IAMG2023

The 22nd annual conference of the IAMG

August 05 - 12, 2023, Trondheim, Norway

co-organized by:

Short Abstracts

IAMG2023

Trondheim, Norway, August 05 - 11, 2023,
The 22th Annual Conference of the International
Association for Mathematical Geosciences

IAMG2023 sponsored by:

Preface

Dear Friends and Colleagues,

The local organizers of the 22nd annual IAMG conference take great and sincere pleasure in welcoming you to Trondheim for IAMG2023 and introduce you to the conference's book of abstracts. We are glad to announce more than 150 oral presentations including awardees and keynote speakers, almost 30 posters and more than 200 participants. This confirms that our association is alive, healthy, and vibrant. We are looking forward to days of networking, tough and rewarding discussions, and social events. Good to be back. Almost 25 years since last time (IAMG99). Trondheim is Norway's mineral-, knowledge-, and technology capital. It is housing the only mining engineering education on master and PhD-level in Norway, the Mining Directorate (DirMin), and the Geological Survey of Norway (NGU). Both historical and active mines are just on the city's doorstep. This includes dormant mining towns like Løkken with more than 300 years long historical roots and the active Verdalskalk-operation just a couple of hours drive north of Trondheim. The city is also the home of Norway's more dominating energy companies like Equinor and AkerBP as well as engineering geology companies like Multiconsult, COWI and Norconsult all of which are working on topics relevant for IAMG. IAMG2023 is organized as a hybrid event with two-way online communication and discussions. It brings practitioners side by side with researchers and promotes interaction and dialog between representatives of various disciplines of mathematical geology. The program has been developed to create synergy between theory and practice. New sessions for our community like Real Options in Geoscience will be organized to further increase the association's ability to exploit its capabilities and capacities. The local committee has focused on developing sessions with varying focus to highlight the cross-boundary applications of mathematical geosciences. Sessions are covering both petroleum, mining, and engineering geology applications. The host of this year's conference, The Norwegian University of Science and Technology (NTNU) has a total of 40.000 students and has historical roots that dates to 1910. The first lecture given at NTNU was in mathematics, and NTNU has long traditions in geomodelling, geostatistics and prospect play analysis. The mining education at NTNU was based on activities in silver mines near the town of Kongsberg and was moved to Trondheim in 1912. The education in petroleum geology and -technology sprung out of the mining education in the 1970s to follow up the Ekofisk discovery in the North Sea. IAMG2023 is the result of a major team effort. We would like to thank all members of the local organising committee for their work, and commitment. We are also grateful to the scientific committee and the officers of IAMG. Would also like to direct our sincere thanks to sponsors; AkerBP ASA, Sibelco Nordic ASA and Rana Gruber AS. Finally, we would like to thank all those who have submitted contributions, the session conveners and to the keynote speakers and awardees that have agreed to enlighten us during the plenaries. Without your contributions – no conference.

Hope to see all of you again at IGC in South-Korea in 2024!

Steinar Ellefmo and Jo Eidsvik
Conference Chair and Conference co-chair

Contents

S00 Plenary Session

S0001. Spatial Analytics for Geoforensics and Geohealth

Jennifer Mckinley (Queen's University Belfast)
Room: R7 2023-08-07 08:40

Human-environmental relations are explored through case study examples involving ground-based geological and geochemical data combined with remotely sensed data to explore the importance of spatial thinking in forensic geoscience. Jennifer will also present recent research which seeks to gain a greater understanding of the link between human health and our environment, the potential impacts of air and traffic pollution and the value of green space and nature-based solutions. The approaches presented acknowledge the compositional nature of the environmental data such as geochemistry data and will include case studies involving the use of compositional data analysis, including the use of log ratios and balances.

S0002. Geochronology as a compositional data problem

Pieter Vermeesch
Room: R7 2023-08-07 13:00

Geology is a historical science, in which timing is of the utmost importance. The geologic time scale is calibrated using isotopic ’clocks’ that are based on the radioactive decay of ’parent’ isotopes (e.g., U-238, Rb-87, K-40) to radiogenic ’daughter’ isotopes (e.g., Pb-206, Sr-87, Ar-40) in the presence of non-radiogenic ’sister’ isotopes (e.g., Pb-204, Sr-86, Ar-36). Together, these three types of isotopes form a compositional data space. A huge edifice of sophisticated mathematical-statistical tools has been constructed to process geochronological data. Unfortunately, these tools ignore the compositional nature of isotopic measurements. Ratio correlation and the Aitchison geometry impose a covariance structure on isotopic data that permeates throughout geochronology. Capturing this covariance structure requires a complete re-evaluation of the isotopic data processing chain, beginning with the acquisition of raw mass spectrometer data. The re-evaluation of isotopic data processing is tedious and branches out into many directions. However, it is an effort worth making because it greatly improves the precision and accuracy of geochronology, which benefits all the disciplines that depend on it.

S0003. Data Analytics and Machine Learning for Geostatistics

Michael Pyrcz (University of Texas at Austin)
Room: R7 2023-08-08 08:30

The subsurface resource industry has a long history of working with large, complicated geoscience and engineering datasets. Our community has been working with 'big data' for decades, driving the development and establishment of the geostatistical toolbox. As a result, we are uniquely situated in this current digital revolution, as we are further down the road of data-driven workflow development and have insights and experiences to share with the other scientific and engineering domains. The foundation of geostatistics, domain expertise, spatial continuity, scale and uncertainty remain essential for machine learning (even advanced deep learning architectures); therefore, we must lead in the advancement of new data analytics and machine learning technologies for subsurface applications. To reflect on this opportunity, I elucidate connections between geostatistics, data analytics and machine learning to frame up a vision for the future of geostatistics.

S0004. Decision-making for follow-up mineral exploration based on spatial uncertainty of geochemical data

Behnam Sadeghi (CSIRO)
Room: R7 2023-08-08 13:00

In mineral prospectivity mapping based on geochemical data, generating interpolated maps is necessary due to limitations such as restricted access or budget constraints that prevent the collection of a higher number of samples. This can result in errors due to differences between the actual concentrations in sampled areas and the estimated concentrations in unsampled areas. Behnam will discuss his current research on quantifying and modeling this spatial uncertainty to facilitate better decision-making for follow-up sampling and mineral exploration. The decision-making model is based on both frequency and Bayesian frameworks, allowing it to identify target areas with the highest concentration frequencies and the greatest probability of mineralization occurrence. This model is useful in a variety of scenarios, including those with a limited number of samples, restricted access to certain areas due to distance, morphology, or budget constraints, and even in the search for mineral potentials on other planets such as Mars. Behnam is currently developing this project in collaboration with CSIRO, Stanford University, and the Carnegie Institution for Science, the Earth and Planets Laboratory (EPL).

S0005. Modeling of coupled processes in fractured geothermal systems

Inga Berre (University of Bergen)
Room: R7 2023-08-09 08:30

Geothermal systems are typically characterized by fractured structures which impacts coupled thermo-hydro-mechanical processes. In this talk, we consider mathematical and numerical models for such systems with explicit representation of fractures. We discuss how the processes are strongly dominated by fractures in the formation, and how fractures can reactivate, deform, and propagate because of the thermo-hydro-mechanical processes. Examples of modeling of fracture deformation and propagation are shown for different scenarios. The first consider how fractures may deform as a response to reservoir fluid injection. This done in development of geothermal resources to actively stimulate fractures to open to enhance reservoir permeability, or during the production phase to re-inject produced geothermal brine. The second shows how natural convection in the systems can give rise to deep migration of fractures because of convective cooling and resulting thermal contraction of the rock, thereby contributing to transport of heat from the deep crust into the geothermal reservoir.

S0006. Harnessing the complexity of minerals: Data-driven exploration of evolving Earth and planetary systems

Shaunna Morrison (Earth and Planets Laboratory, Carnegie Institution for Science)
Room: R7 2023-08-09 13:00

The key to answering many compelling and complex questions in Earth, planetary, and life science lies in breaking down the barriers between scientific fields and harnessing the integrated, multi-disciplinary power of their respective data resources. We have a unique opportunity to integrate large and rapidly expanding data resources, to enlist powerful analytical and visualization methods, and to answer multi-disciplinary questions that cannot be addressed by one field alone.

Rapidly expanding mineral data resources have created an opportunity to characterize changes in near-surface mineralogy through deep time and to relate these findings to the geologic and biologic evolution of our planet over the past 4.5 billion years. Data-driven studies employing advanced analytical and visualization techniques such as mineral ecology, network analysis, and association analysis, allow us to begin tackling big questions in Earth, planetary, and biosciences, including those related to

(1) the relationships of mineral formation and preservation with large-scale geologic processes, such supercontinent assembly, the oxidation of Earth's atmosphere, and changes in ocean chemistry.

(2) the abundance and likely species of as-yet undiscovered mineral, as well as the probability of finding a mineral or mineral assemblage at any locality on Earth or another planetary body.

(3) exploring the possibility that Earth's mineral diversity and distribution is a biosignature.

(4) characterizing the origins of all mineral species through the development of the Evolutionary System of Mineralogy – a system that will provide a framework for predicting the formational conditions of mineral species of unknown origin.

(5) lastly, integrating across disciplines and exploring ideas that one field alone cannot fully characterize (e.g., how the geochemical makeup of our planet affected the emergence and evolution of life, and, likewise, how life influenced chemical composition and geological processes throughout Earth history).

S0007. Non linear and non Gaussian geostatistical models using Copulas

Andras Bardossy (University of Stuttgart)
Room: R7 2023-08-10 08:30

Spatial dependence of natural parameters can vary strongly due to the complex underlying geological, environmental or meteorological processes responsible for their generation. Copulas offer a flexible framework for the analysis of multivariate data and can be extended to random fields for geostatistical investigations. Copulas define dependence in the rank space and are independent of the marginal distributions of the variables, thus can be considered as descriptors of the spatial dependence only. One of the specific properties of spatial dependence which can be investigated using copulas is asymmetry. Many spatial processes such as advection-diffusion lead to asymmetric dependence – high values have different dependence structures then low ones, and directions of processes can be detected. Appropriate copula descriptor tools – asymmetry functions can be used to detect such structures. Models with changing spatial dependence can be used to mimic this kind of behaviour. The corresponding simulation methods are relatively simple and fast. Theoretical and real-life examples of groundwater quality and rainfall are used to illustrate the methodologies.

S0008. A journey into covariance models for spatial data

Xavier Emery (University of Chile)
Room: R7 2023-08-10 12:40

Covariance (positive semidefinite) functions are an essential tool in spatial statistics, machine learning, mathematics and computer sciences, among others. This work gives an overview on characterizations and properties of matrix-valued covariance models for vector random fields defined in a Euclidean space, a sphere, the product of two Euclidean spaces, or the product of a Euclidean space and a sphere, which can be used to represent variables that are coregionalized in a limited portion of the geographical space or on a large portion of the Earth's surface, and whose variations may also depend on time or on the direction along which the variables are measured. We emphasize spectral representations that play an important role in the definition of valid parametric models and in the simulation of spatial or spatiotemporal random fields. We then provide matrix-valued versions of well-known covariance models, such as the Matérn covariance in the Euclidean space, the Gneiting covariance in space cross time or the exponential covariance on the sphere, and analytical expressions for more general parametric families of matrix-valued covariance functions, many of which involve special functions or depend on basic auxiliary functions that control the spatial or temporal marginal covariances. Of particular interest are the definition of compactly supported models or of nonseparable models in product spaces, and the identification of sufficient validity conditions allowing for more flexible parameterizations than those currently available. An example to the modeling of geotechnical data is presented to illustrate the usefulness of regionalizing these data in a product space to account for the dependence of the variable on the measurement direction.

