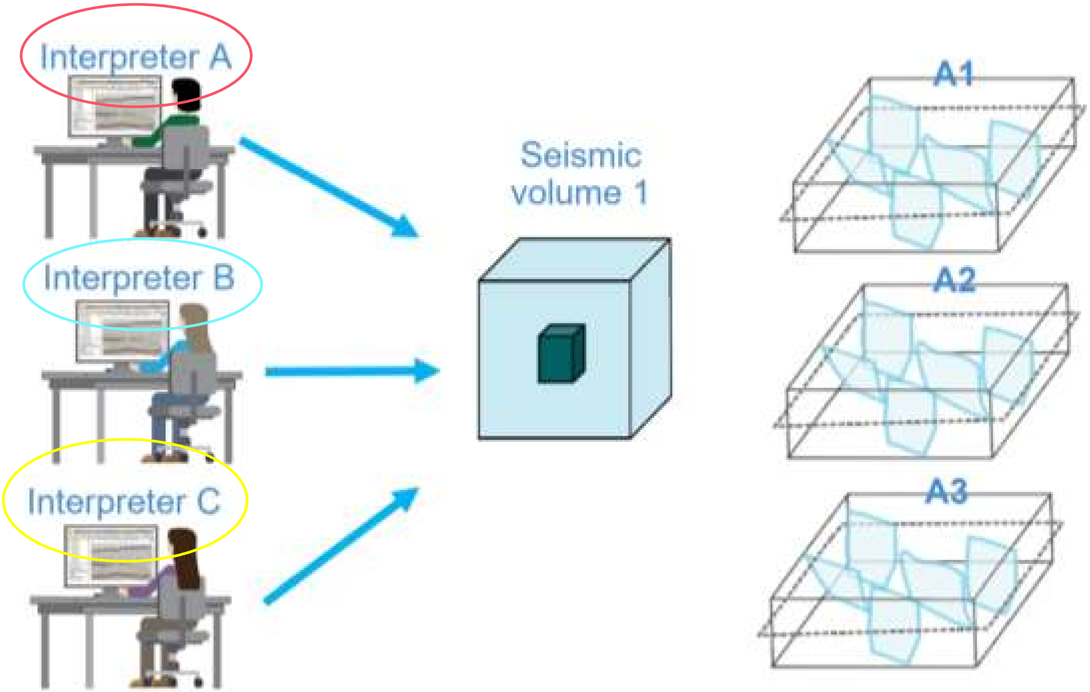
Uncertainty Quantification from Seismic Data (Current Work)

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



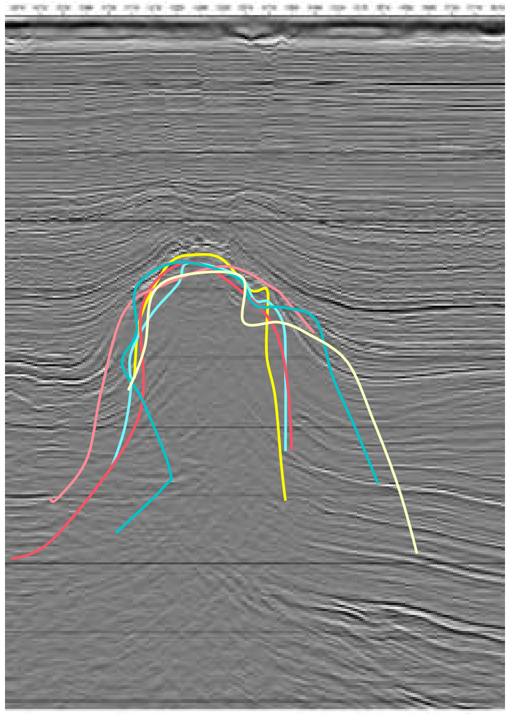


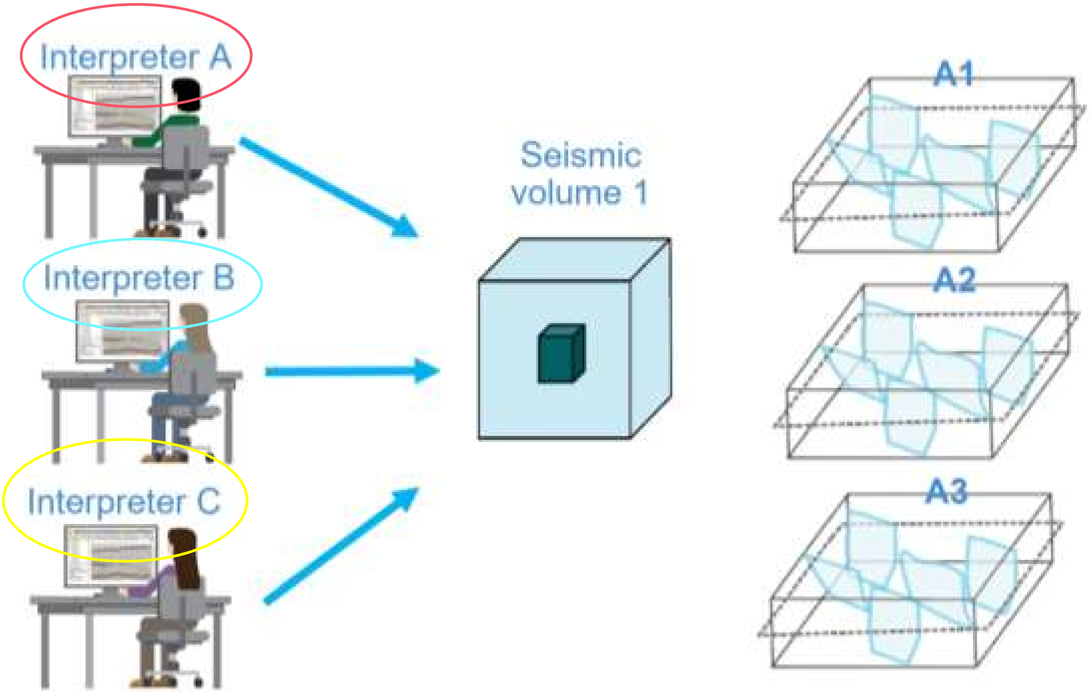


Uncertainty Quantification from Seismic Data (Current Work)

