GeoBlocks: Building blocks for the quantification of uncertainties in geological models

Friedrich Carl (**LIH)** Lehrstuhl für Ingenieurgeologie und Hydrogeologie, RWTH Aachen
Jian Yang (**CG3**) Computational Geoscience, Geothermics and Reservoir Geophysics, RWTH Aachen
Carlos Colombo (**CG3**) Computational Geoscience, Geothermics and Reservoir Geophysics, RWTH Aachen
Marlise Colling Cassel (**GIA**) Lehrstuhl für Geologie und Paläontologie und Geologisches Institut, RWTH Aachen

13.06.2023

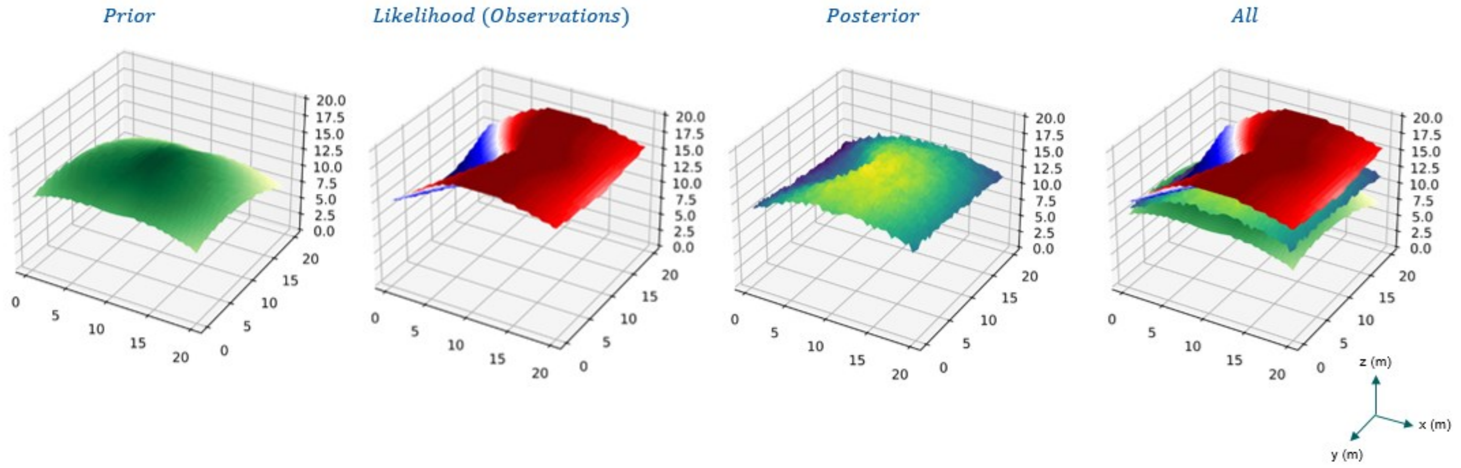
URS2024

Outline

- Recapitulation
- Work Progress - Uncertainty Analysis
- Discussion
- Next Steps

Recapitulation

- Data Management Plan (DMP) (first version issued now).
- GIS Database (ongoing).
- Revision of explicit and implicit geomodelling methods (ongoing).
- Model building with GemPy (ongoing).
- Probabilistic modelling.



Python GemPy PyMC3 NumPy matplotlib

Uncertainty Analysis - Types of Uncertainty

Epistemic Uncertainty

State/lack of knowledge, can be reduced with observations

Aleatory Variability or Uncertainty

Unpredictability due to inherent randomness, cannot be reduced

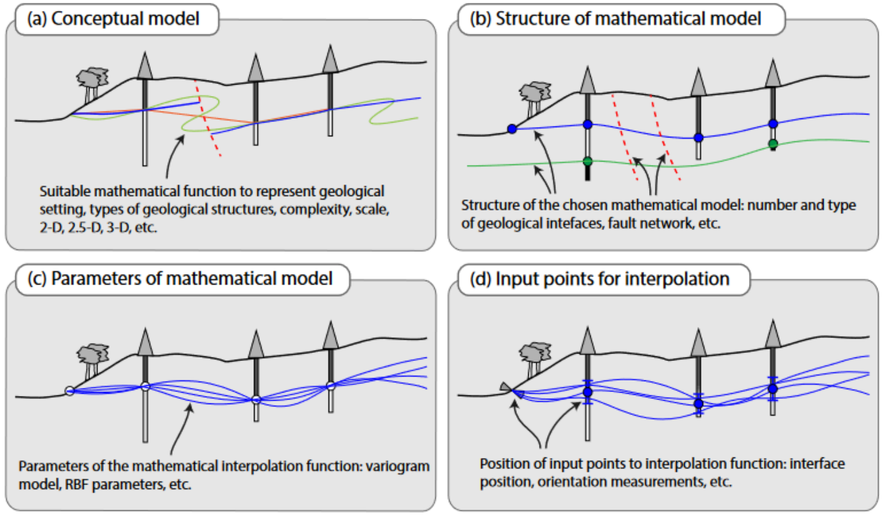
Witter et al. 2018, Wellmann and Caumon 2018

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

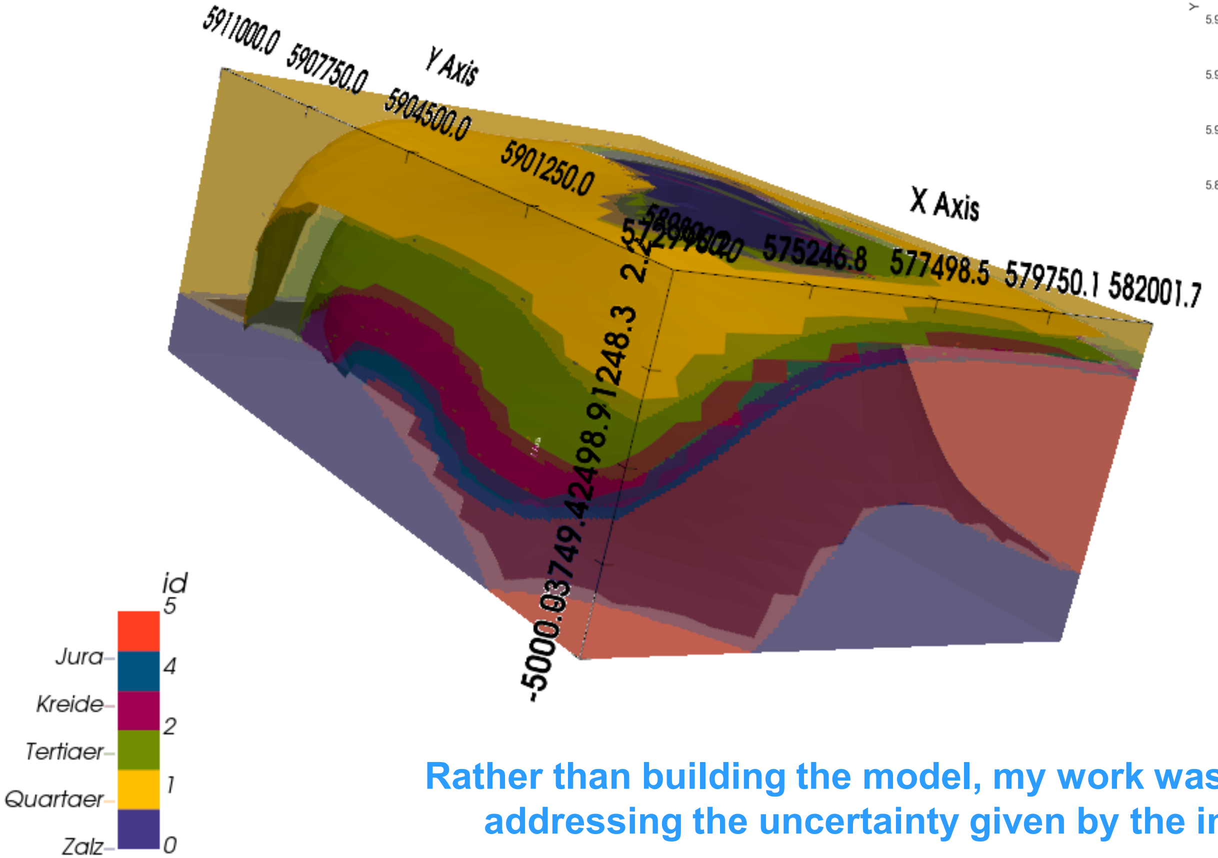
Type 3 - Imprecise knowledge
Problem of incomplete knowledge of structures in subsurface