S01 Reservoir Characterization and Machine Learning: From Pore to Field Scale

Suihong Song (Stanford University),
Yuqi Wu (China University of Petroleum (East China)),
Mingliang Liu (Stanford University),
Senyou An (Imperial College London)

Characterization and modeling of subsurface reservoirs at different scales (i.e., from pore scale to field scale) is of great significance to understand inner geological structures, reveal transport mechanisms of various phenomena, predict physical properties, and forecast fluid flow in the future. In recent years, machine learning has exhibited great potential to improve the characterization and modeling accuracy of subsurface reservoirs. In the session, we welcome innovative machine learning approaches and exciting applications related to reservoir characterization and modeling of various scales. The key topics include, but are not limited to:

(1) Characterization and modeling of reservoirs at the core or field scale based on traditional geostatistics;

(2) Machine learning methods for geo-modeling of macro-scale reservoirs and reconstruction of micro-scale digital cores;

(3) Machine learning methods for the upscaling of reservoirs from pore to field scale;

(4) Machine learning methods related to the digital rock images (e.g., denoising, segmentation, and enhancement);

(5) Data-driven or physics-informed methods for geophysical/flow simulation and inverse problems;

(6) Application of the above methods for various fields (e.g., petroleum, mine, groundwater, CO2 storage).

S0101. Application of Reinforcement Learning in Geostatistical Modeling Workflows

Baran C Yucel (Pennsylvania State University),
Sanjay Srinivasan (Pennsylvania State University),
Morteza E Naraghi (Pennsylvania State University)
Room: R9 2023-08-08 10:00

Numerous geostatistical methods have been proposed to model spatial and temporal variations in earth science phenomena. There are several algorithms for modeling complex spatial heterogeneity, data assimilation, and uncertainty quantification. The inference and modeling of semi-variogram(s) are crucial for implementing many of these algorithms. Robust variogram inference is crucial for understanding the spatial correlation of structures, aiding in subsurface modeling and characterization. Fitting a reliable variogram model is crucial for robust simulation of the spatial distribution of various geologic properties like porosity, permeability, and saturation. Even though variogram inference and modeling are crucial for assessing the spatial structure of the geologic attributes, the inference and modeling process is mostly subjective. It depends strongly on the interpretations based on sparse data. To render the variogram inference and modeling process more objective and somewhat automated, this paper proposes using reinforcement learning methods for improving spatial predictions while also gaining a better understanding of underlying spatial processes that may have given rise to heterogeneities.

Reinforcement learning can facilitate the identification of legitimate variogram parameters for various nested structures and enable the learning agent to determine an optimal policy for variogram modeling by iteratively adjusting variogram parameters and observing the resulting estimations at sampled locations, which may yield different rewards based on discrepancies between estimated and actual sampled outcomes. The primary advantage of using reinforcement learning lies in its capacity to develop a policy for maximizing the cumulative reward for optimal variogram parameters. A suitable reinforcement learning algorithm must be selected and utilized appropriately for this. This paper applies the Q-learning algorithm since it is model-free and quite flexible. Moreover, the Q-learning algorithm brings more benefits in being computationally efficient and easily implemented with conditions where state-action spaces are well defined. The Q-learning algorithm is applied to find optimal parameter values for different anisotropic cases. In addition to determining an optimal automated variogram inference modeling framework, the Q-learning framework also allows us to determine an optimal decision-making policy for manual variogram inference. The expert will also explore if injecting expert knowledge about geological processes into the variogram inference and modeling process is possible.

S0102. Graph Generative Deep Learning for Uncertainty Quantification of Brugge field

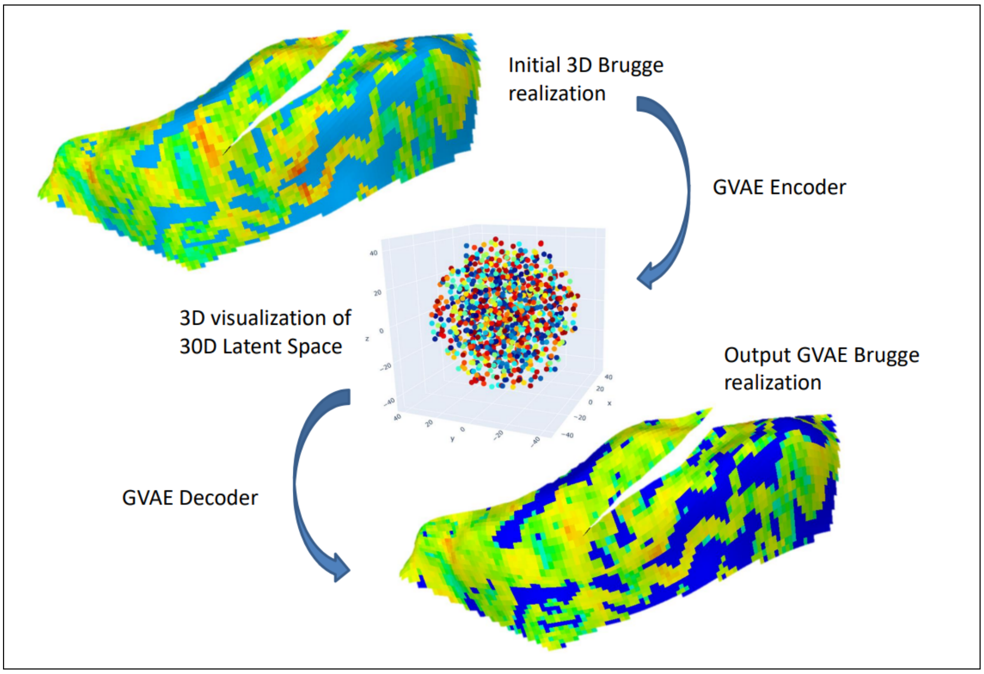
Gleb Shishaev (Heriot-Watt University),
Vasily Demyanov (Heriot-Watt University),
Daniel Arnold (Heriot-Watt University),
Roman Vygon
Room: R9 2023-08-08 10:20

This abstract is an extension of an approach presented in IAMG 2022 conference presenting generative capabilities of Graph Variational Autoencoders (GVAE) to model geological uncertainty of a benchmark Brugge field. Brugge field constitutes an assemblage of realistic depositional environments - fluvial and shallow marine – that realistically represents geological variability and posts a challenge to evaluate uncertainty of spatial property distributions. The work demonstrates the parameterization process of different depositional environment zones of Brugge field through the Latent Space of decreased dimensionality. The novel approach demonstrates a way to build 3D geological property models with respect to the associated and uncertainty.

The key concept of the work is the Geometric deep learning approach that generalizes lattice-based deep neural nets to non-regular domains, such as graphs, and performs convolutions on connected topological graph data. Graph modelling approach provides more flexibility to handle irregularities of geological structural grids representing the topology and connectivity of geological structures. Every node of a graph geological model represents a vector of reservoir properties (e.g., porosity, permeability), while the connections between the nodes are adjusted according to the structure (e.g., faults, unconformities, etc.).

In this work, we describe the GVAE training process and generative capabilities with Brugge benchmark reservoir case study. Brugge is a synthetic oil field which comprises four depositional layers: fluvial, lower shoreface, upper shoreface, sandy shelf). An ensemble of Brugge realizations represent the distribution of prior geological uncertainty across multiple model parameterizations and geological assumptions. We demonstrate how GVAE is able to generate an ensemble of conditional realizations that capture prior geological uncertainty and remains in good agreement with initially allocated geological concepts and statistical parameters.

We conclude that trained GVAE can reliably generate new geomodels through latent space representation of complex geological objects given initial uncertainty. This result paves the way to solve an inverse problem of matching the geological model to dynamic data by perturbing the generate GVAE ensemble in the latent space.

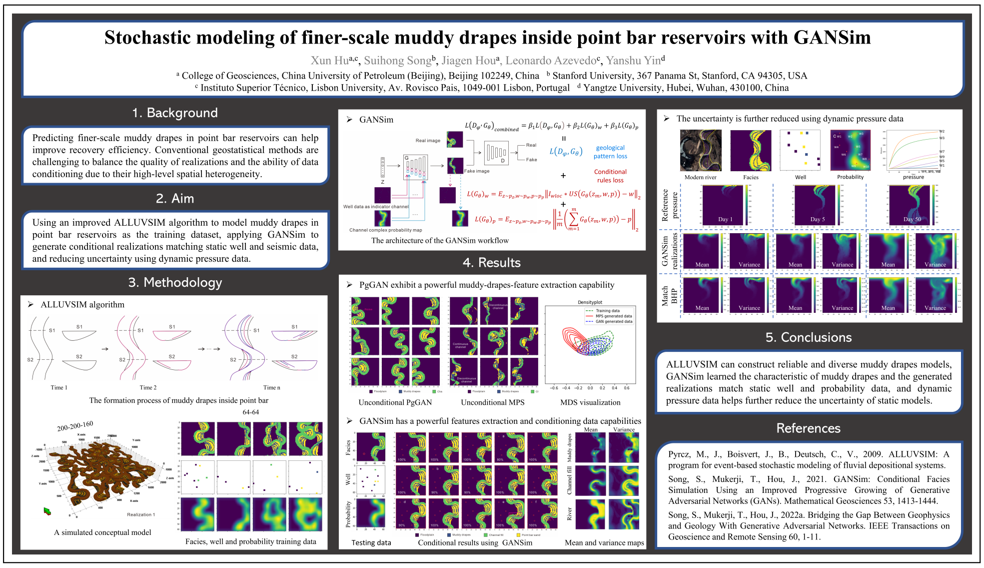


Initial 3D Brugge realization
GVAE Encoder
3D visualization of 30D Latent Space
Output GVAE Brugge realization
GVAE Decoder

S0103. Stochastic modeling of finer-scale muddy drapes inside point bar reservoirs with GANSim

Xun Hu (China University of Petroleum (Beijing)),
Suihong Song (Stanford University),
Jiagen Hou (China University of Petroleum (Beijing)),
Leonardo Azevedo (Instituto Superior Técnico, Lisboa),
Yanshu Yin (Yangtze University)
Room: R9 2023-08-08 10:40

Modeling finer-scale muddy drapes inside point bars using conventional deterministic or stochastic geostatistical (e.g., object-based, event-based, and pixeled-based) methods to improve the recovery efficiency in point bar reservoirs is challenging due to their high-level spatial heterogeneity and millimeter-thick characteristics. Nowadays, Generative Adversarial Networks (GANs) methods have been successfully applied to reproduce several geological facies modeling scenarios (e.g., braided river and carbonate reservoirs). However, few people apply GANs to modeling small-scale muddy drapes and evaluate their modeling capabilities. Thus, this study proposes using an improved ALLUVSIM algorithm (a process-based method) to model muddy drapes in point bar reservoirs as the training dataset, and then applying GANSim (an advanced state-of-art conditional GANs method proposed by Song et al., 2021, 2022) to generate conditional realizations matching static well and probability data, and further reduce uncertainty using dynamic pressure data. The results show that the improved ALLUVSIM can construct reliable and diverse muddy drapes models with specific sinuosity. In synthetic data tests, the progressive growing of Generative Adversarial Networks (PgGAN) captures both large-scale features of channels and small-scale curvilinear features of muddy drapes well, which was challenging for conventional MPS methods. Note that the local noise in PgGAN results may not significantly hurt the global spatial pattern reproduction. Similarly, GANSim (conditional PgGAN) exhibits the powerful muddy-drapes-feature extraction, computational, and conditioning of the different types of information capabilities in producing multiple realizations. The conditional trained generator is also applied to a modern meandering river application with a pressure measurement experiment. The results show that the generator has broad applicability, and the uncertainty of the realizations has been further reduced under the constraints of dynamic pressure data. In the future, the muddy drapes generator trained using Improved ALLUVSIM and GANSim is expected to promote fine reservoir characterization of meandering rivers to guide the development of the remaining oil.