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

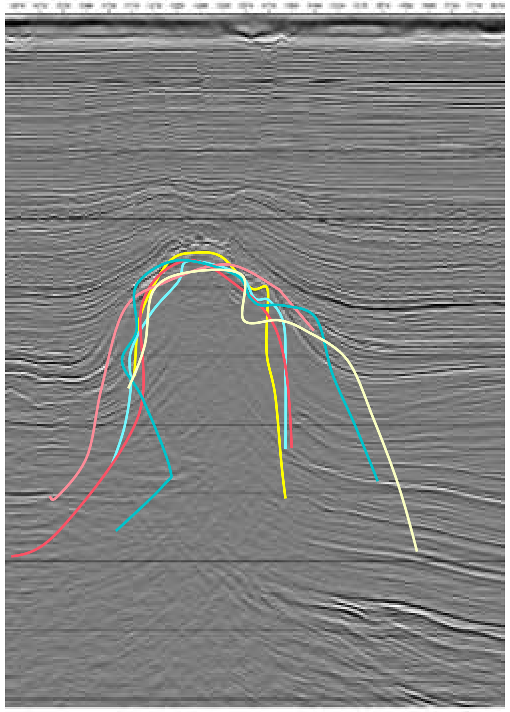




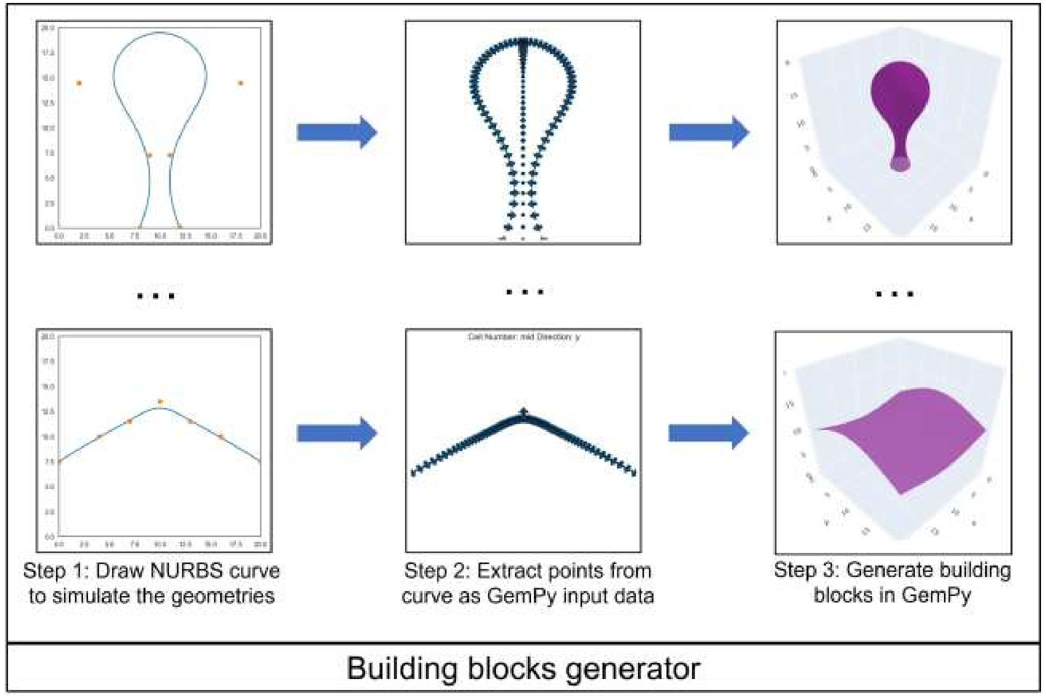
Uncertainty Quantification from Seismic Data (Current Work)

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



+



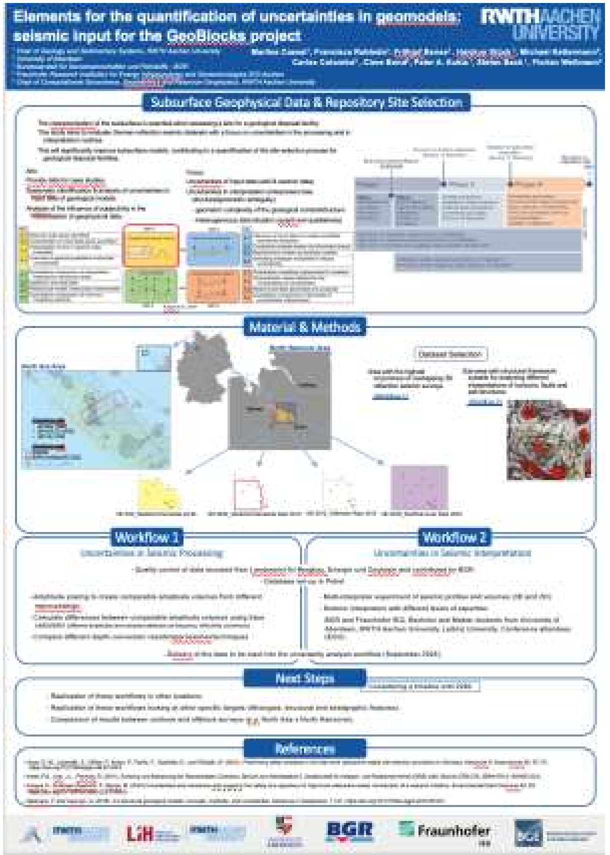
MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

Uncertainty Quantification from Seismic Data (Current Work)

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



Computational Geoscience and Reservoir Engineering

RWTH AACHEN UNIVERSITY

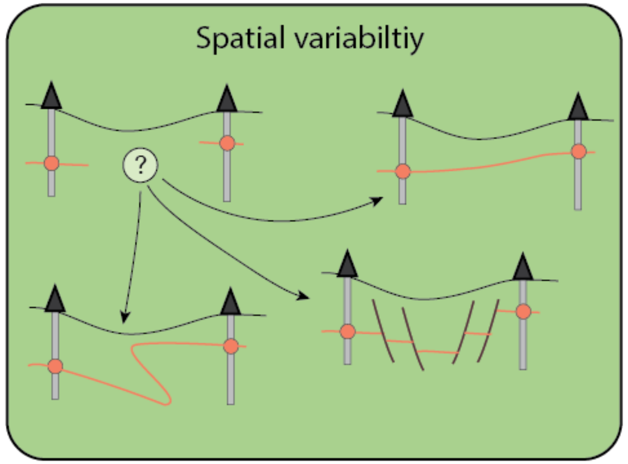
Uncertainty Quantification from Seismic Data (Current Work)

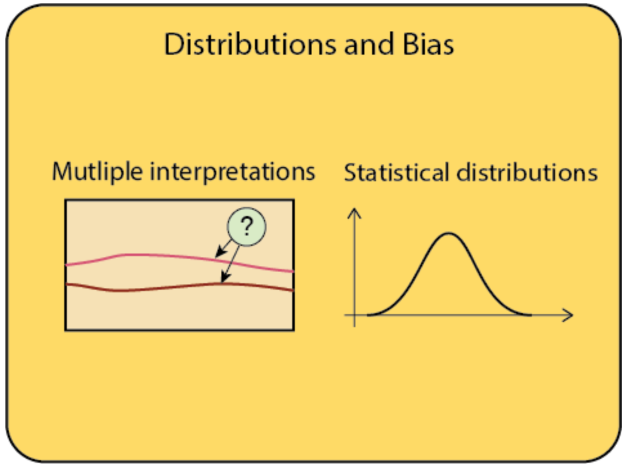
MS 2.1	Verfügbare Daten gesichtet	M12
MS 3.1	Relevante Datentypen identifiziert	M6