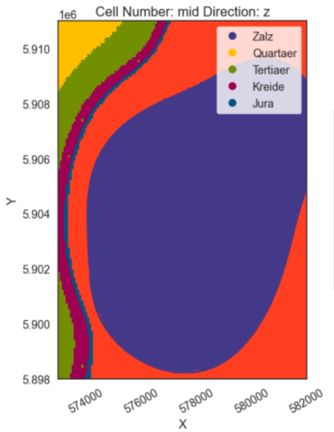


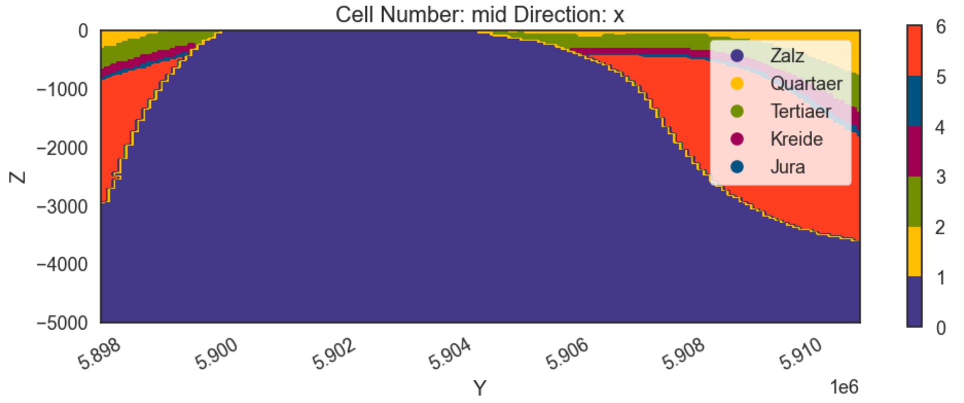
Wellmann and Caumon 2018

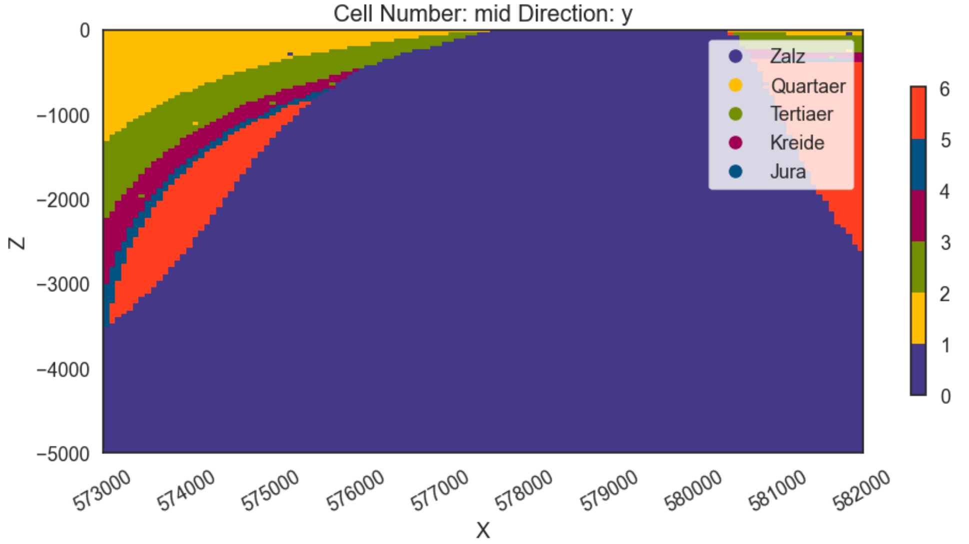
Uncertainty Analysis - Model building with GemPy 3

- Borehole data









Rather than building the model, my work was focused on addressing the uncertainty given by the input data

Uncertainty Analysis - Model Building and UQ Methods Tested

1. An initial model, with boreholes and an edited salt dome structure was created.

2. **Uncertainty from the data (Type 1)**

   A 2D section was extracted from the model in point 1 to perform the following independent approaches:

   a) Quantification of uncertainty with an entropy computation of the model: 20 realizations in a 1000x1000 grid and 20 in a 100x100 grid.

   b) ~~Quantification of uncertainty with a Gaussian process, using the interfaces/layer boundaries of model in point 1 for training, in a 100x100 grid.~~

   c) Quantification of uncertainty with a Gaussian process classifier, using the formations of model in point 1 for training, in a 100x100 grid, resulting in probabilities for being in a specific formation (Salt in the examples that follows).

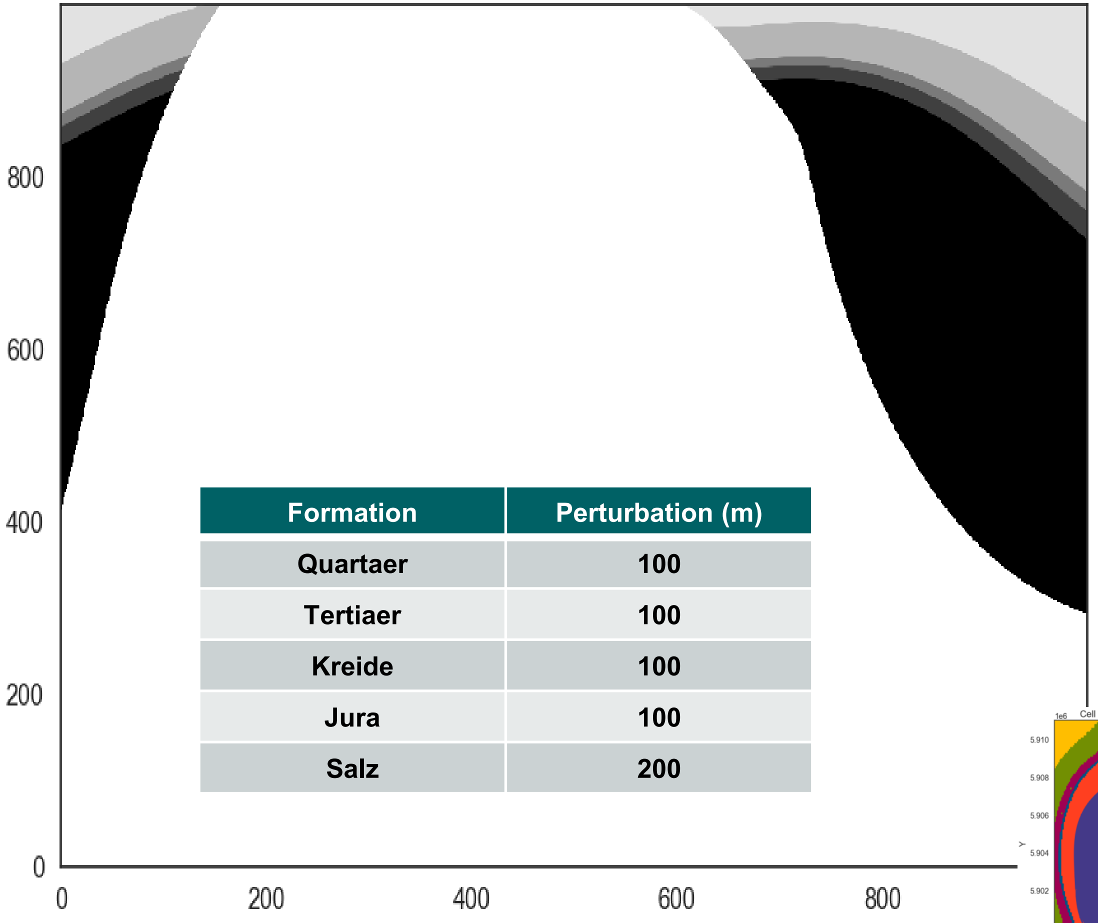
3. **Uncertainty from the interpolation method (Type 2)**

   a) Gaussian process.

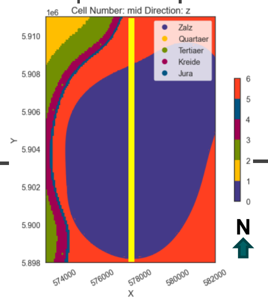
   b) Variational Gaussian process.