Stochastic modeling of finer-scale muddy drapes inside point bar reservoirs with GANSim
Xun Hu[a,c], Suihong Song[b], Jiagen Hou[a], Leonardo Azevedo[c], Yanshu Yin[d]
[a] College of Geosciences, China University of Petroleum (Beijing), Beijing 102249, China [b] Stanford University, 367 Panama St, Stanford, CA 94305, USA
[c] Instituto Superior Técnico, Lisbon University, Av. Rovisco Pais, 1049-001 Lisbon, Portugal [d] Yangtze University, Hubei, Wuhan, 430100, China
1. Background
Predicting finer-scale muddy drapes in point bar reservoirs can help improve recovery efficiency. Conventional geostatistical methods are challenging to balance the quality of realizations and the ability of data conditioning due to their high-level spatial heterogeneity.
2. Aim
Using an improved ALLUVSIM algorithm to model muddy drapes in point bar reservoirs as the training dataset, applying GANSim to generate conditional realizations matching static well and seismic data, and reducing uncertainty using dynamic pressure data.
3. Methodology
ALLUVSIM algorithm
Time 1
Time 2
Time n
The formation process of muddy drapes inside point bar
200-200-160
64-64
A simulated conceptual model
Facies, well and probability training data
GANSim
Real image
Fake image
Real
Fake
geological pattern loss
Conditional rules loss
Well data as indicator channel
Channel complex probability map
The architecture of the GANSim workflow
4. Results
PgGAN exhibit a powerful muddy-drapes-feature extraction capability
Densityplot
Training data
MPS generated data
GAN generated data
Unconditional PgGAN
Unconditional MPS
MDS visualization
GANSim has a powerful features extraction and conditioning data capabilities
Testing data
Conditional results using GANSim
Mean and variance maps
The uncertainty is further reduced using dynamic pressure data
Modern river
Facies
Well
Probability
pressure
Reference pressure
Day 1
Day 5
Day 50
GANSim realizations
Match BHP
Mean
Variance
5. Conclusions
ALLUVSIM can construct reliable and diverse muddy drapes models, GANSim learned the characteristic of muddy drapes and the generated realizations match static well and probability data, and dynamic pressure data helps further reduce the uncertainty of static models.
References
Pyrcz, M., J., Boisvert, J., B., Deutsch, C., V., 2009. ALLUVSIM: A program for event-based stochastic modeling of fluvial depositional systems.
Song, S., Mukerji, T., Hou, J., 2021. GANSim: Conditional Facies Simulation Using an Improved Progressive Growing of Generative Adversarial Networks (GANs). Mathematical Geosciences 53, 1413-1444.
Song, S., Mukerji, T., Hou, J., 2022a. Bridging the Gap Between Geophysics and Geology With Generative Adversarial Networks. IEEE Transactions on Geoscience and Remote Sensing 60, 1-11.

S0104. Accurate reconstruction of reservoir models in spectral domain using compressed sensing guided by reinforcement learning

Corey Hoydic (Pennsylvania State University),
Sanjay Srinivasan (Pennsylvania State University),
Morteza Naraghi (Pennsylvania State University)
Room: R9 2023-08-08 11:00

To make accurate predictions of the subsurface response to various stimuli, it is imperative that accurate models of spatial variations in subsurface properties are available. Since geologic structures and formations exhibit natural order that manifests in the form of spatial patterns at scales ranging from the pores to the basin, it is necessary that geologic models accurately depict the spatial variability at multiple scales, ensuring consistency of predicted responses with actual field observations. However, the data available to develop reservoir models is sparse in most cases. This research addresses the issue of accurate reproduction of spatial patterns of variability when sparse data is available.

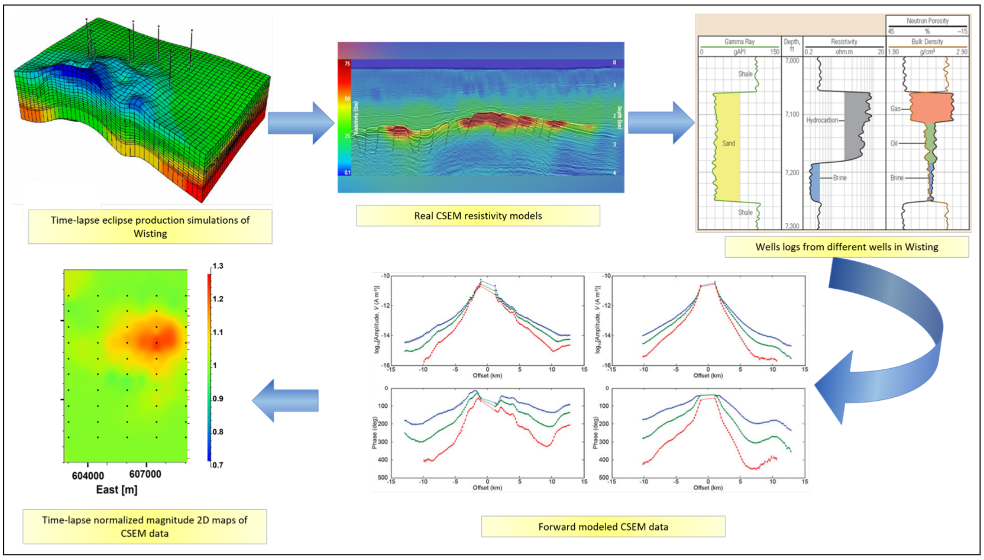
The inference of spatial patterns of variability and subsequent construction of accurate models of spatial variability can be performed in the spectral (frequency) domain. When sparse data is available, the construction of reservoir model can be solved as a compressed sensing problem. Efficient and accurate reconstruction of models is possible by retaining only the non-zero Fourier coefficients. This implies definition of optimal subspaces (of nonzero Fourier coefficients) to reconstruct the models. In this research, a Q-learning reinforcement learning (RL) framework is implemented for optimizing the search space in the frequency domain. Fundamentally, the proposed approach explores the characteristics of the optimal frequency-domain representation of a reservoir model based on sparse data and relates that to the essential features exhibited by geology. The definition of an optimal search space for model construction through the application of a reinforcement learning algorithm is a unique contribution made by this research.

The ultimate goal of this research is to extract quantitative information from available sparse data and geologic analog models such that the appropriate information can be transferred to a target subsurface model. The application of the proposed RL framework will also allow us to compute a suite of spatial and spectral statistics directly from sparse data that can also act as a basis by which we can compare different geologic analogs. This will enable us to develop a method for screening a database of geologic analogues and cataloguing them on the basis of their statistical characteristics.

S0105. Time-lapse 3D CSEM for reservoir monitoring based on rock physics simulation

Mohammed Ettayebi (NTNU),
Shunguo SW Wang (Norwegian University of Science and Technology),
Martin ML Landrø (Norwegian University of Science and Technology)
Room: R9 2023-08-08 11:20

Marine controlled-source electromagnetic (CSEM) method measures EM fields transmitted by an active source either at seafloor or a few tens of meters above the seafloor, therefore it can image the electrical resistivity beneath the seafloor. Being able to resolve resistivity anomalies of high saturations of petroleum, marine CSEM sounding has gained increasing popularity after its introduction in 1981 by Cox. Nevertheless, the EM research community starts to extend its applications from petroleum exploration to an active production monitoring tool. Our study focuses on the latter. By utilizing various dynamic reservoir properties made available through reservoir simulation of the Wisting field located in the Norwegian part of the Barents Sea, a realistic geoelectric model was created. To this end, we develop geologically consistent rock physics models, such that the available simulation results can be transformed into resistivity maps. We show that the resistivity map pertaining to each time-step can be used as an input model in a Finite Difference Time Domain (FDTD) forward modeling workflow to produce synthetic EM data. This synthetic EM data can be studied and analyzed in light of production induced changes in the reservoir for different production phases. This will allow us to acquire insights towards developing a technically feasible reservoir monitoring workflow suitable for time-lapse CSEM. Our result shows that at different production phases, the CSEM responses are different. Therefore, the method can be effectively used for production monitoring purposes. Moreover, this study will enable the testing of other time-lapse workflows with realistic complexities evaluating the potential of this technology for field application, by investigating the resolution limitations and the repeatability requirements.



S0106. Pore Pressure Uncertainty Characterization Coupling Machine Learning and Geostatistical Modelling

Amilcar Soares (Tecnico U. Lisboa),
Leonardo Azevedo (CERENA, Instituto Superior Tecnico),
Amilcar Soares (CERENA, Instituto Superior Tecnico),
Ruben Nunes (Mind2Earth, Portugal),
Teresa Martins (GALP, Portugal),
Mario Santos (GALP, Portugal),
João Costa (UFRGS),
Paulo Salvadoretti (UFRGS, Brasil)
Room: R9 2023-08-08 11:40

Pore pressure prediction is fundamental when drilling or managing deep and geologically complex targets, for example, for Carbon Capture and Sequestration projects. Even with a considerable number of existing wells, when located in challenging geological environments a poor prediction of abnormal pore pressure might result in high risk technological failures. To better quantify geomechanical risks, the uncertainty associated with the pore pressure prediction should be integrated within the geo-modelling workflow. Leveraging a challenging real case from the Brazilian pre-salt reservoir, the work presented herein proposes a seismic-driven gradient pore pressure modelling workflow based on machine learning regression and stochastic sequential co-simulation. First, existing angle-dependent seismic reflection data are inverted for P- and S-wave velocity and density models. Then, a regression model between pore pressure gradient and P- and S-wave velocity, density and depth is created based on the well-log information. The trained model is applied to the predicted a gradient pore pressure model from the model retrieved from the seismic inversion. This gradient pore pressure model is a smooth representation of the highly variable subsurface and is used as secondary variable in a stochastic sequential co-simulation to generate multiple realizations of gradient pore pressure. The ensemble of predicted models can be used to assess the spatial uncertainty about the predictions. The results of the application example show the ability of the method to reproduce the spatial patterns observed in the seismic and to reproduce existing GPP well-logs at two "blind" well locations.

S0107. Wavelet Operational Matrices and Lagrange Interpolation Differential Quadrature-Based Numerical Algorithms for Simulation of Nanofluid in Porous Channel

Aisha M. Alqahtani (PNU)
Room: R7 2023-08-09 14:00

Wavelet Operational Matrices and Lagrange Interpolation Differential Quadrature-Based Numerical Algorithms for Simulation of Nanofluid in Porous Channel

In this work we analyse the features of nanofluid flow and thermal transmission (NFTT) in a rectangular channel which is asymmetric by developing two numerical algorithms based on scale-2 Haar wavelets (S2HWs), Lagrange's interpolation differential quadrature technique (LIDQT), and quasilinearization process (QP).