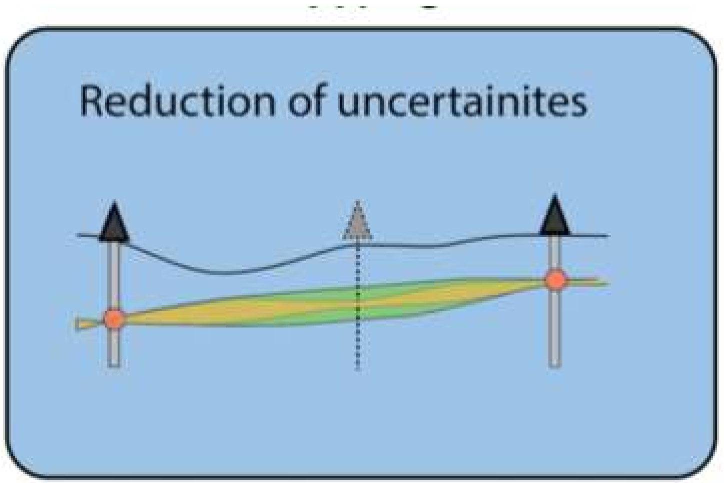
Set of case studies and associated data

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)

START

↓

Interpret seismic data, extract horizon of boundary structure and create acoustic impedance model

↓

Sample variogram range, sill and nugget from uniform distributions

↓

Use previous variogram parameters to create a Gaussian field and add it to the interpreted acoustic impedance model from first step to perturb it

↓

Compute forward modelling to obtain seismic data

↓

Measure difference with seismic data from first step and assign a likelihood

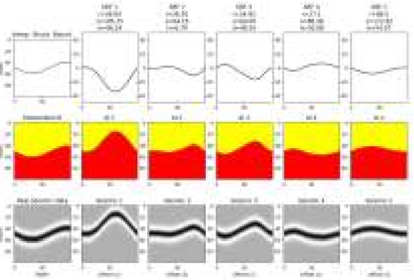
↓

Identify optimal parameters or their distributions

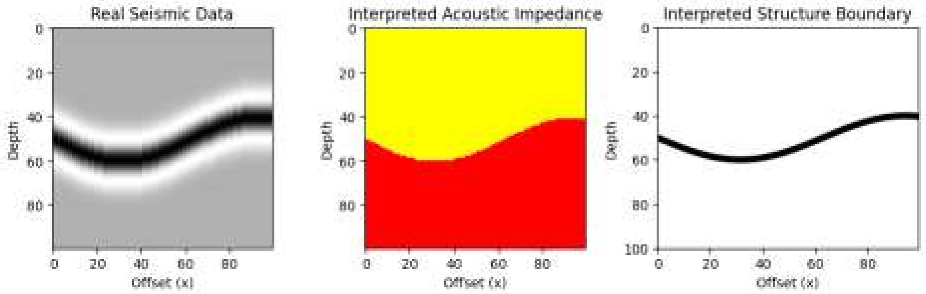
↓

Use SGS to characterize uncertainty

Repeat **n** times



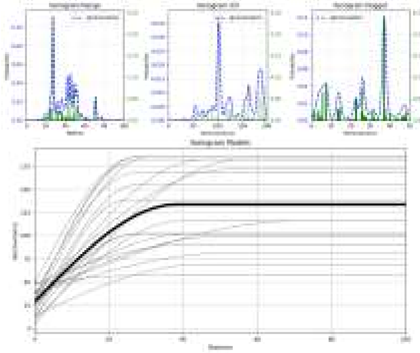
Input Data



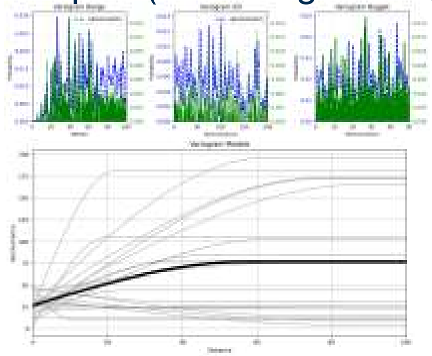
Forward operator: $$\boldsymbol{S_i} = \boldsymbol{WS_i}\,(\mathbf{d}, \theta = 0) = w(\mathbf{d}) * \frac{d\, ln\, \boldsymbol{M_i}(\mathbf{d})}{d\mathrm{d}}$$

Different error measure methods being tested at the moment

Least-squares



Opt. Transport (Sambridge et al., 2022)



Also contributing to:

MS 3.1	Relevante Datentypen identifiziert	M6
MS 2.1	Verfügbare Daten gesichtet	M12
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18

MS 3.2	Ungewissheiten auf Eingangsdatentypen quantifiziert	M18
MS 3.3	Interpretationsfehler bei seismischen Daten bewertet	M30
MS 4.1	Quantitativer Vergleich Interpolations-methoden für Strukturmodell	M18
MS 5.1	Probabilistische Modellierung in Workflow implementiert	M18

Uncertainty Quantification from Seismic Data (C

START

Interpret seismic data, extract horizon of boundary structure and create acoustic impedance model

Sample variogram range, sill and nugget from uniform distributions

Use previous variogram parameters to create a Gaussian field and add it to the interpreted acoustic impedance model from first step to perturb it

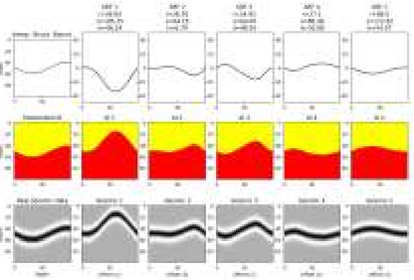
Compute forward modelling to obtain seismic data

Measure difference with seismic data from first step and assign a likelihood

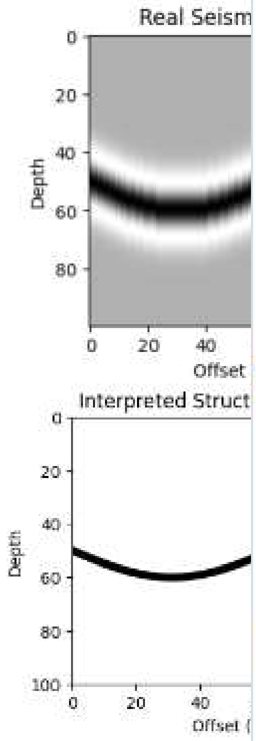
Identify optimal parameters or their distributions