Uncertainty Analysis - Entropy

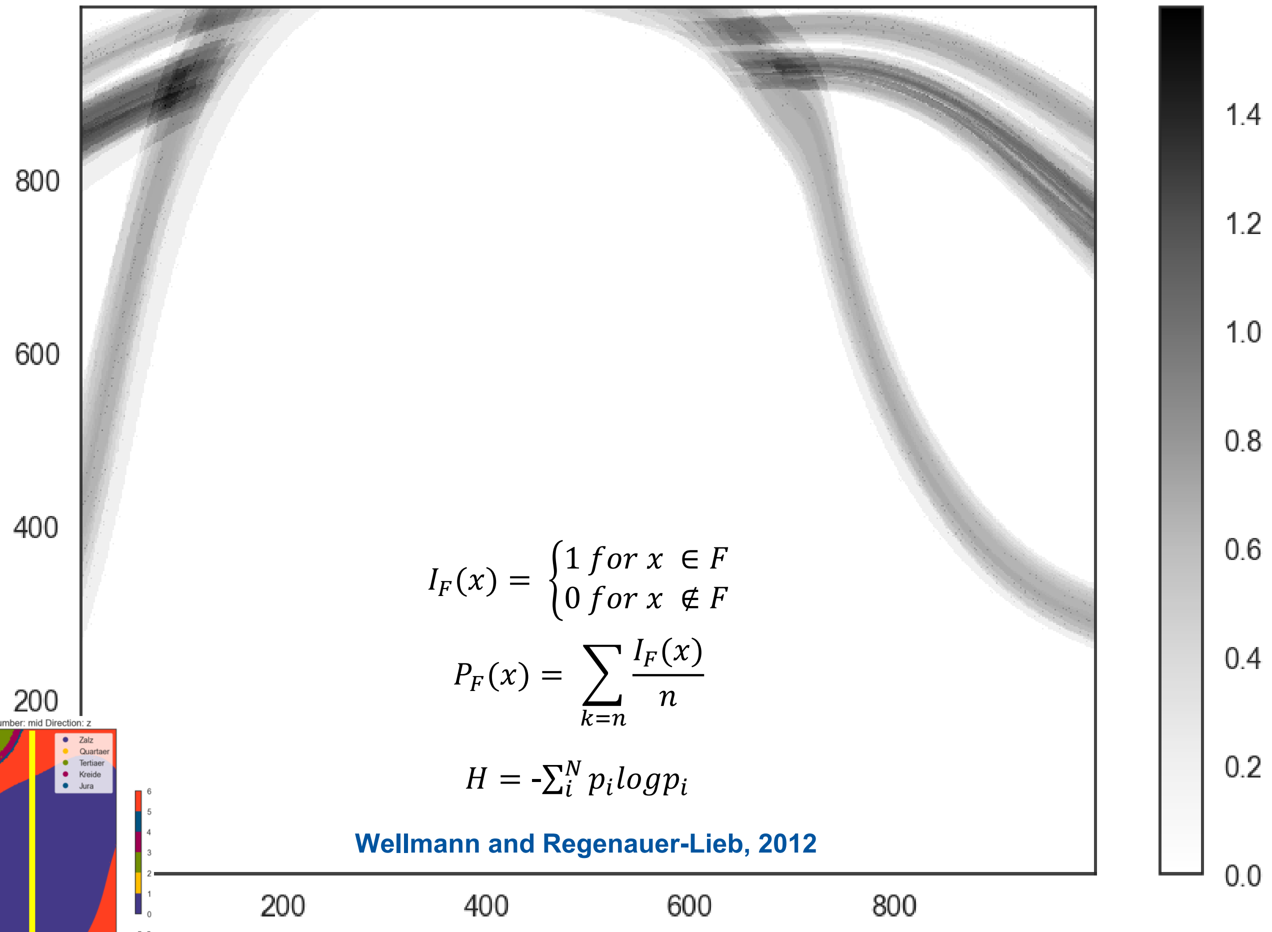
Section



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



Entropy



$$I_F(x) = \begin{cases} 1 \; for \; x \in F \\ 0 \; for \; x \notin F \end{cases}$$

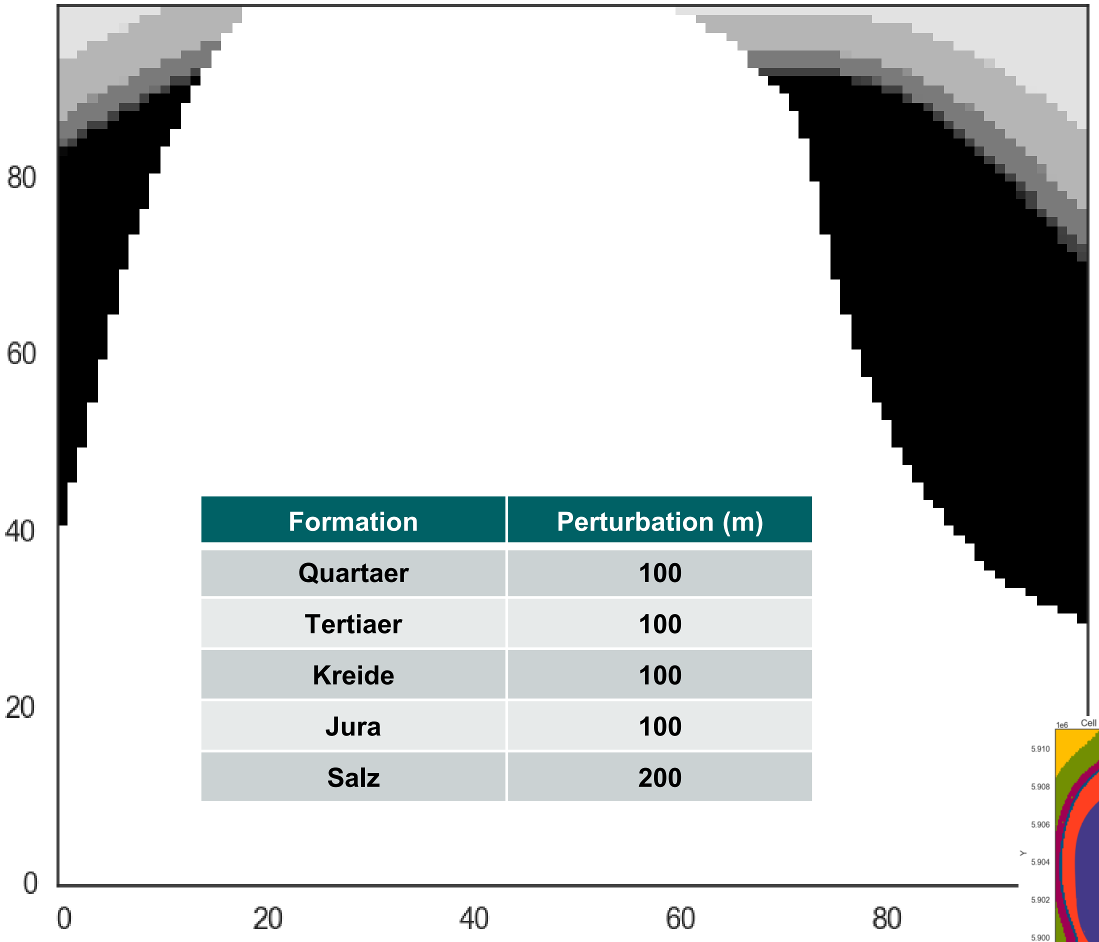
$$P_F(x) = \sum_{k=n} \frac{I_F(x)}{n}$$