In the simulation procedure, first of all, using similarity transformation (ST), the governing unsteady 2D flow model is changed into two highly non-linear ODEs. After that, QP is applied to linearize the non-linear ODEs, and finally (S2HWs) and (LIDQT) are used to simulate the non-linear system of ODEs. In results and discussion section, the parameters Reynolds number (R), expansion ratio and Nusselt (Nu), and nanoparticle volume fraction ( ) are analysed with respect to velocity and temperature profiles. The proposed techniques are easy to implement for fluid and heat transfer (FHT) problems.

S0108. Machine learning-based reservoir characterization at multi-scales: current research status, future directions, and challenges

Tapan Mukerji (Stanford University),
Suihong Song (Stanford University),
Mingliang Liu (Stanford University)
Room: R7 2023-08-09 14:20

Machine learning methods have been deeply integrated with reservoir characterization at multiple scales. Representative topics include: multimodal and multiscale fusion of digital images using GANs for reconstruction of porous rocks at the pore scale, 3D reservoir facies classification from seismic data based on discriminative CNN method, field-scale reservoir geomodelling conditioned to global features (e.g., facies proportion, channel sinuosity), well data, and geophysics-produced facies probability maps with GANs (i.e., GANSim) and its filed applications, flow simulation at various scales with data-driven or physics-informed approaches and surrogate training, inversions, and links between these topics, etc.

Challenges and future research directions include: construction of a large library of real or at least geologically reasonable datasets to create the equivalent of "ImageNet" for geology including 2D/3D digital rock images/cubes, different types of 3D reservoir geomodels etc., improvement of physics-informed approaches and injection of physics into reservoir and digital rock generation processes, with linkage among reservoirs of various scales.

S0110. Facies modelling in fluvial environments with Self-Attention Generative Adversarial Networks conditioned to well data

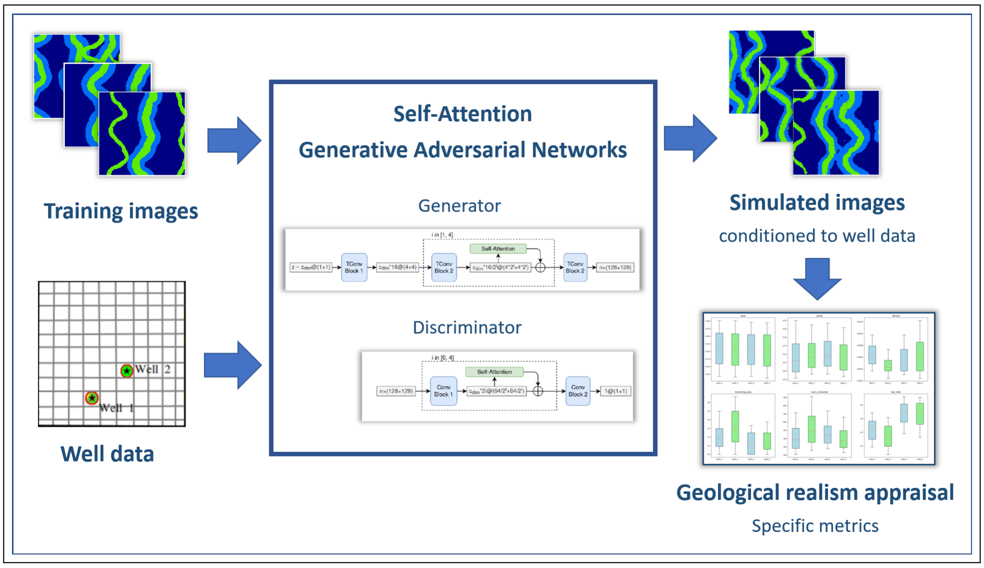
Valentin GOLDITE (Paris-Saclay University),
Antoine BOUZIAT (IFP Energies nouvelles),
Jean-François LECOMTE (IFP Energies nouvelles),
Thibault FANEY (IFP Energies nouvelles)
Room: R7 2023-08-09 14:40

Subsurface reservoir modelling is routinely used in various contexts dealing with the underground environment: energy resources production, subsurface storage, water resources management... Simulating the spatial distribution of geological facies is then a cornerstone piece in most reservoir modelling workflows, as this distribution directly impacts key physical variables such as the porosity, permeability and strength of the underground rocks. However, modelling facies heterogeneities can be particularly challenging for channelized sedimentary environments, and more generally for any environment where the facies distribution is driven by specific genetic units or architectural elements, often called "geobodies". Indeed, classical geostatistical approaches imply a difficult trade-off between conditioning the model to the available hard data and simulating geobodies which are geologically realistic.

In this work, we use an emerging Deep Learning method based on Generative Adversarial Networks (GAN) to facilitate facies modelling in fluvial environments. With this method, multiple realizations of the facies spatial distribution can be simulated in each layer of the reservoir model grid, after training the GAN with a set of reference images. This approach appears as a promising way to reconciliate hard-data conditioning and geological realism, and thus a possibly useful complement to classical geostatistical methods.

More precisely, we introduce a new GAN architecture and a simulation strategy gathering several assets: (1) consideration of long-range dependencies between grid cells thanks to a self-attention mechanism, (2) conditioning to hard data at well locations thanks to a specific loss function, and (3) advanced geological realism assessment thanks to more than 10 tailored metrics based on the shape, size and connectivity of the channels. Each of these elements taken independently has been used in previous studies, but this is the first time to our knowledge they are associated together in the same facies modelling methodology.

Finally, we apply the proposed strategy on two different datasets and document the results, aiming at illustrating its promises for simulating lithological heterogeneities in situations where standard geostatistical approaches are challenged.


Training images
Well 2
Well 1
Well data
Self-Attention
Generative Adversarial Networks
Generator
Discriminator
Simulated images
conditioned to well data
Geological realism appraisal
Specific metrics

S0111. Unsupervised Machine Learning for Well Log Depth Alignment: A Case Study of Well 16/1-9 in the Ivar Asan Field, Norwegian North Sea

Sushil Acharya (NTNU),
Karl Fabian (Norwegian University of Science and Technology)
Room: R7 2023-08-09 15:20

Well logs provide valuable information about underlying geological formations and are used in the oil and gas industry to estimate the potential of hydrocarbon reservoirs and track the progress of drilling operations. But well logs may not be at the same depth reference due to factors such as formation variations, rate of penetration, cable tension, and tool deviation. It is crucial to align them to allow effective interpretation of geological data.

Well logs are often aligned manually by experts in this industry. But, with to advances in machine learning, we can now utilize machine learning methods to automatically align well logs to a common depth reference. In this case study, we explain the options for aligning the logging while drilling (LWD) and Electrical wireline logging (EWL) well log data from the Well 16/1-9 in the Ivar Asan Field, Norwegian North Sea using K means clustering.

The suggested algorithm divides the data into clusters using K-means clustering. The cluster boundary is identified by the algorithm by analyzing rapid changes in cluster allocations along the depth axis, and pairs of neighboring data points with different cluster assignments are evaluated as probable data boundaries.

The technique was applied to raw LWD and EWL log data from well 16/1-9, which included various log suites such as gamma-ray, density, acoustic, neutron, and micro resistivity to predict boundaries. These boundaries may also be considered as bed boundaries or lithological boundaries. Therefore, the predicted data boundary on both logs should refer to the same depth level. After predicting the data boundaries on both LWD and EWL logs, the gamma-ray log from the LWD was aligned to the same depth reference as the reference gamma-ray log from EWL. All the other logs from the LWD log file were then adjusted to the appropriate depth values based on the adjusted depth of the EWL gamma-ray log.

The result demonstrate that the proposed approach properly aligns the LWD and EWL logs. Overall, using unsupervised machine learning methods to align well log depths has the potential to improve objectivity, repeatability, and efficiency in the oil and gas industry.

S0113. Geomodelling with GANSim: from idea to practice

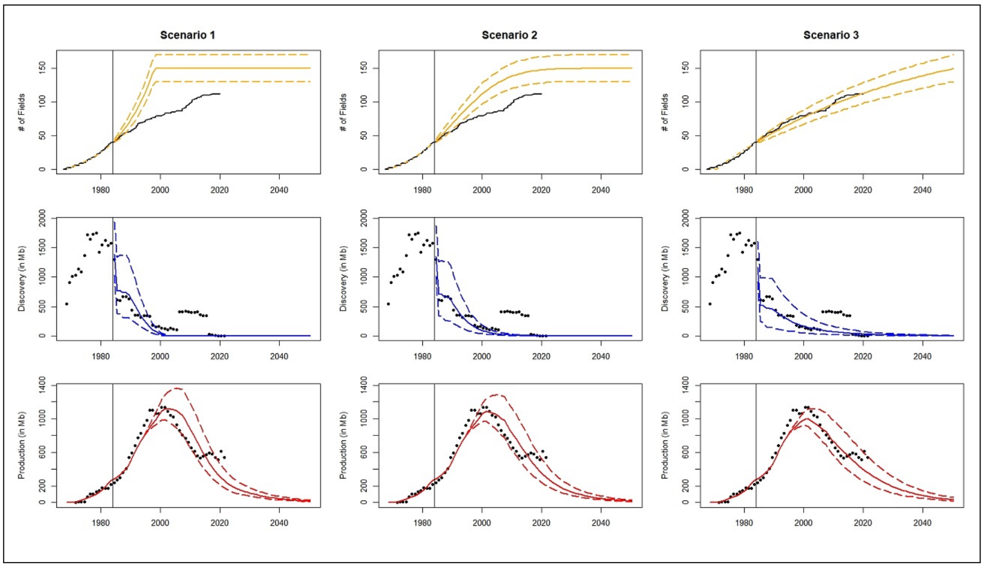
Suihong Song (Stanford University),
Tapan Mukerji (Stanford University),
Dongxiao Zhang (Eastern Institute of Technology, China),
Jiagen Hou (China University of Petroleum (Beijing))
Room: R7 2023-08-09 15:22

Facies geomodelling with GANSim has been drawing increasing attention. GANSim can directly produce multiple realistic geomodels conditioned to global features (e.g., facies proportion, channel width), sparse well data, and geophysics-produced facies probability maps. The original idea behind GANSim is borrowed from how geologists draw facies maps. A geologist learns geological patterns by observing lots of outcrops, modern sedimentary systems, etc. and then uses the learned patterns to directly draw reasonable facies maps consistent with the given conditioning data. Such a process corresponds to the training of a CNN-based generator and the production of geomodels directly from the given conditions with the trained generator in GANSim. The GANSim workflow has been validated with synthetic channelized reservoirs, and has been practically used for geomodelling of 3D karst caves, inner architecture of point bars, and submarine turbidite reservoirs. It has also been combined with physics-informed CNN surrogate to further achieve conditioning of geomodels to well flow data. The future research directions of GANSim include expanding to more reservoir types and depositional systems, digital rocks at the pore scale, etc. and training the generator to learn PDE-dominated physics.