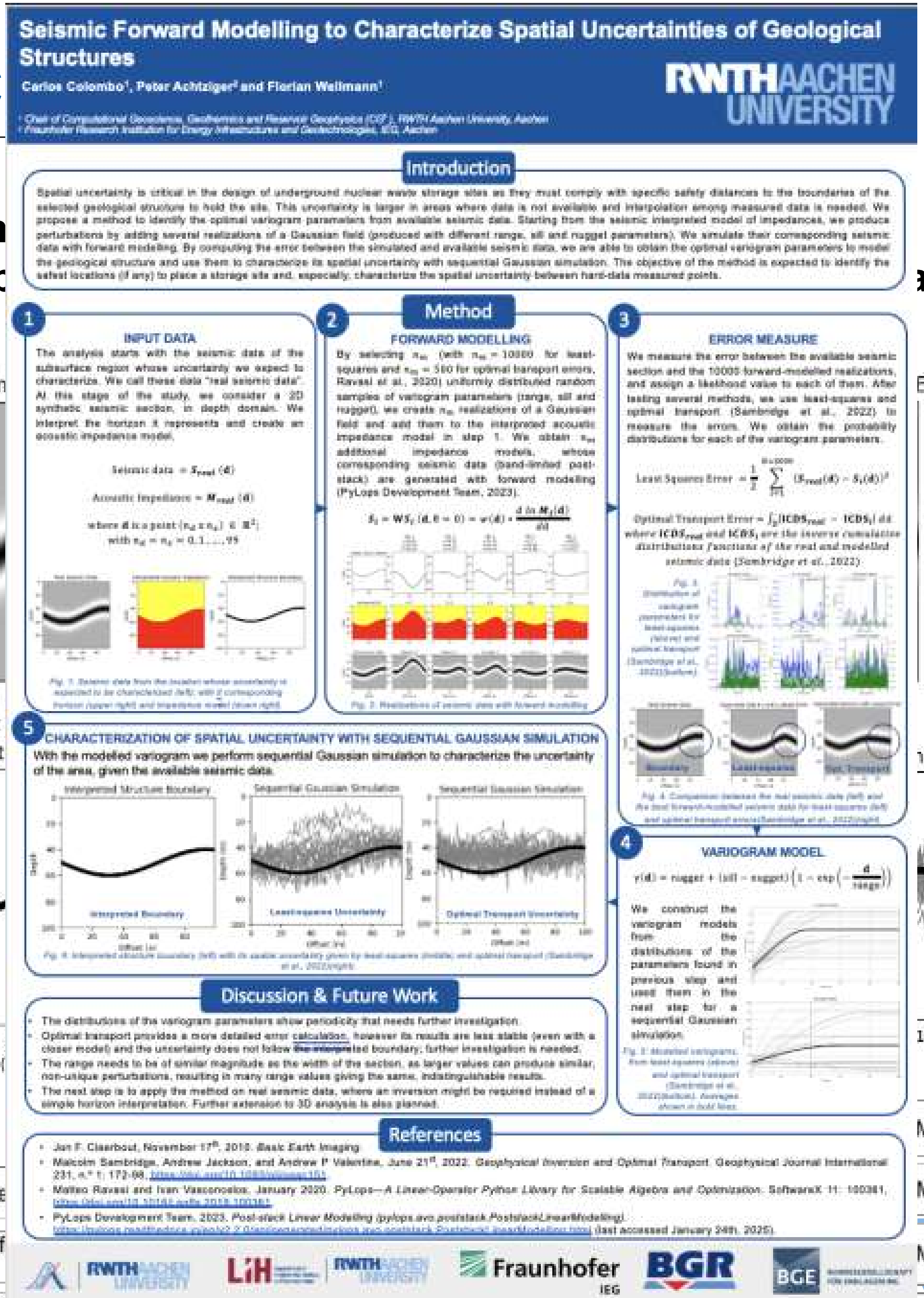
Use SGS to characterize uncertainty

Repeat **n** times



Spatial un
variogram p





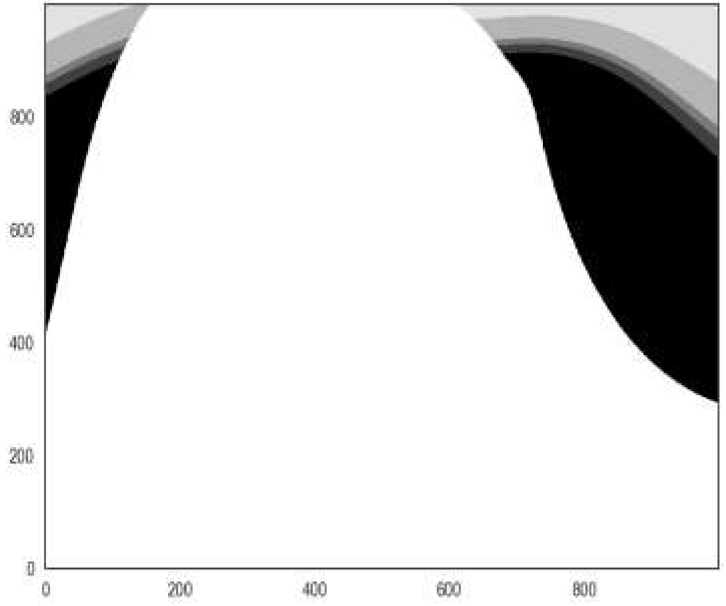
MS 3.1	Rele				M18
MS 2.1	Verf				M30
MS 2.2	Repräsentative Modelle ausgewählt/erstellt	M18	MS 5.1	Probabilistische Modellierung in Workflow implementiert	M18