$$H = -\sum_i^N p_i log p_i$$

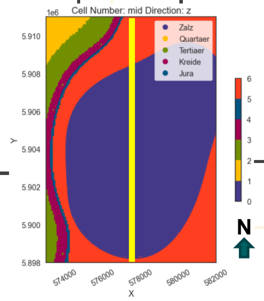
Wellmann and Regenauer-Lieb, 2012

Uncertainty Analysis - Entropy

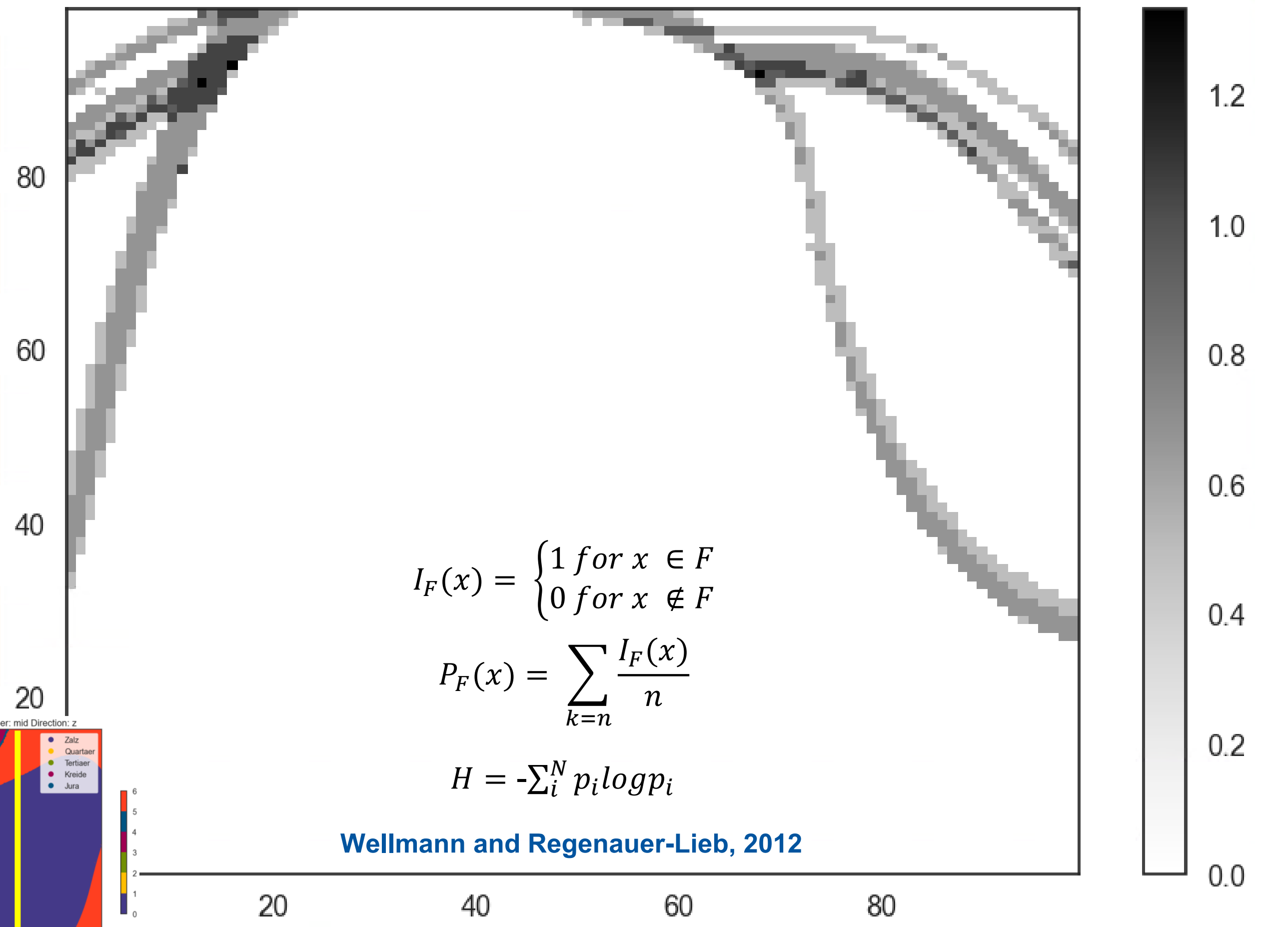
Section



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



Entropy



$$I_F(x) = \begin{cases} 1 \; for \; x \in F \\ 0 \; for \; x \notin F \end{cases}$$

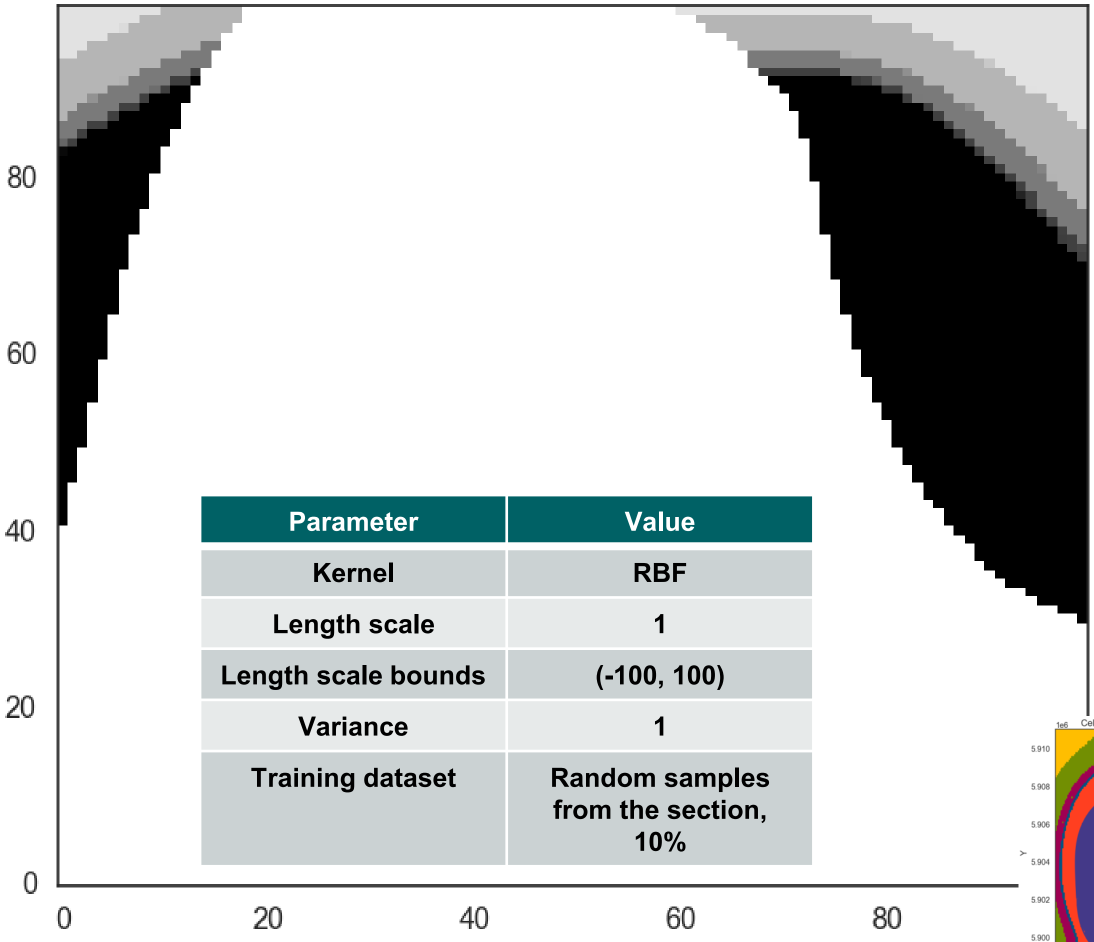
$$P_F(x) = \sum_{k=n} \frac{I_F(x)}{n}$$