S0114. A Stochastic Bottom-Up Model for Regional Oil Production Combining Size-Biased Sampling, the Poisson Process, and Production Profiles: With an Application to the Gulf of Mexico and Norway

Moritz Tarach (Technical University Darmstadt)
Room: R7 2023-08-09 15:24

In this paper, we formulate a stochastic bottom-up model for the rate of oil production in a region which is composed of three elements. The first element are empirically founded field-level production profiles (see, e.g., Höök et al. (2009)), which convert the size of an oil field into a continuous time path for production. The second element is an extension of the size-biased sampling model from Barouch and Kaufman (1976) as a model for the sequence of field sizes in discovery order. In essence, this allows the field sizes to diminish (or increase) with the succession of fields while still being stochastic variables, and the degree to which this occurs can be estimated from the data. Here, we assume a truncated lognormal distribution for the field sizes and add the assumption of a Poisson distribution for the total number of fields. We show how this alters the usual estimation procedure (see, e.g., Nair and Wang (1989)) which is based on the expectation-maximization algorithm. The third element of the overall model is a model for the discovery times of new fields, which we treat as stochastic variables generated by a birth process, or in particular, a Poisson process. For this, a cumulative intensity function, which essentially maps the expected discovery numbers onto the time axis, needs to be assumed exogenously or extrapolated from a past trend. Combining the three elements, a continuous time path for the overall production path in a region emerges simply as the sum of the individual field-level production paths. We repeat the parameter estimation for Norway and the U.S. Gulf of Mexico (for the deepwater part and the shelf part as separate regions) based on the oil fields which were discovered before a certain year. Based on the parameter estimates we calculate quantiles for the undiscovered resources in each region, which are mostly lower than the point estimates found in official resource accounts. Combining the estimation results with the other components of the model, we conduct Monte Carlo simulations to project the the rate of oil discoveries and the rate of oil production into the future.



Scenario 1
Scenario 2
Scenario 3
# of Fields
Discovery (in Mb)
Production (in Mb)
1980
2000
2020
2040

S02 Mathematical methods and approaches for monitoring and predicting ground movements

Jan Blachowski (Wroclaw University of Science and Technology),
Jörg Benndorf (TU Bergakademie Freiberg),
Steinar L. Ellefmo (NTNU)

The session focuses on topics related to theoretical and applied studies of phenomena related to the effect of mining and post-mining on their surroundings. Such as: - application of GIS, spatial statistics, machine learning and deep learning for modelling of spatial relationships in mining areas, - applied earth observations for monitoring of mining activity, - multi-source data fusion for analysis and modelling of mining areas, - assessment of mining activities to mitigate their environmental impacts, - case studies of innovative methods and applications.

S0201. Ground surface deformation modelling with spatial and random forest regression approach

Jan Blachowski (Wroclaw University of Science and Technology),
Steinar L. Ellefmo (NTNU)
Room: R8 2023-08-08 10:00

Sublevel caving mining causes ground surface deformation in the form of caved, fracture and continuous deformation zones. The external boundary of the furthest continuous deformation zone represents the limit of the mine influence. The most of the published studies focus on the caved and fracture zones. Whereas, the total extent of the mining related deformation, as well as the factors contributing to the deformation are not yet fully recognised.

This study has been aimed at estimating the extent of the ground surface deformation in the hanging-wall, identifying the driving factors and assessing their statistical significance.

A methodology based on the combination of PSInSAR Line-of Sight (LOS) ground surface movements, spatial statistics, weighted spatial regression and random forest regression in geographic information system has been proposed to investigate the problem. A geodatabase comprising of two dependent variables (LOS movements recorded between 2016-06-03 and 2021-10-11 for two ascending tracks A102 and A175) and 27 explanatory variables describing topographical and mining factors was developed. All the variables have been represented as attributes of a regular hexagon grid of reference units.

The study has been done for a case study of sub-level mining operation in Norway.

Global Linear Regression (GLR), Geographically Weighted Regression (GWR) and Random Forest Regression (RFR) models have been developed and tested for both the dependent variables and various combinations of explanatory variables. The RFR approach produced the best performing models with the coefficient of determination of 0.718 for track A102 and 0.769 for track A175. On average the RFR models were 9% more accurate than the GWR ones. The GLR models performed the worst due to non-stationary relations in the data.

The results of RFR modelling show that the following variables are the most significant driving factors to the ground surface deformation: distance to the deepest sub-level, depth to the deepest sub-level, distance to the caved zone, slope of the terrain. Depth and distance to the second deepest sub-level, digital elevation model and distance to neighbouring disused open-pit are also important factors.

The study contributes the better knowledge of ground surface deformation process in sub-level caving mining.

S0202. Data – Driven Prediction of Spatial Settlements on Mining Waste-Dumps – Combining Geotechnical Process Knowledge with InSAR-Data

Jörg Benndorf (TU Bergakademie Freiberg)
Room: R8 2023-08-08 10:20

Mining waste dumps belong can be characterized as risky construction ground. With the ongoing transition to a CO2-neutreal energy supply in Europe, most coal mines terminate their operations during the next years. What remains, are large areas of waste dumps. Alone in Germany this area has a spatial extend of about 3000 km^2. For a safe, environmentally and economically friendly usage of the post-mining landscape, the geotechnical risk needs to be well understood. One particular aspect are spatial settlements, which are characterized by regionally extended slow but long-lasting vertical deformations of the surface. Over the past decades, geomechanical models have been developed for predicting these movements. Challenges arise one the one side from the large number of parameters required, on the other side from the spatial heterogeneity of waste dumps. With the availability of satellite-based radar-monitoring data of ground movements, the deformation behaviour can be well observed. This provides a rich data base for a spatial-timely characterization of the deformation behaviour and an investigating of the settlement process influencing factors. In comparison to geomechanical models, in this way, data - driven complexity – reduced prediction models can be defined that allow to predict settlement behaviour in both, the time and space dimension. The contribution will introduce to the topic, present the methodology and demonstrate the method on a real example in German coal mining.

S03 Mining geostatistics, optimization and geometallurgy

K. Gerald van den Boogaart (HZDR, HIF),
Julian Ortiz (Queens University Kingston),
Raimon Tolosana Delgado (Helmholtz Institute Freiberg),
Jörg Benndorf (TU Bergakademie Freiberg)

The sessions aims to bring together all aspects of mining-relevant mathematical methods, along the whole mining cycle from exploration targeting to mine closure and on all scales from microstructure characterization to the long term mine planning and multi mine site operation planning. Important areas are: potential mapping, microstructural modelling and observation, geostatistics of geometallurgical variables, process modelling, high order geostatistical simulation for geometallurgical variables, structural modelling with uncertainty, stochastic mine planning, real time mining updating, and predictive process optimization. Contributions from all fields of application or development of geomathematical methods for mining are welcome.

S0301. Local Transformations using kernel density estimation for simultaneous detrending and Gaussian anamorphosis

Ilnur Minniakhmetov (BHP)
Room: R9 2023-08-08 14:00

Traditional methods for addressing non-stationarity in mineral deposits, such as trend removal, are limited in their ability to account for the wide range of non-stationarity observed in real-world deposits. This paper proposes a new approach that uses kernel density estimators in four-dimensional space (XYZ and variable) to perform local distribution transformations that result in locally Gaussian factors. This allows for the replication of complex non-stationary patterns without creating artificial sharp boundaries and replicating smooth transitions from one area to another. The resulting factors have locally Gaussian distributions, allowing for the reproduction of complex non-stationary patterns after back-transformations. In addition, the proposed approach demonstrated its ability to highlight areas of non-stationarity due to changes in continuity and provide more stable values for variogram calculations in factor space. The method was tested on the Olympic Dam deposit, one of the larges deposits in the world, with high degree of non-stationarity, such as changes of the shapes of local distributions. The results compared to conven-

tional trend subtraction method and modern trend-residual decorrelation technique. The paper concludes with a discussion on the extension of the method to multivariate decorrelation.

S0302. Integration of Machine Learning and Geostatistics to improve the concentration mapping (Case study: Bauxite mining wastes in Puglia-Italy))

Sara Kasmaeeyazdi (University of Bologna),
Emanuele Mandanici (Associate professor),
Enrico Pivato,
Thorkild Rasmussen (Full professor)
Room: R9 2023-08-08 14:20

Machine learning (ML) approaches are widely used in geoscience applications. Many of them mainly are used for clustering and classifications of land cover features in mapping and mining exploration. On the other hand, the well-known geostatistical methods are still the best to produce the concentration mapping. In this work, the two methods are integrated to enhance the mapping quality. The methodology is performed over a historical mining residue in Apulia Region (South of Italy). The main challenge of mapping in this area is the low number of available samples (with Iron and Aluminum concentrations). To salve this challenge the high-resolution image of WordView-2 is used as the auxiliary variable. Thanks to the high-regular data from the satellite image, it has been possible to learn the pattern of the land-cover features for the concentration mapping. From ML, the self-organizing map (SOM) is performed to classify the area and to learn the pattern of the mining wastes. Then, each class is used as indicators and input for the Indicator Co-Kriging (ICK). Based on the existent correlation between the image pixels and the concentration sample values, the ICK is used to map the concentration maps of mining wastes. The SOM indicators help the estimation results to follow the land cover patters. Comparison of results with a simple Ordinary Kriging (OK), using concentration samples without satellite image data, demonstrates the possibility of re-creating the image patterns within the concentration variability in the selected area. Final maps indicate the advantage of integrating remote sensing data in mapping concentration variability, even when few samples are available.

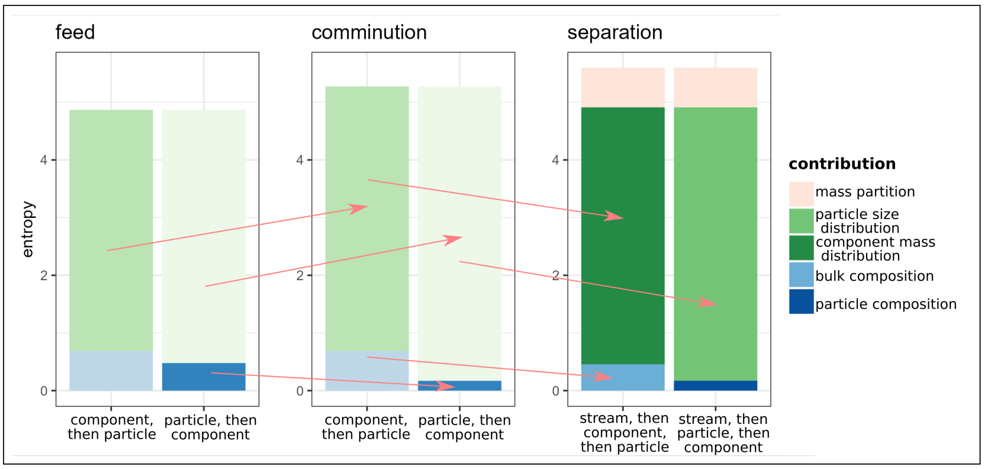
S0303. Concepts of statistical entropy for the evaluation of comminution and separation processes

Raimon Tolosana Delgado (Helmholtz Institute Freiberg),
Edgar Schach (TU Bergakademie Freiberg),
Urs Peuker (TU Bergakademie Freiberg)
Room: R9 2023-08-08 14:40

Statistical entropy, corresponding to Shannon entropy in information theory or Gibbs entropy in physical thermodynamics, can be used to evaluate the additional disorder in a system induced by its subdivision in components or subsystems. Several variations of this concept have been introduced in the last 20 years to quantify the efficiency of a separation process (or a chain of it) within the context of minerals processing both in mining and recycling. This contribution presents a coherent frame connecting these concepts, and extends them to evaluate comminution processes, in such a way that joint thermoeconomic zoptimization of whole beneficiation plants becomes possible.