Error

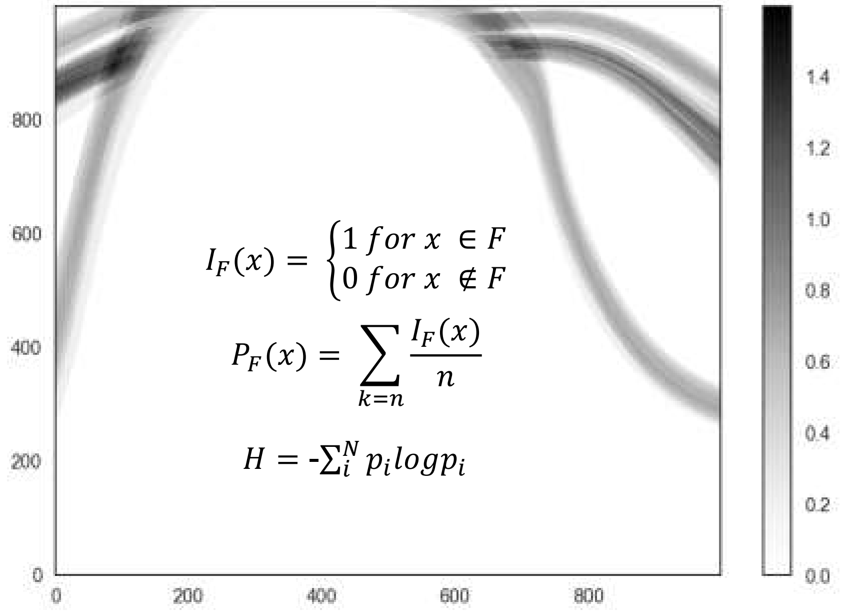
100

Uncertainty Quantification with Gaussian Processes and Information Entropy

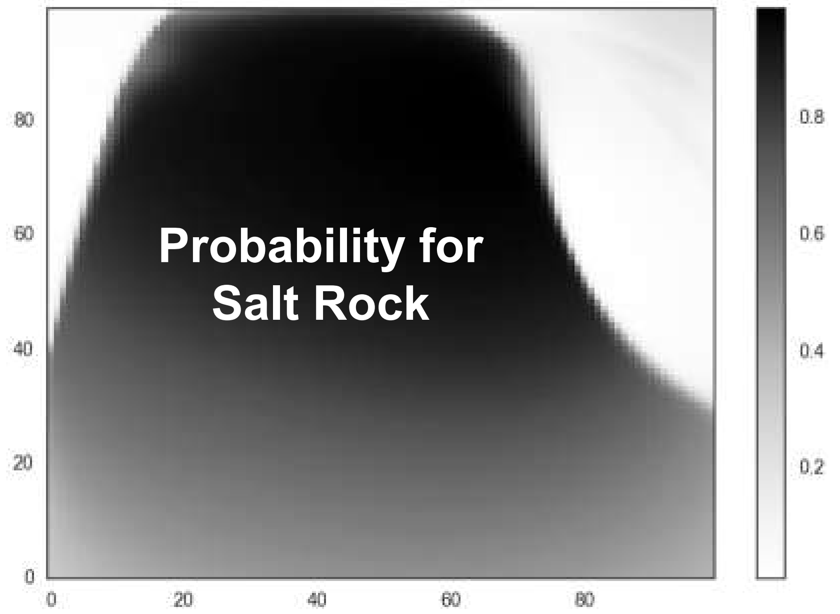
Salt Dome Structure



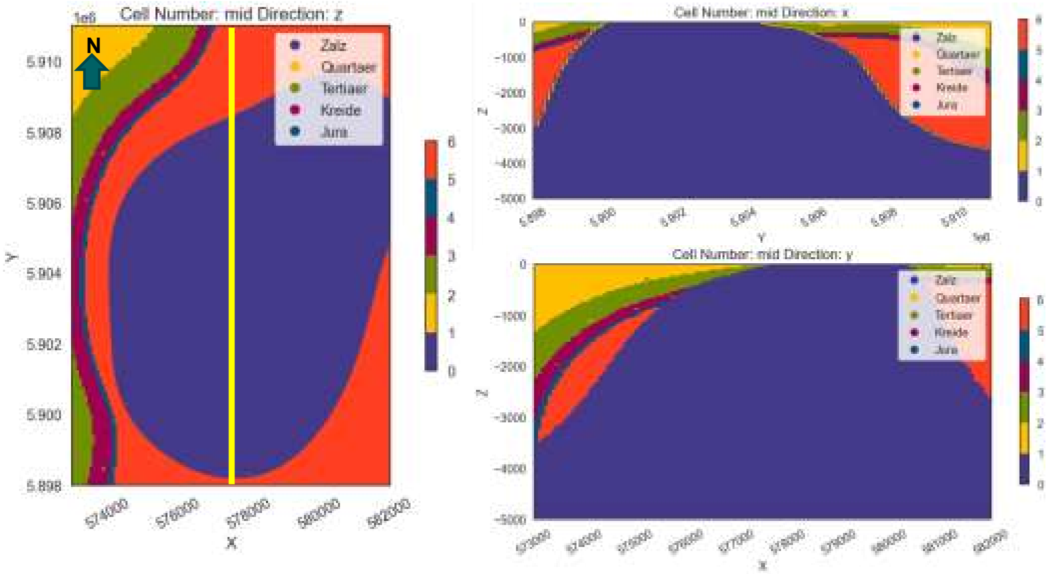
Information Entropy



Gaussian Process Classifier



Formation	Perturbation (m)
Quartaer	100
Tertiaer	100
Kreide	100
Jura	100
Salz	200



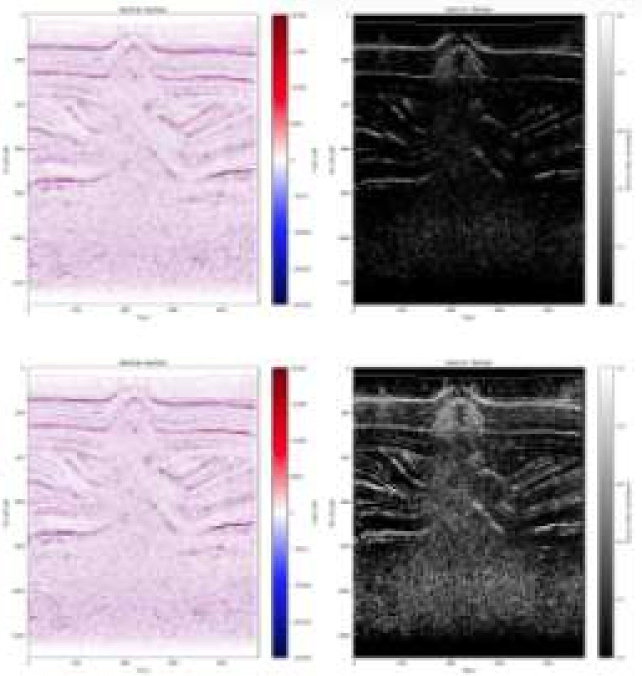
Also contributing to:

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 Interpolations-methoden 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).