$$H = -\sum_i^N p_i log p_i$$

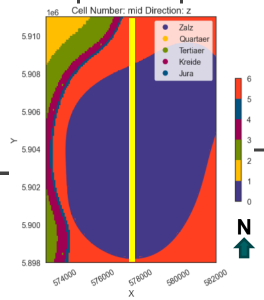
Wellmann and Regenauer-Lieb, 2012

Uncertainty Analysis – Gaussian Process - *Dismissed*

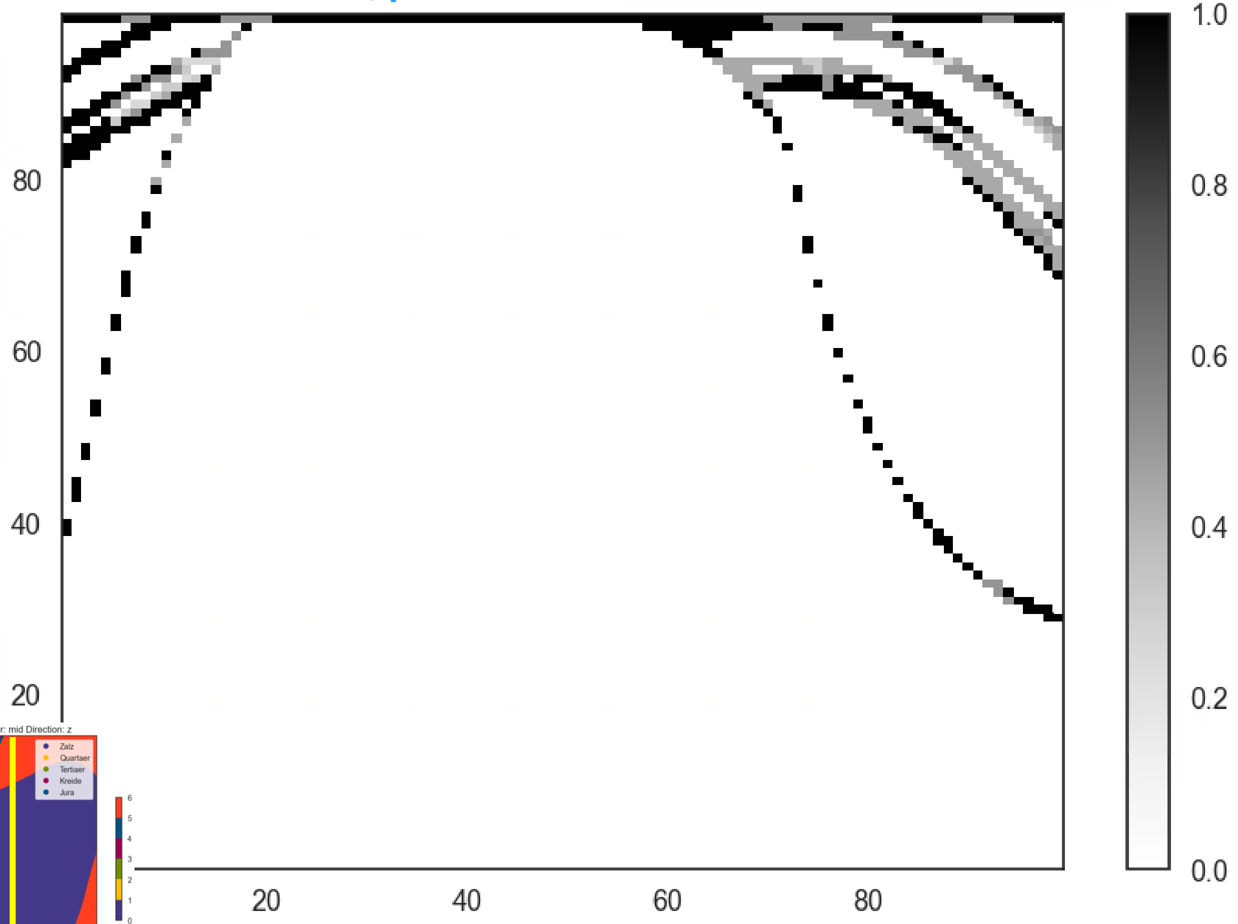
Section



Parameter	Value
Kernel	RBF
Length scale	1
Length scale bounds	(-100, 100)
Variance	1
Training dataset	Random samples from the section, 10%

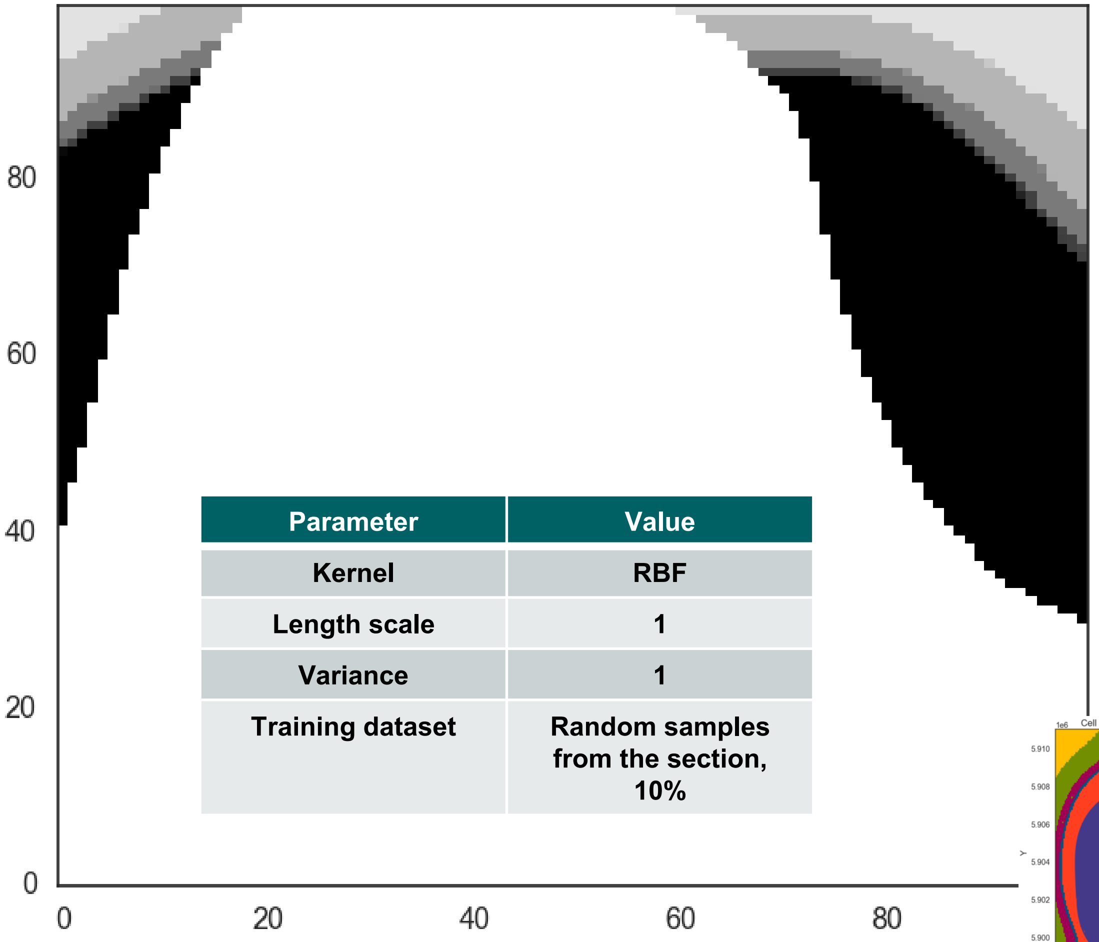


Gaussian Process, probabilities for the formation interfaces

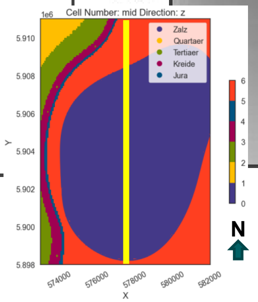


Uncertainty Analysis – Gaussian process classifier

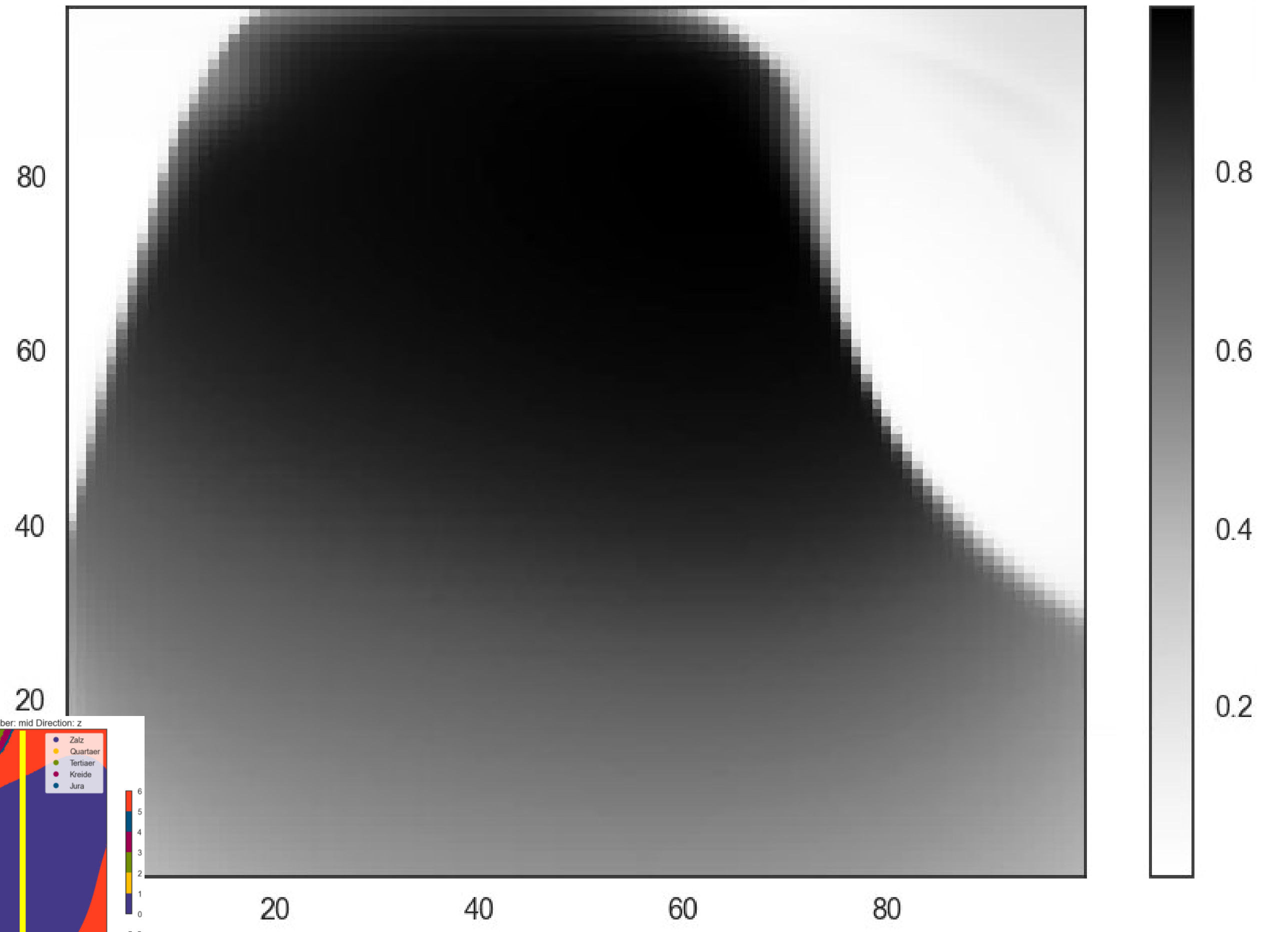
Section



Parameter	Value
Kernel	RBF
Length scale	1
Variance	1
Training dataset	Random samples from the section, 10%