The main idea is to split the system simultaneously into three dimensions: (1) mineral or chemical components, (2) particles or particle classes, and (3) output streams. In each of these dimensions, a different ratio can be defined to describe the way the mass (or volume or matter) of the system is split into the subsystems: (1) a composition; (2) a particle-wise distribution; a (3) set of partition coefficients. Entropy contributions can then be defined for each of these dimensions at different levels of integration, e.g. (1) component-wise or for a total, (2) particle-wise or bulk-wise, (3) per stream or per stage feed.

By systematically considering all possible entropy decompositions, the following results were found. Total stage entropy does not depend on the ordering of integration through the three dimensions. Comminution cannot decrease the stage entropy; a comminution that would generate perfectly liberated particles without overgrinding would keep the stage entropy constant. A separation process cannot decrease stage entropy, and perfect separation would keep it constant. However, once the masses of the output streams are measured, the output entropy does indeed decrease for any moderately good separation process. The difference between the stage entropy with and without controlling output masses can then be understood as the information value of the output mass flow measurements.



S0304. High-Order Simulation of Processing Responses Applied to Adaptive and Reinforcement Learning-Based Short-Term Planning of Industrial Mining Complexes

Joao Pedro de Carvalho (McGill University),
Roussos Dimitrakopoulos (Professor - McGill University)
Room: R9 2023-08-08 15:00

Processing responses relate to attributes or requirements measured at a processing plant facility or equipment, for example, mill throughput, energy consumption, metallurgical recovery, and reagent consumption, amongst others. Several studies have demonstrated these attributes are a function of the material feed characteristics beyond metal grades. However, the impact of processing plant responses over production planning is often underestimated. Industrial practices often consider a single estimated factor to represent these attributes. Only recently, some approaches used production data to model these processing responses through multivariate regression or non-linear approaches as a function of the characteristics of the material feeding the processor. Yet, the uncertainty associated with these attributes is not thoroughly assessed. This work proposes a new method that simulates processing plant responses as a function of the input characteristics and past performance of the processor using high-order statistics. This allows capturing more relevant statistics and patterns from the historical production data. The uncertainty characterizing the ore material is also included, related to metal grades, deleterious elements and geometallurgical attributes. Then, the approach is combined with a policy gradient reinforcement learning approach that assigns shovels to grade-control diglines while considering this uncertainty in processing responses. The approach is applied to an operating gold mining complex, where the addressed processing responses include ball mill throughput and consumables usage. The results show the benefits of the uncertainty assessment related to processing responses and how the proposed short-term approach profits from this more informed mining complex operation.

S0305. Integration of Machine Learning for grade estimation in mineral resource modelling.

Ilyas Ongarbayev (University of Oulu),
Jukka-Pekka Ranta (Universoty of Oulu),
John Carranza (University of the Free State),
Nasser Madani (Nazarbayev University)
Room: R9 2023-08-08 15:20

Resource modeling is an integral part of geological exploration. Resource modeling traditionally employs geostatistical methods, such as kriging, to estimate grade distribution within a mineral deposit. However, these methods may lead to suboptimal results due to assumptions of first and second-order stationarity, linear nature, and subjectivity.

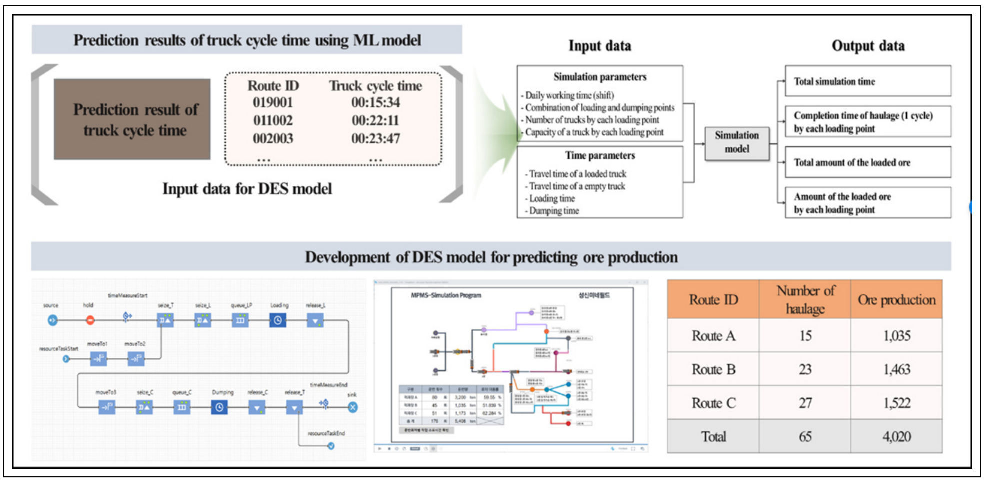
The integration of machine learning in geological applications has gained popularity due to a drastic increase in computational power and relative simplicity compared to geostatistical methods. Machine learning offers several advantages over geostatistical tools such as the ability to model complex, non-linear relationships between different variables, providing potentially more accurate predictions compared to geostatistical tools, especially with complex or poorly understood systems.

This Ph.D. project is focused on investigating the potential of machine learning in grade estimation during mineral resource modeling. The project is part of the European Union-funded SEMACRET project, focusing on understanding magmatic Ni-Cu-PGE-Cr-V-Ti systems in regional and deposit scale contexts. Reference deposits include Akanvaara Cr-V-Ti (Finland) Suwalki Fe-Ti-V (Poland) and Ransko Ni-Cu-PGE-Co (Czech Republic). The research highlights the potential of machine learning techniques for improving geological modeling and producing more accurate grade estimation for resource models.

S0306. A hybrid model of machine learning and discrete event simulation techniques to predict ore production in mining operations

Yosoon Choi (Pukyong National University),
Sebeom Park (Pukyong National University),
Dahee Jung (Pukyong National University),
Jeong-Gi Um (Pukyong National University),
Jieun Baek (Pukyong National University)
Room: R9 2023-08-08 15:22

This study proposes a new model for predicting ore production by combining machine learning (ML) and discrete event simulation (DES) techniques. After selecting a limestone underground mine as the research area, time data of truck operation were collected using tablet computers and Bluetooth beacons for 15 weeks. Training of the ML model was conducted using the collected data to predict the truck cycle time. As a result, the support vector regression with particle swarm optimization (PSO-SVM) model exhibited the best performance. The PSO-SVM model showed the lowest error at a mean-absolute-error (MAE) of 2.79 min, mean-squared-error (MSE) of 14.29 min, root-mean-square-error (RMSE) at 3.79, and R2 at 0.68. The cycle time for each truck and section, as predicted using the PSO-SVM model, was linked with the DES model to predict ore production by time period. Verification of the DES model was performed by comparing the production log created at the site with the simulation results, which showed that the simulation model properly simulated the haulage system in the study area. It was confirmed that the developed ML-DES hybrid model could be utilized to enhance the productivity of mining haulage systems. In particular, it can predict ore production according to truck and section to determine the optimal combination of equipment for each workplace.



S0308. Resource Modeling in Copper Deposits using Multivariate Geostatistics

Artur Korniyenko (Nazarbayev University),
Nasser Madani (Nazarbayev University)
Room: R9 2023-08-08 15:24

This work addresses the problem of implementing geostatistical co-simulation for mineral resource estimation (MRE) of a multi-element copper deposit. MRE is an important step throughout all the stages of the mine life cycle which predicts the properties of an orebody such as grades of the elements to construct a geological block model of an orebody. The accuracy of MRE affects the evaluation of recoverable resources in a deposit. In general, MRE influences the designing, strategic mine planning and profitability of downstream mining activities. In order to make the MRE more accurate, geostatistical interpolation approaches could be used since they take into account the spatial continuity of the regionalized variables (e.g. anisotropy). For instance, kriging is one of those methods which is currently used in the mining industry in order to estimate the desirable variable at unsampled locations. Despite the fact that kriging could estimate regionalized variables, it has some drawbacks such as the smoothing effect. Therefore, geostatistical simulations were proposed as an alternative to kriging, which also considers the spatial continuity of the regionalized variables without the problem of the smoothing effect. However, in deposits with the existence of co-variates, the independent simulation might be suboptimal, because it ignores the intrinsic correlation between the variables. Therefore, a co-simulation paradigm can be advocated in such complex situations. A copper deposit is considered in this study, where copper (Cu) is correlated with molybdenum (Mo). The linear correlation coefficient between Cu and Mo grades on sampling locations is measured as 0.47. Reproduction of this intrinsic correlation is of paramount importance from the geological perspective. Hence, co-simulation for modeling the Cu and Mo grades in this copper deposit was utilized as an alternative algorithm to independent simulation. Next, co-simulation and independent simulations were compared based on the reproduced correlation coefficients between realizations of Cu and Mo grades. To make this comparison, 100 realizations for both cases were produced. Other statistical parameters were also examined and it was deduced that co-simulation performed better in this deposit than the independent simulations, since the correlation coefficient between Cu and Mo is considered in the simulation process.

S0309. Application of Multivariate Geostatistics for Geospatial Modelling of Tailings Storage Facilities in Kazakhstan

Aidyn Tileugabylov (School of Mining and Geosciences, Nazarbayev University), Nasser Madani (Nazarbayev University)

Room: R9 2023-08-09 10:00

Tailings storage facilities (TSFs) play a critical role in the management and disposal of mining waste. In Kazakhstan, TSFs are mainly distributed in the southern, central, and eastern parts of the country. This study focuses on the geospatial modelling of the Kazakhstani TSFs to understand their degree of environmental impact for better management in terms of economic interest. Throughout each TSF, continuous and categorical variables are quantified, including the facility's capacity, crest width, safety factor (SF), tailings hazard index (THI), water and overall environmental hazards. In order to identify the spatial distribution of such models across the country, a multivariate geostatistical method, known as turning bands co-simulation, and an indicator-based approach, known as sequential indicator simulation were used for modelling of continuous and categorical variables, respectively. To better understand the situation of TSFs throughout the country, after the geostatistical modelling process, clustering algorithms were implemented to identify the regions that show the same characteristics. To accomplish these goals, machine learning algorithms such as Geostatistical Hierarchical Clustering (GHC) and k-means have been implemented. The outcomes of GHC clustering were compared with those of the k-means clustering algorithm. As a result, GHC showed a stronger association with each variable included in the modelling and it provided more contiguous and compact clusters. However, the failure of the k-means algorithm to illustrate patchy clusters leads to a conclusion that GHC is superior to the k-means clustering method for spatial data modelling and clustering in terms of reproducing the compact clusters. The results of modelling using GHC showed that TSFs with larger capacities and with high crest widths are mostly distributed in the central part of Kazakhstan. The TSFs with the highest SF are disseminated in the eastern region of Kazakhstan near Ust'-Kamenogorsk city. The most hazardous TSFs, as quantified by the highest THI throughout the clusters, are situated in the southern part of the country near Shymkent city. The results obtained from the GHC could help to identify the potential environmental impact of TSFs and ensure that proper remediation measures are put in place to prevent soil, water and air contamination.