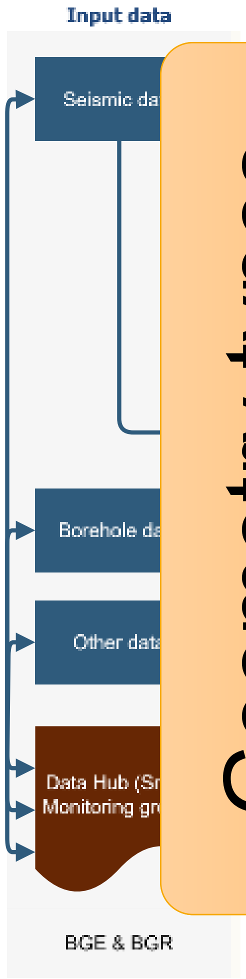
Seismic Attribute Testing with PyLops



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



Geometry types

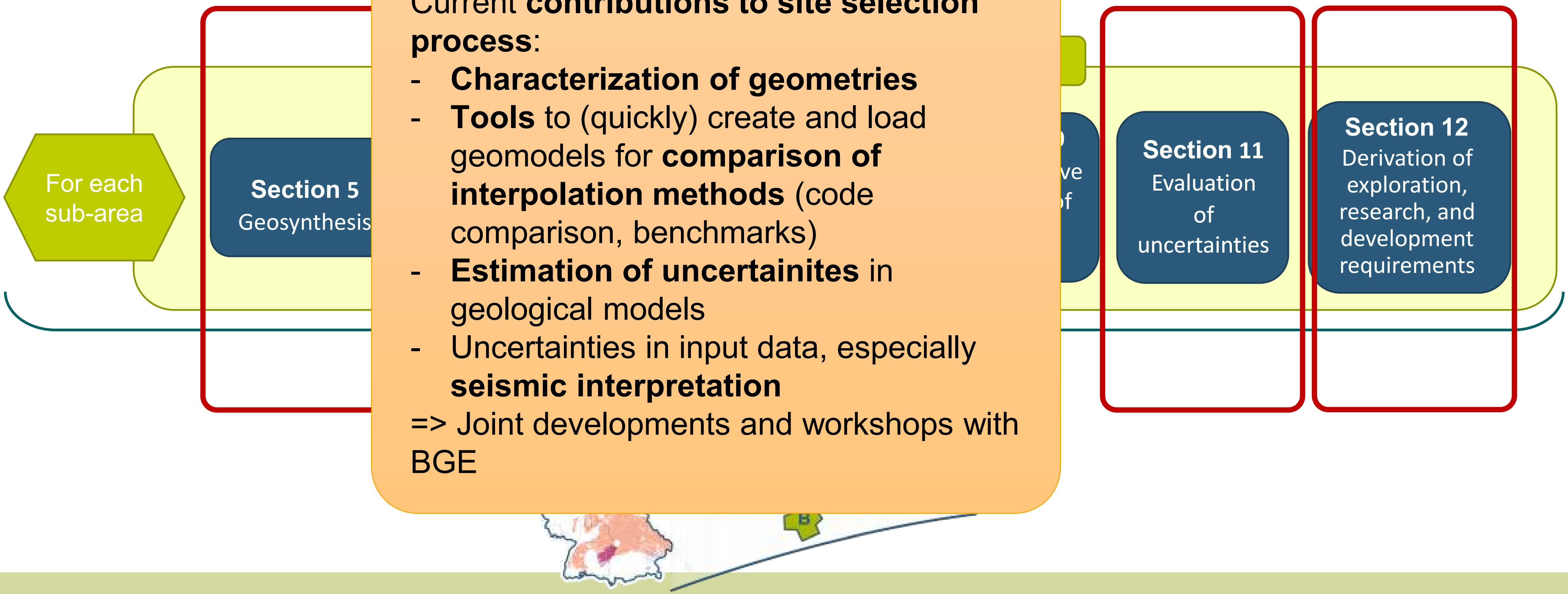
Input data uncertainties

Interpolation uncertainties

Comparison and reduction

Method and process | Outcome and result | Possible collaboration from other URS project | Possible input from other researchers

The representative preliminary safety assessments (rvSU)



Current **contributions to site selection process**:

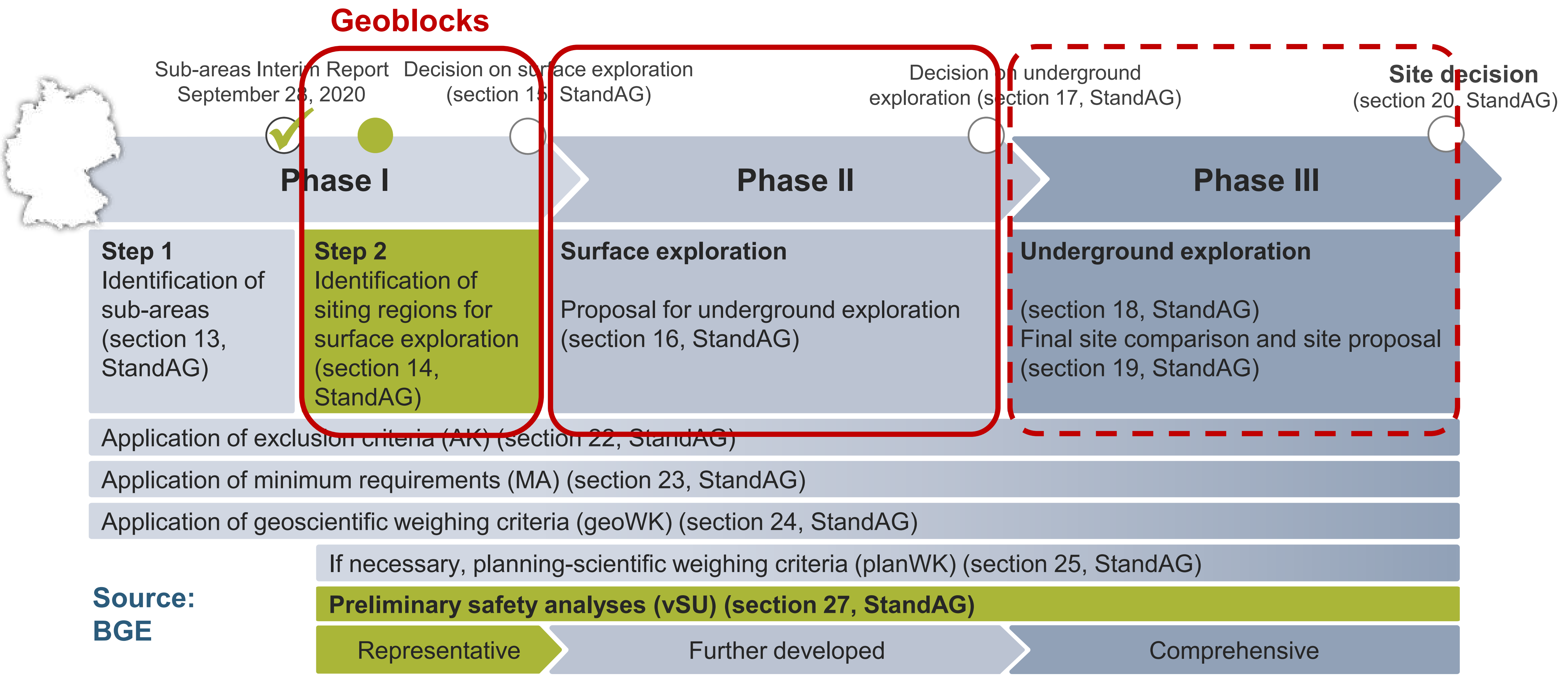
- **Characterization of geometries**
- **Tools** to (quickly) create and load geomodels for **comparison of interpolation methods** (code comparison, benchmarks)
- **Estimation of uncertainites** in geological models
- Uncertainties in input data, especially **seismic interpretation**

=> Joint developments and workshops with BGE

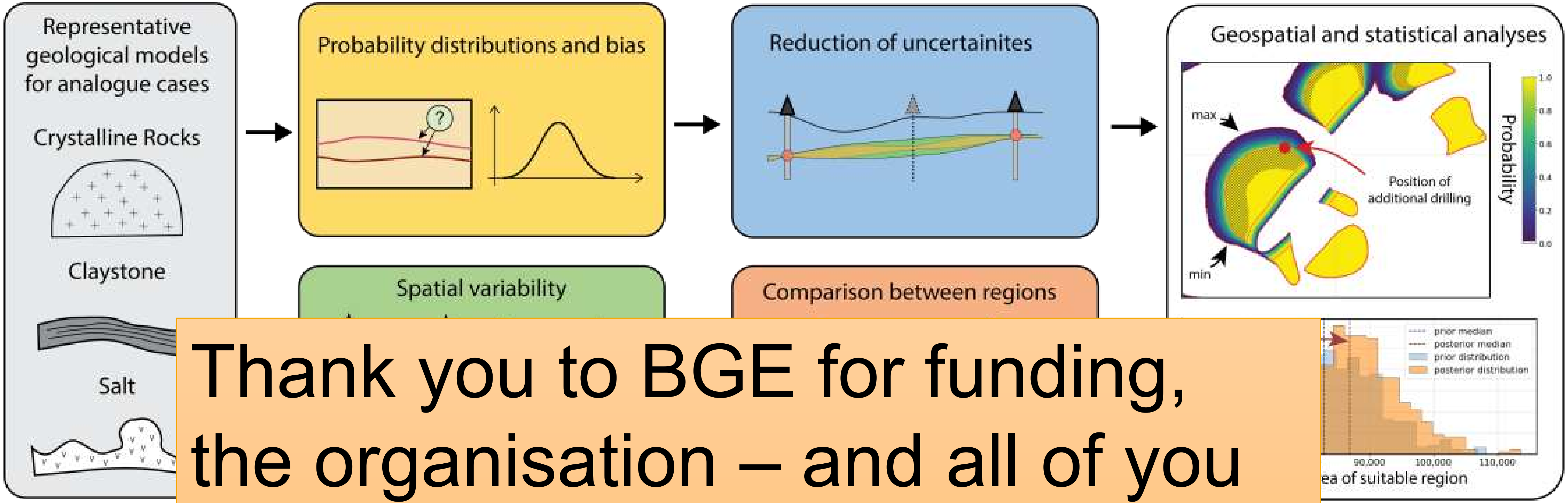
Step-by-step narrowing down of sub-areas to site regions using criteria-based approach