Probability for Salt



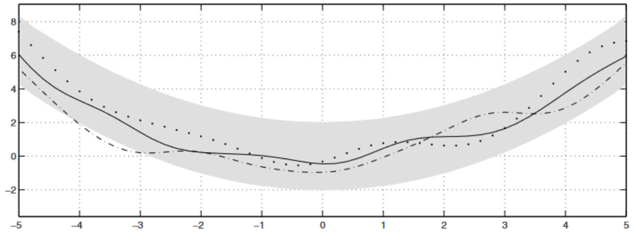
Uncertainty Analysis - Gaussian Process

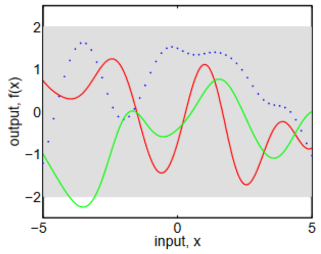
- Three dimensional model

$$f(x) \sim GP\,(m(x), k(x, x'))$$

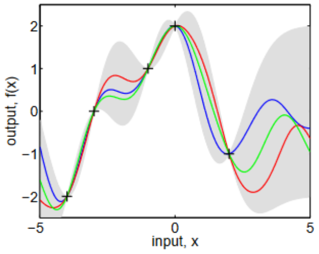
$$m(x) = \mathbb{E}\,[f(x)]$$

$$k(x, x') = \mathbb{E}\,[(f(x) - m(x))(f(x') - m(x'))]$$



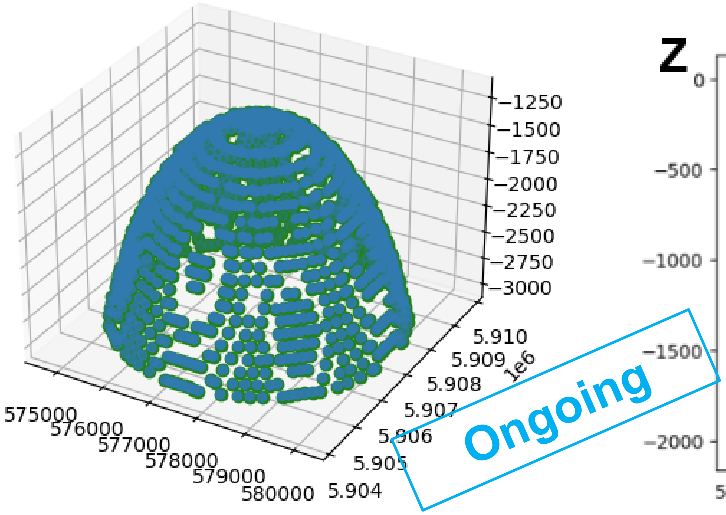


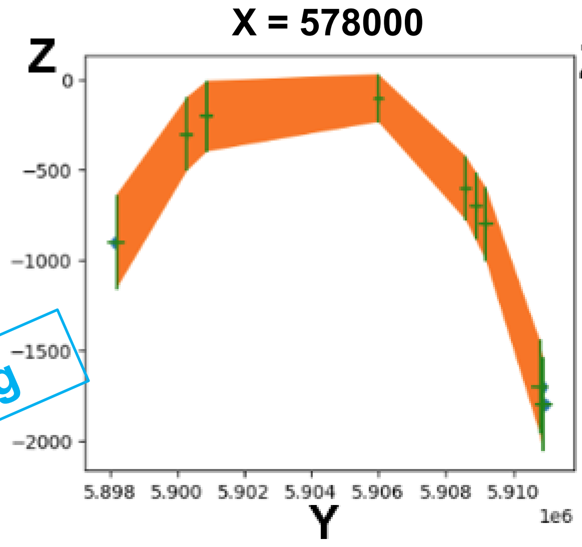
(a), prior

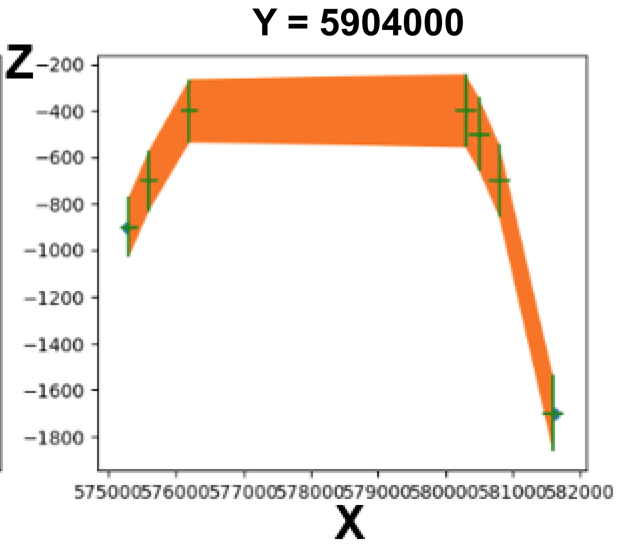


(b), posterior

Rasmussen 2004; Rasmussen and Williams 2006

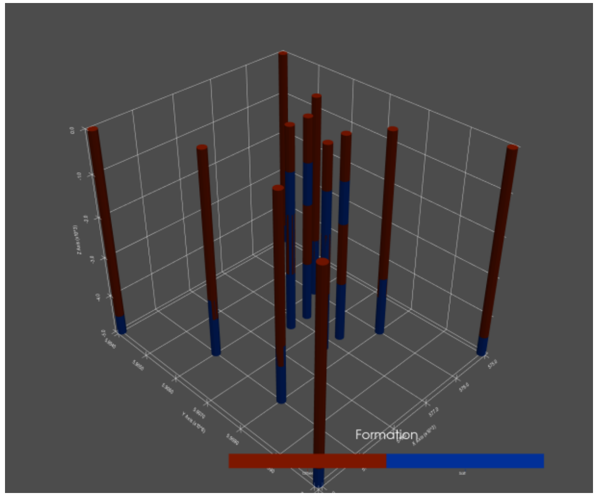


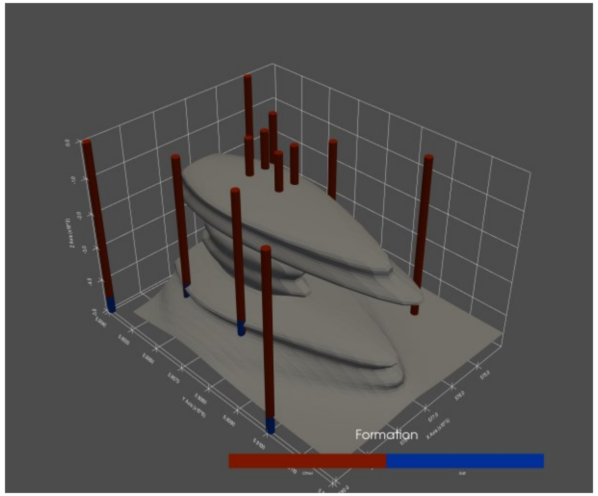




Uncertainty Analysis - Variational Gaussian Process with GeoML

- Synthetic salt dome





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

$$p(z_* | \mathbf{z}) = N(z_*; m_*, \sigma_*^2)$$

$$m_* = \mathbf{k}_{*\mathbf{f}}(\mathbf{K}_{\mathbf{ff}} + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{z}$$

$$\sigma_*^2 = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_{*f}(\mathbf{K}_{\mathbf{ff}} + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{k}_{\mathbf{f}*} + \sigma^2$$

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

$$p(\mathbf{f}, \mathbf{u}) = N_{N+U}\left(\begin{bmatrix}\mathbf{f} \\ \mathbf{u}\end{bmatrix}; \mathbf{0}, \begin{bmatrix}\mathbf{K}_{\mathbf{ff}} & \mathbf{Kfu} \\ \mathbf{K}_{\mathbf{uf}} & \mathbf{Kuu}\end{bmatrix}\right)$$

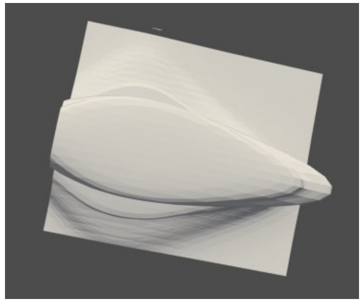
$$p(\mathbf{z}, \mathbf{f}, \mathbf{u}) \propto p(\mathbf{z}|\mathbf{f}, \mathbf{u})p(\mathbf{f}, \mathbf{u}) = p(\mathbf{z}|\mathbf{f})p(\mathbf{f}|\mathbf{u})p(\mathbf{u})$$