S0310. Geostatistics for mineral resource assessment of critical battery elements

Milena Nasretdinova (Nazarbayev University),
Nasser Madani (Nazarbayev University),
Mohammad Maleki (Department of Metallurgical and Mining Engineering, Universidad Católica del Nor)
Room: R9 2023-08-09 10:20

Elements for the construction of batteries are currently experiencing growing demand due to the need to create lithium-ion batteries in electronics, electric car engines and other vehicles as well as energy distribution and storage systems against the green development of the economies of the countries. Frequently, the target metals for batteries are cobalt, copper, lithium, nickel and molybdenum. The increased interest for these elements may cause a risk of supply in the upcoming years, which calls these elements critical. To come up with this difficulty in the future, sulfidic deposits such as copper porphyries can be considered as the primary sources of extracting these elements. This study is devoted to the development of an algorithm based on integration of machine learning and geostatistics for assessing mineral resources of the abovementioned critical battery elements in a copper porphyry deposit, where their local and spatial distributions are being regulated by geological properties such as rock type, mineralization, and alteration. The methodology includes, first loosening the compositional nature of these continuous variables by using additive log-ratio (alr) transformation. After this step, the log-ratios were subjected into principal component analysis (PCA) by also considering the indicators of the geological domains. The first two PCs were selected and estimated over the target blocks using the simple kriging and two domains were inferred using fractal analysis of the estimated PCs. Then, the domains were migrated from block to the sample location, and a hierarchical resource modeling was realized. For this purpose, the log-ratios belonging to each domain were modelled at the target domain already obtained by fractal analysis. The sequential Gaussian co-simulation and simple co-kriging were utilized for the former and latter modelling approaches, respectively. Furthermore, a mineral resource assessment was considered to quantify the recovery functions for the whole deposit to compare the modelling results. It is shown that sequential Gaussian co-simulation reproduced much higher mean grade and metal quantity for all five critically battery elements with better reproduction of statistical parameters. The results of this study demonstrated that copper-porphyry deposits possibly can be considered as one of the primary sources of critical battery elements.

S0311. Validating non-stationary Conditional Simulations at Olympic Dam Deposit: Benchmarking against Estimation Techniques using data from the recent drilling

Ilnur Minniakhmetov (BHP),
David Clarke (BHP)
Room: R9 2023-08-09 10:40

Accurately simulating the geology and mineralization of the Olympic Dam (OD) deposit for underground mine planning purposes poses many challenges. This paper discusses how the authors addressed these challenges using several geostatistical techniques, including projection pursuit multivariate transform, local Gaussian transformations, hierarchical truncated Gaussian boundary simulation, and spectral simulation method. The resulting non-stationary conditional simulation models were validated against various benchmarks and historical samples to assess their accuracy. The validation includes comparison of univariate statistics, grade-tonnage curves, swath plots, variogram reproduction at validation samples locations. The authors compared their models to alternate estimation techniques, such as ordinary kriging and local uniform conditioning, and found that their models provided more robust predictions of in-situ resources. The paper has demonstrated the superiority of the conditional simulation method in predicting in-situ resources in OD deposit without the limitations of smoothing or conditional bias that are inherent in conventional deterministic approaches and provides an accurate estimation of uncertainty, which is crucial for making informed decisions in downstream processes. Overall, the authors present an innovative approach to accurately modelling a complex deposit using state-of-the-art geostatistical techniques.

S0312. Resource risk quantification through automation and sensitivity analysis – an Alcoa's case study

Hadrien Meyer (DeepLime),
Johann Dangin (DeepLime, Perth WA 6008),
Claude Cavelius (Development Lead, DeepLime, Paris France),
Lucas Tuckwell (Alcoa, Perth WA 6008),
Quentin Swart (Senior Resource Geologist, Alcoa, Perth WA 6008),
Thomas Green (Resource Development Manager, Alcoa, Perth WA 6008)
Room: R9 2023-08-09 11:00

This paper will address novel ways of quantifying the uncertainty of a resource model. It is only through the automation of the modelling, that we can go beyond the traditional use of grade simulations, and look into a much more encompassing study of the impact of geological assumptions through sensitivity analysis.

The art of automating parts of the geologist's workflow has always proven difficult, due to the endless variety of geological settings every mine has to offer. Some of the instruments most recently adopted by the industry, such as implicit modelling, are supplementing the geologist's tool box on its journey towards computerisation of its workflow. However, when it comes to fully putting the geological data into processes, very little solutions exist, and the integrity of the modelling workflow is still a very patchy and manual process.

We will discuss in this paper how the use of common data-science tools became an inspiration and solution to tackle the problem of flexible automation within the diversity of modelling methods.

An Alcoa bauxite ore body, located in Western Australia's Darling Ranges, is used to illustrate the implementation of an automated routine going from drill hole data validation to block model estimation.

The use of common and trade-specific Python tools enable a very modular design of many, but not all, processes, leaving the space to more traditional 'drawing' tools in some instances. However, for the parts they do cover, automation of modelling offer at the very least an important time gain in the decision making, as well as process transparency.

A focus is then drawn to uncertainty analysis. The repeatability of modelling allows for a much more comprehensive understanding of influential factors and risk profile of a model. The spread of results is appraised, together with the quantification of the impact of each variable towards that spread.

S0313. Simultaneous stochastic optimization of mining complexes and geostatistical simulations: Key aspects and interactions

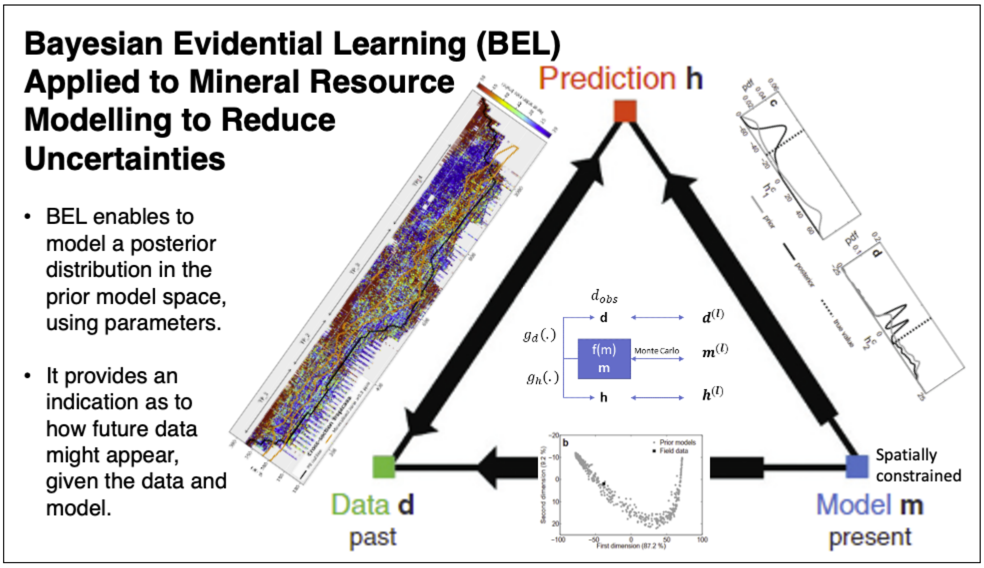
Roussos Dimitrakopoulos (Professor - McGill University)
Room: R9 2023-08-09 11:20

Mining complexes or mineral value chains are integrated engineering systems where raw materials extracted from mineral deposits are transformed into a set of sellable products. The supply of extracted materials and their quality from available mines represents a major source of uncertainty and technical risk that requires management. The mathematical framework for the simultaneous stochastic optimization of mining complexes is first overviewed, and outlines the stochastic integer programming with recourse utilized and its direct link to fandom field models and geostatistical simulations. Subsequently, pertinent aspects of the geostatistical simulation methods used to generate orebody realizations are explored. In doing so, the different geostatistical modelling frameworks and their effects on optimization results are shown, including the (i) established Gaussian models, and (ii) high-order ones that address complex spatial connectivity. The issue of how many scenarios of simulated mineral deposit realizations are needed to provide stable results in the simultaneous stochastic optimization of mining complexes is also overviewed. For single mines, about 10-15 simulated scenarios are sufficient. Meanwhile, for mining complexes with more than one mine, a relatively small proportion of combinations, like 15 to 20%, of the initial simulations are sufficient, thus facilitating the efficient solution of exceptionally large and complex applications.

S0320. Bayesian Evidential Learning (BEL) Applied to Mineral Resource Modelling to Reduce Uncertainties

Aafje A.L. Houben (Delft University of Technology),
Masoud Soleymani Shishvan (Delft University of Technology)
Room: R9 2023-08-09 11:40

Mineral resource modelling (MRM) requires enough geological information to define the geological model. The success of a mining project is supported by the accuracy of this model and its interpretation. Major failures can occur as a result of an incorrect degree of uncertainty quantification in the geological/geometallurgical models. There are different techniques in the industry today to reduce uncertainty in MRM. All techniques respect the statistical and geostatistical properties of the constraining data, although they vary in the specifics and the approach and they all rely on the stationarity assumption, which is not a testable hypothesis but rather the choice to collect data from a certain area or domain. This paper aims at developing a framework for reducing associated uncertainties with MRM using Bayesian Evidential Learning (BEL). BEL enables to model posterior distributions in prior model spaces using predefined parameters. It provides an indication of how future data might appear, given the data and model. The Walker Lake dataset is used to test the framework. The objective is to reduce the uncertainty in the prediction of the hardness of the orebody. First, the model is built with data from lithology, mineralogy, penetration rate and grade. These properties are obtained from samples that are spatially correlated. Then, Monte Carlo realizations are obtained based on the exploration data and the assumed uncertainty range. A relationship needs to be obtained between lithology, grade and mineralogy and hardness variables. PCA is applied to get a better visualization by looking at the most influential properties. The observed data are used to compare and see if the prior model needs to get falsified. It is determined that the penetration rate and the lithology are the most influential properties. After that, Canonical Correlation Analysis (CCA) is applied to find the combination of the variables that have the maximized linearity between the penetration data and the prediction data. The predictions are made and then back-transformed to their original space. Finally, the hardness predictions are not falsified by the observed data from the drillholes. These predictions are used to domain the orebody into soft, medium or hard materials.


Bayesian Evidential Learning (BEL) Applied to Mineral Resource Modelling to Reduce Uncertainties
• BEL enables to model a posterior distribution in the prior model space, using parameters.
• It provides an indication as to how future data might appear, given the data and model.
Prediction h
Data d
past
Model m
present
Spatially constrained
d_{obs}
d
$g_d(.)$
f(m)
m
Monte Carlo
$g_h(.)$
h
$d^{(l)}$
$m^{(l)}$
$h^{(l)}$
Prior models
Field data

S0321. Developing a New Data-driven Multiple-point Simulation Method Based on Copulas

Babak Sohrabian (Hacettepe University),
A. Erhan Tercan (Hacettepe University)
Room: R9 2023-08-09 14:00

Multiple-point simulation methods have been developed to reproduce complex patterns and curvilinear structures. These methods are based on the calculation of the probability of occurrence of a value regarding a bunch of samples using higher-order statistics. Multiple-point statistics are calculated based on some predefined training images or data of fully explored similar deposits. The use of training images that are often suggested by specialized experts is a controversial issue. Moreover, it is very rare and perhaps impossible to come across two deposits with the same statistical characteristics. Therefore, it may not be reasonable to apply the multiple-point statistics of a fully explored deposit to another one.