Source: BGE

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 uncertianty?

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

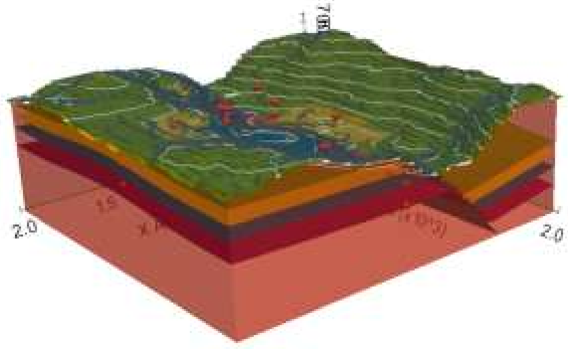
- Revision of geomodelling concepts and interpolation methods
- Review Bayesian inversion
- Started building general GemPy models
- Write draft of project data management plan

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

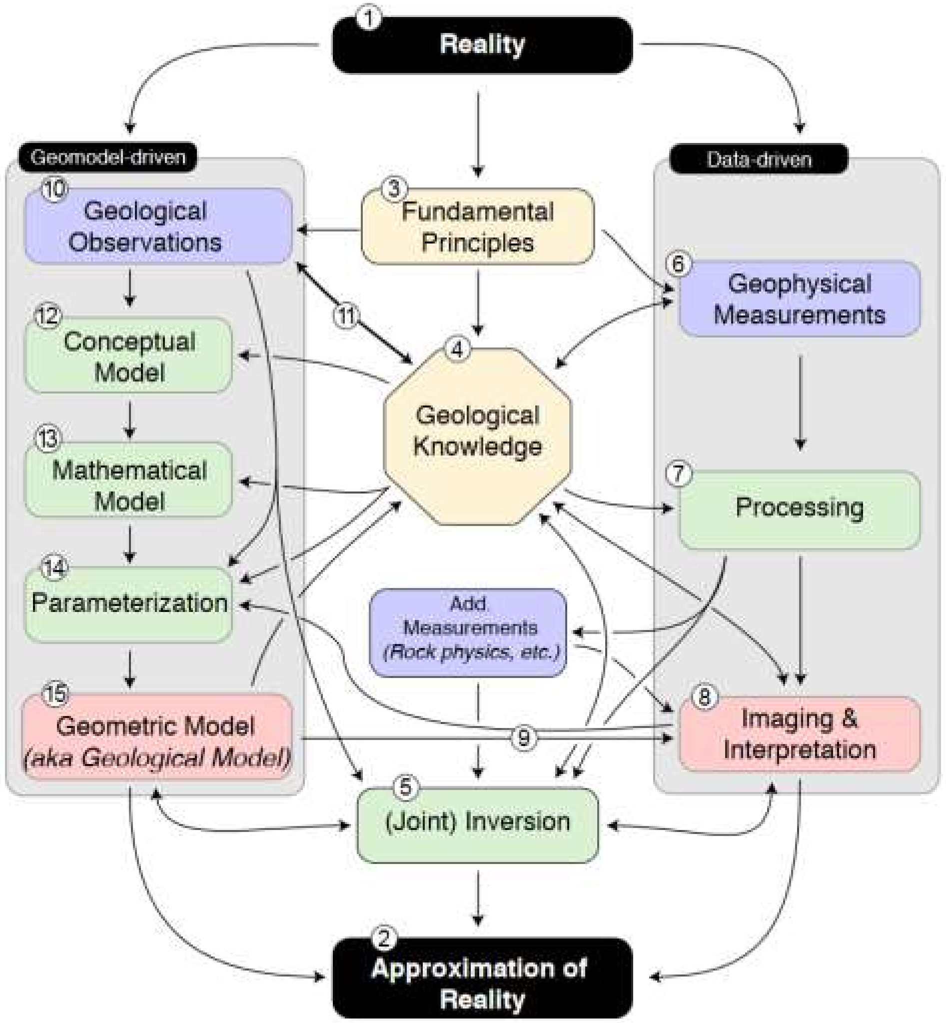
- Built GIS Database

Also contributing to:

MS 1.1	Plan für Forschungsdaten-management 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 Interpolations-methoden für Strukturmodell	M18



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



Wellmann and Caumon, 2018

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

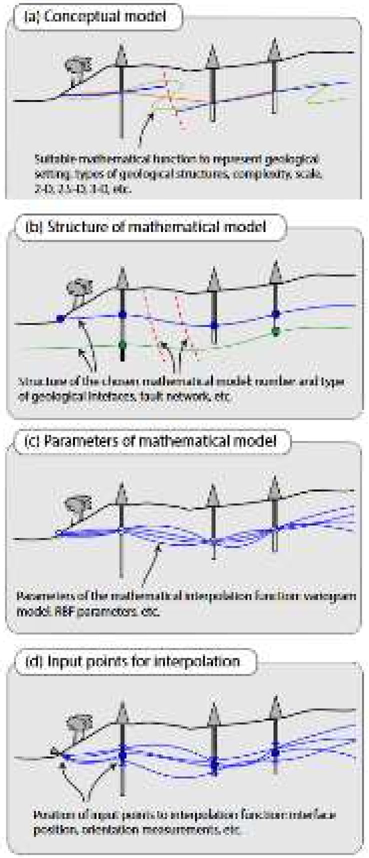
Witter et al. 2018, Wellmann and Caumon 2018

<u>Classification of Wellmann and Caumon, 2018, after Mann 1993 and Cox 1982</u>:

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



Wellmann and Caumon 2018

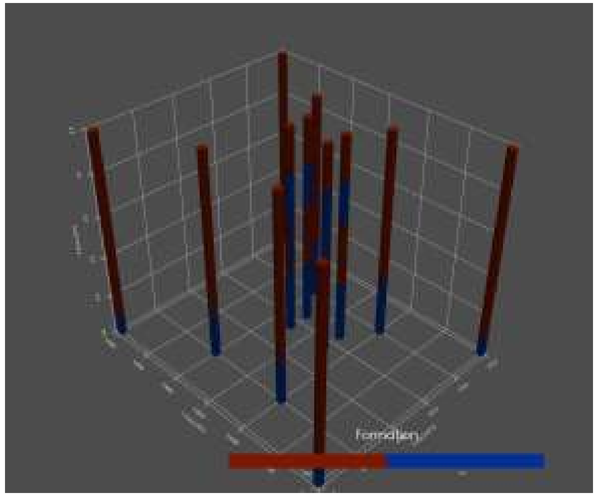
Reviewed Concepts and Methods:

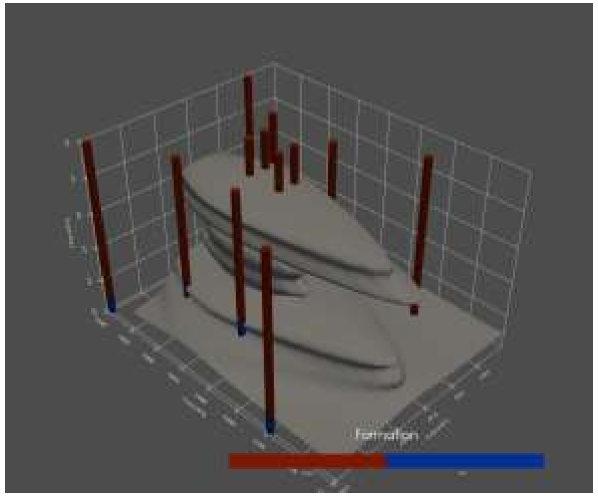
Bayesian Inference
Approximate Bayesian Computation
Geological Model-based Uncertainties
Information Entropy (Shannon Entropy)
Gaussian Processes
Sequential Gaussian Simulation
Monte Carlo (Model Based)
Markov chain Monte Carlo
Uncertainty from Seismic Data
Spatial Diffusion
Neural-Network-based Multi-step 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

<u>Also contributing to</u>:

MS 3.1	Relevante Datentypen identifiziert	M6
MS 4.1	Quantitativer Vergleich Interpolations-methoden für Strukturmodell	M18

Synthetic salt dome





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

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

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

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

$$\log p(\mathbf{z}) = -\frac{1}{2}\mathbf{z}\boldsymbol{T}(\mathbf{Kff} + \sigma^2\mathbf{I}N) - 1 \ \mathbf{z} - \frac{1}{2}\log|\mathbf{Kff} + \sigma^2\mathbf{I}N| \ -\frac{N}{2}log 2\pi$$

$$p(\mathbf{f}, \mathbf{u}) = NN + U \left(\begin{bmatrix}\mathbf{f}\\ \mathbf{u}\end{bmatrix}; \mathbf{0}, \begin{bmatrix}\mathbf{Kff} & \mathbf{Kfu}\\ \mathbf{Kuf} & \mathbf{Kuu}\end{bmatrix}\right)$$

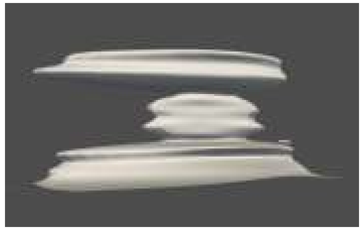
$$p(\mathbf{z}, \mathbf{f}, \mathbf{u}) \propto p(\mathbf{z}|\mathbf{f}, \mathbf{u})p(\mathbf{f}, \mathbf{u}) = p(\mathbf{z}|\mathbf{f})p(\mathbf{f}|\mathbf{u})p(\mathbf{u})$$

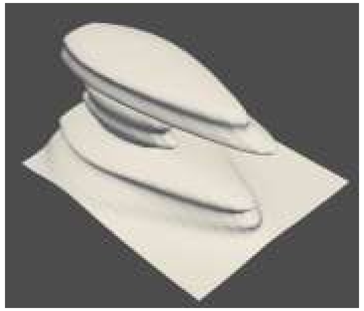
to maximize $p(\mathbf{z}), p(\mathbf{f}|\mathbf{u})$ is replaced by $q(\mathbf{f}|\mathbf{u})$

Variational inference approximates the posterior as a Gaussian $\boldsymbol{q}(\boldsymbol{f}, \boldsymbol{u}) = \boldsymbol{p}(\boldsymbol{f}|\boldsymbol{u})\boldsymbol{q}(\boldsymbol{u})$
by introducing the variational distribution
$\boldsymbol{q}(\boldsymbol{u}) = NU\ (\boldsymbol{u};\ \boldsymbol{m}, \boldsymbol{S})$, whose mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{S}$
are variational parameters to optimize

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

Gonçalves et al. 2022





Also contributing to:

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

Probabilistic Modelling for Uncertainty Quantification

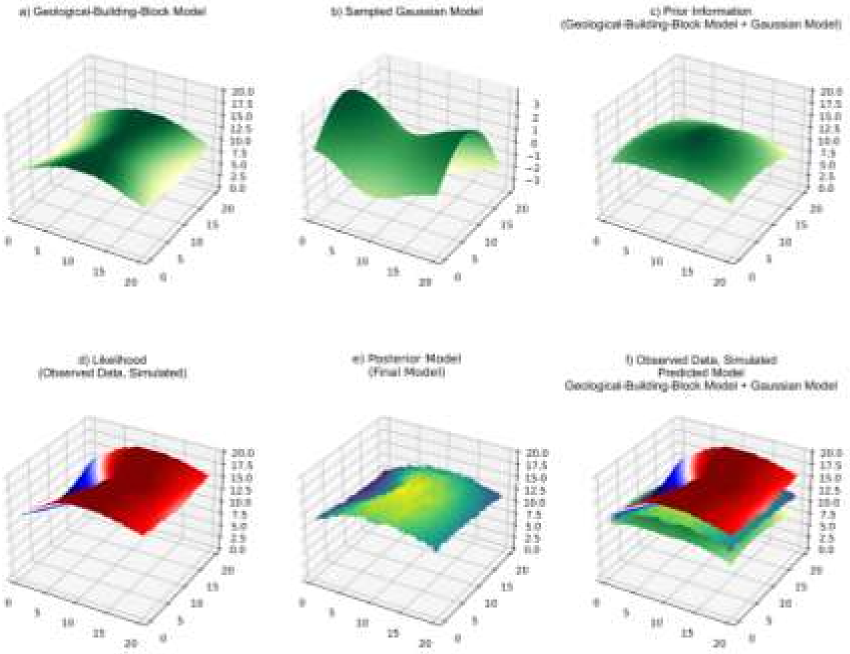
Coordinates & Orientation → **GemPy Model** → **ODSIM Approach** → **Spatial Uncertainty Quantification** → **Uncertainty Reduction**

Seismic Borehole data → **Large-scale features** → **Small-scale features**

IAMG 2023

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

Python Programming language · GemPy · PyMC3 · NumPy · matplotlib