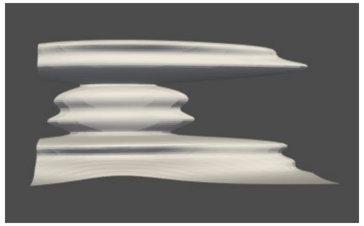
to maximize $p(\mathbf{z}), p(\mathbf{f}|\mathbf{u})$ is replaced by $q(\mathbf{f}|\mathbf{u})$

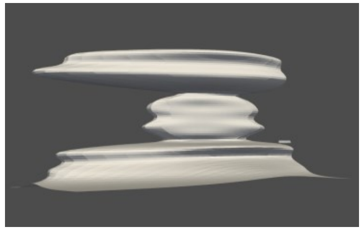
Variational inference approximates the posterior as a Gaussian $\boldsymbol{q}(\boldsymbol{f}, \boldsymbol{u}) = \boldsymbol{p}(\boldsymbol{f}|\boldsymbol{u})\boldsymbol{q}(\boldsymbol{u})$
by introducing the variational distribution
$\boldsymbol{q}(\boldsymbol{u}) = N_U(\boldsymbol{u};\ \boldsymbol{m},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

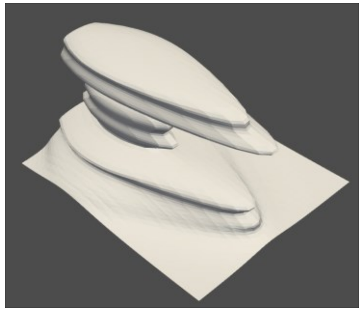
$$\log p(\mathbf{z}) \geq \mathbb{E}_{q(\mathbf{f})}[log\ p(\mathbf{z}|\mathbf{f})] - \mathbf{KL}\ [q(\boldsymbol{u})\ ||\ p(\boldsymbol{u})] = \mathbf{ELBO}$$

Gonçalves et al. 2022



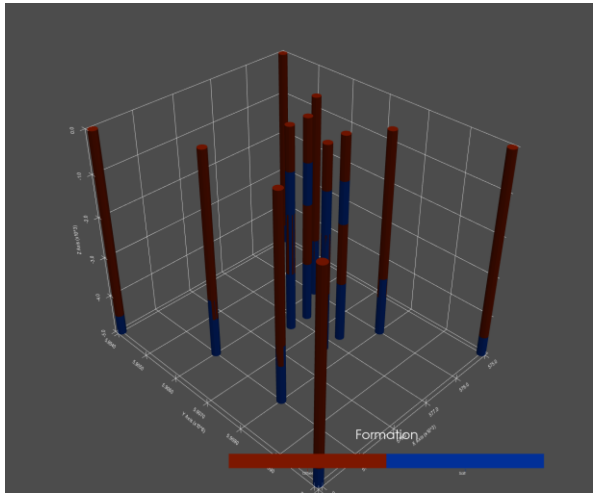


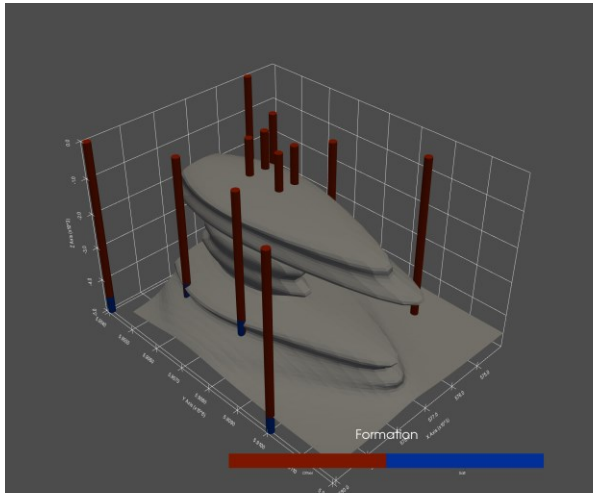


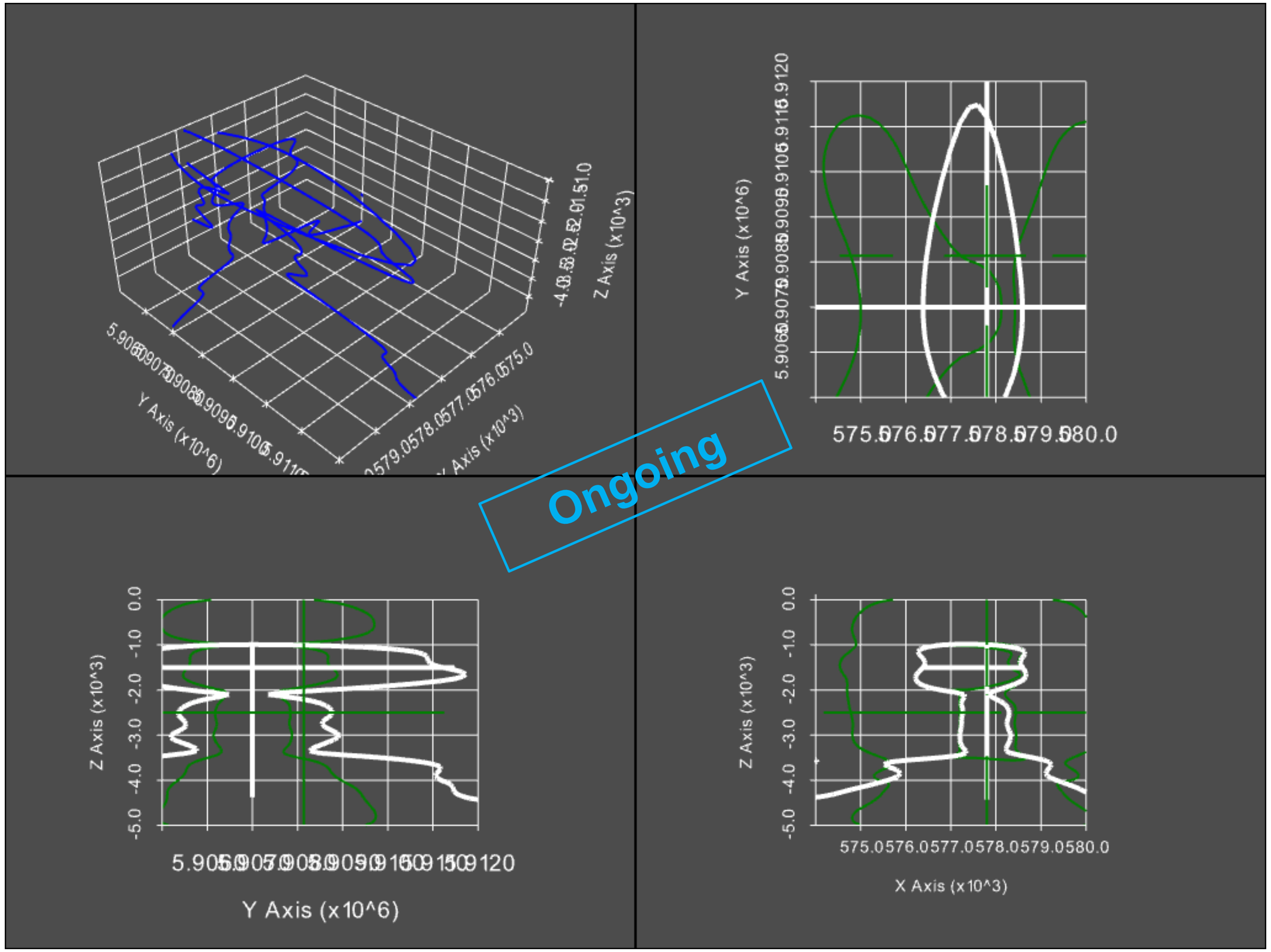


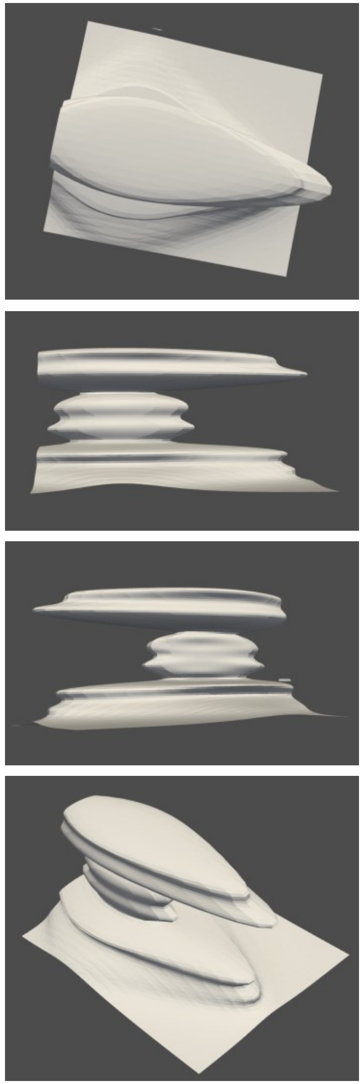
Uncertainty Analysis - Variational Gaussian Process with GeoML

- Uncertainty analysis









Discussion

Discussion

- The objective of this work is create or contribute on a decision analysis framework, where each type of uncertainty analysis conditions the outcome.

- Review of the influence of the covariance.

- Incorporation of seismic data.

- <u>Tentative</u>: investigate the influence of the velocity model on the UQ approach (e.g., use probabilistic approach for the velocity model and the corresponding spatial/geometrical shape it builds).

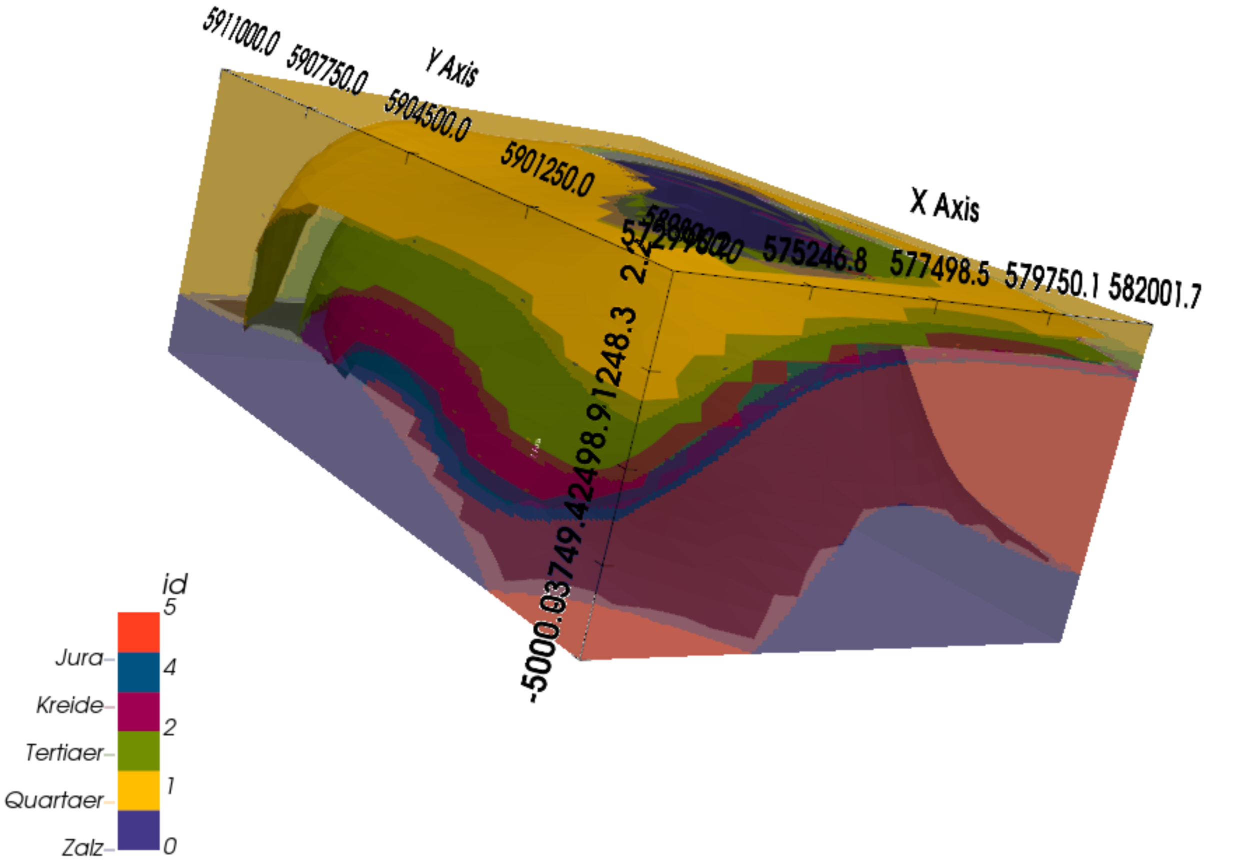
Next Steps

Next Steps

- Review of the influence of the covariance.
- Incorporation of seismic data, initially as horizons.
- Incorporation of the uncertainties given by the velocity model used for depth conversion.

GemPy v3 Released

- Python-based, open-source library for 3D geological modelling.
- Optimized, rigorous testing.



- GitHub
  https://github.com/gempy-project

- Tutorials
  https://gempy-project.github.io/temp_gp3_docs/tutorials/index.html

References

Colombo C. and Wellmann F., “Implicit Modelling with GemPy and Spatial Uncertainty Reduction of Potential Nuclear Waste Storage Sites”, International Association for Mathematical Geosciences (IAMG), 2023.

De la Varga M. and Wellmann J. F. (2016). *Structural geologic modeling as an inference problem: A Bayesian perspective*. Interpretation 4(3):SM1-SM16. https://doi.org/10.1190/INT-2015-0188.1

De la Varga M., Schaaf A., Wellmann F. (2019). GemPy 1.0: open-source stochastic geological modeling and inversion. Geoscientific Model Development, 12, 1–32, 2019. https://doi.org/10.5194/gmd-12-1-2019

Gonçalves, Ítalo Gomes, Felipe Guadagnin, and Diogo Peixoto Cordova. ‘Learning Spatial Patterns with Variational Gaussian Processes: Regression’. Computers & Geosciences 161 (April 2022): 105056. https://doi.org/10.1016/j.cageo.2022.105056.

GemPy Project, https://github.com/gempy-Project.

GemPy Project, Tutorials, https://gempy-project.github.io/temp_gp3_docs/tutorials/index.html.

Köhn D., De Nil D., al Hagrey S. A. and Rabbel W., ‘A Combination of Waveform Inversion and Reverse-Time Modelling for Microseismic Event Characterization in Complex Salt Structures’, Environmental Earth Sciences 75, no. 18 (September 2016): 1235, https://doi.org/10.1007/s12665-016-6032-4.

Probabilistic Programing (November 15th, 2020). Link: https://probabilistic-programming.org/ . Accessed on January 12th, 2023.

Probtorch. Link: https://github.com/probtorch/probtorch .

PyMC, https://www.pymc.io/welcome.html. © Copyright 2022, PyMC Team. Accessed on January 12th, 2023.

PyMC, probabilistic programming library. Link: https://www.pymc.io/welcome.html

Pyro. Link: https://pyro.ai/.

Python. Link: https://www.python.org/ .

References

Rasmussen, C. E., 'Gaussian Processes in Machine Learning', in Advanced Lectures on Machine Learning, ed. Olivier Bousquet, Ulrike von Luxburg, and Gunnar Rätsch, vol. 3176, Lecture Notes in Computer Science (Berlin, Heidelberg: Springer Berlin Heidelberg, 2004), 63–71, https://doi.org/10.1007/978-3-540-28650-9_4.

Rasmussen, C. E. and Williams, C. K. I., Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning (Cambridge, Mass: MIT Press, 2006).

Von Berlepsch T. and Haverkamp B., 'Salt as a Host Rock for the Geological Repository for Nuclear Waste', Elements 12, no. 4 (August 2016): 257–62, https://doi.org/10.2113/gselements.12.4.257.

Wellmann F. (n.d.). *Modeling geological interfaces using implicit surface representations with gempy. Why, when and how?* Computational Geoscience and Reservoir Engineering, RWTH Aachen University, Germany. Presentation.

Wellmann F. and Caumon G. (2018). *3-D Structural geological models: Concepts, methods, and uncertainties*. Advances in Geophysics. Elsevier. DOI: 10.1016/bs.agph.2018.09.001

Wellmann, F. and Liang, Z., Structural Geological Models, WS 2022/23 (Lecture).

Witter J. B., Trainor-Guitton W. J., and Siler D. L., 'Uncertainty and Risk Evaluation during the Exploration Stage of Geothermal Development: A Review', Geothermics 78 (March 2019): 233–42, https://doi.org/10.1016/j.geothermics.2018.12.011.

Languages, packages and libraries:

- GemPy: https://gempy.org/, https://docs.gempy.org; https://github.com/gempy-project
- Matplotlib: https://matplotlib.org/
- NumPy: https://numpy.org/)
- PyMC3: https://www.pymc.io/projects/docs/en/v3/index.html
- Python: https://www.python.org/

Thank you for your attention

Questions?