This study proposes a new data-driven multiple-point simulation approach that extracts multiple-point statistics from sparse data without using any training images or predefined other sources of information. The suggested approach takes advantages of Archimedean copulas in calculating the probability of occurrence of different values at unknown locations. It uses the conditional independence assumption to avoid modelling the joint distribution of conditioning data. Application of the proposed approach to two satellite images demonstrated that it is capable of extracting the multiple-point statistics of a fully known image based on its small subsets (almost 7% of the whole pixels). Comparing to Filtersim and sequential Gaussian simulation, the method showed reasonable performance in reproducing cumulative distribution function, variograms, and directional connectivity.

S0322. Plurigaussian collocated co-simulation: an innovative technique for stochastic modeling of geological domains

Collins Gogoe Adoko (Nazarbayev University),
Nasser Madani (School of Mining and)
Room: R9 2023-08-09 14:20

Modelling of the rock units of a mineral deposit is crucial to accurately estimating the amount of mineral resources as it informs the mineralization controls of the deposit. The deterministic modeling methods make assumptions about the geology, particularly between drillholes, and as a result, they might biasedly predict the deposit's complexity. In addition, these methods are not able to quantify the uncertainty of rock units at unsampled locations. These shortfalls are better addressed by stochastic methodologies such as sequential indicator simulation and plurigaussian simulation. Plurigaussian simulation has become a vital simulating technique for delineating rock units of complex geometry and contact relationships because of the ability to accurately depict the spatial correlation between rock units and quantify their uncertainties. However, using this method, long-range geological features like veins, faults and fractures are often loosely modelled thus necessitating the use of soft data from a deterministic interpretive geological model to avert this inaccuracy. In this research, the plurigaussian simulation is applied to a gold deposit of four rock units (including two veins) to assess the spatial variability and to quantify the uncertainty in the occurrence of each rock unit. A new methodology is presented where a collocated co-simulation algorithm is incorporated in traditional plurigaussian simulation to model the rock units at target block locations. The realizations are conditioned to drillhole data and a deterministic geological model that is already constructed using support vector machine in Leapfrog software program. The method is then compared with plurigaussian simulation that considers the local proportions of rock units obtained from the same deterministic geological model. In all, 50 realizations are produced from the model which is validated with a split-sample validation procedure to authenticate the plurigaussian model. The outcome allude to the correct reproduction of the proportions of the rock unit, as representing the drillhole data and accurately quantifying the uncertainty thereof. The advantages and disadvantages of these two methods are discussed and proper guideline is provided.

S0323. Generalized Laguerre Mosaics as versatile toolbox for microstructure modelling

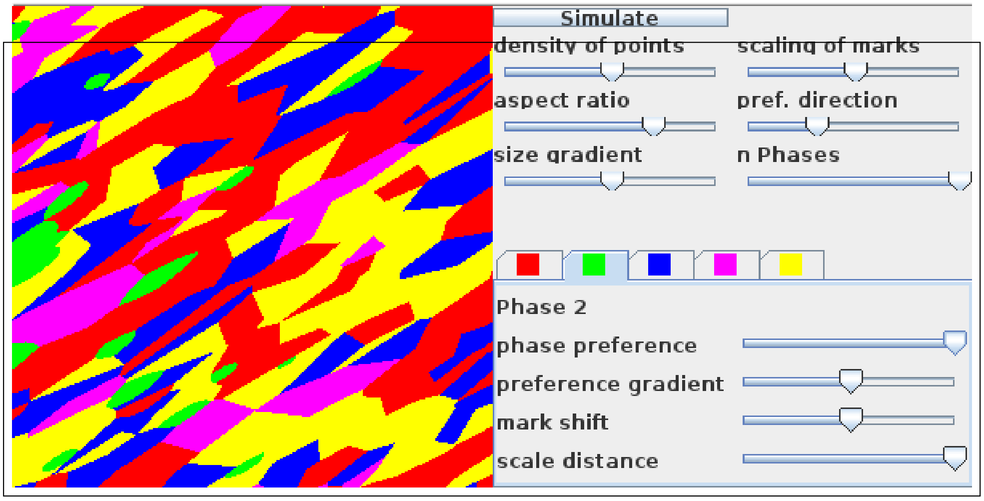
K. Gerald van den Boogaart (HZDR, HIF),
Raimon Tolosana Delgado (Helmholtz Institute Freiberg),
Felix Ballani (HZDR)
Room: R9 2023-08-09 14:40

The particle based simulation of minerals processing often requires models of 3D particles, while only insufficiently resolved 3D information from CT or sufficiently resolved 2D information on the structure of the material is available. Fitted and simulated microstructure models can support us with relevant 3D microstructures.

The contribution proposes several steps to generalize 3D Laguerre mosaics that allow to recreate various typical features of microstructures with meaningful parameters to the simulation process: Multiple phases with different abundance, preferred contacts, different grain sizes, spatial variation of properties, flattened crystallites, preferred orientation, flat and rounded boundaries, etc.

A special feature of the model is the possibility to change the parameters gradually for the same simulation, such that one can track the effect of the parameters visually and for estimation algorithms with lower variability than with a resimulation.

The same microstructure can be computed for 3D voxel spaces, in 2D sections and on MPS type patterns for estimation of high order statistics.



S0325. M3G: Music as a Metaphor in Mining Geostatistics

Steinar L. Ellefmo (NTNU),
Torkjell Breivik (Norwegian University of Science and Technology),
Micah Nehring (University of Queensland),
Julie Ballantyne (University of Queensland)
Room: R9 2023-08-08 15:24

Minerals and metals extracted through mining activities are essential components of various everyday items and are crucial for transitioning to renewable energy sources and e-mobility. The complexity of mining requires handling uncertainties related to social, legal, ecological, technical, and geological factors. Geostatistics, a field that deals with quantifying uncertainties and associated risks, is used by geologists and mining engineers to estimate, and assess the value of minerals in the ground, which is fundamental to any mining project.

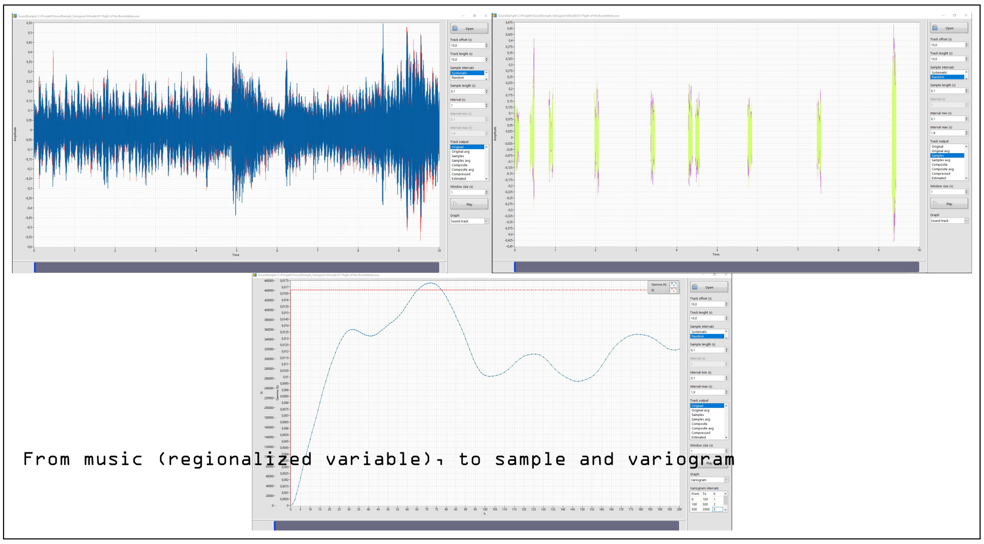
Geostatistics involves the concept of ”regionalized variables” which can be modelled as the sum of a structured aspect (drift) that represents large-scale geological changes and a spatially correlated uncorrelated random aspect, both representing smaller scale geological variability. The variogram is used to model the spatial correlation of the regionalized variable and is used to find the best estimate.

Traditionally, geostatistics is taught at universities through lectures, tutorials including hands-on exercises applying software solutions for geostatistics to realize estimation and simulation using geological data. However, this conventional approach may limit students' understanding of fundamental concepts like the 'regionalized variable' and implications of different sampling techniques and -strategies related to estimating the ”grade in the ground.”

To address this, a new teaching approach called ”Music as a Metaphor in Mining Geostatistics” (M3G) has been developed. This approach uses music as a metaphor to help students grasp the concept of regionalized variables in geostatistics and visualize the impact of different sampling strategies on the estimation process. Instead of using geological data, students interactively experience the effect of sampling techniques through sound (e.g., music, but also “random noise” like the sound from a river), engaging them in a constructivist learning approach within a group setting.

This metaphorical exploration approach tries to relate music characteristics such as dimensions, rhythm, brightness, melody, pitch, and timbre with ore deposit characteristics like ore grade, strength, hardness, and grain size. This creative approach aims to enhance students' understanding and application of geostatistical concepts in mining engineering.

Sessions with students have shown that the M3G-aid increases student's understanding of the differences between random and systematic sampling, how and why different geological settings / music genres have different spatial dependencies and lead to different variograms, the importance and effect of sampling density and sample size and the importance of a-priori knowledge and understanding of the regionalized variable. Eventually, our contribution is meant to encourage IAMG to put more emphasis on efforts teaching geologists and mining engineers to apply mathematics properly.


From music (regionalized variable), to sample and variogram

S05 Rare events detection for risk management in geoengineering

Alla Sapronova (Graz University of Technology),
Thomas Marcher (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)

Proposed title: Detecting rare events with data-driven modeling. Session description: The session will address the problem of rare event detection: with a vast amount of collected observations, the events having a significant impact and long-lasting consequences are often the most uncommon in frequency. We propose to discuss tools, methods, and current best practices for rare event detection in geosciences, with a particular focus on risk management and its application in geoengineering. The session will present new proposals and findings from the scientific community that include but are not limited to data preprocessing, feature engineering, rebalancing, application of statistical or machine learning methods, and demonstrations of progress in rare event detection that can promote a proactive approach in risk management. This section will cover: 1. overviews of methods and recommended workflows for handling data sparsity and uncertainty; 2. reviews of new findings and techniques for rare events' detection and recommendations on data preprocessing methods that help to ensure both data quality and novelty in patterns; 3. reports on how various methods for handling data imbalance affect the accuracy of forecasting data-driven models; 4. approaches used for detecting perceptions and biases in data labeling and interpretations; 5. demonstrations of methods and tools successfully employed for risk modeling and rare event detection.

S0501. Advanced numerical model for landslides: from quick clay to submarine landslides

Quoc Anh Tran (NTNU)
Room: R8 2023-08-08 10:40

In this abstract, we present recent developments in the Material Point Method applied to the simulation of landslides, focusing on practical applications rather than theoretical developments. There are a variety of applications, ranging from the 3D simulation of Gjerdrum landsides to earthquake-induced submarine landslides.

Material Point Method (MPM) is a continuous-based method for modeling large deformations. A background mesh is used in MPM to discretize the Lagrangian particles (or integration points in FEM). As in FEM, weak differential equations are also solved on a background mesh. Unlike FEM, solutions (e.g., deformations) are updated using particles rather than meshes, and the background mesh is reset to its original configuration. Furthermore, the method is coupled with CFD to study the interaction between soil, fluid, and structure, such as earthquake-induced submarine landslides. The abstract does not present the algorithm but here we only demonstrate the application of the Material Point Method in two examples:

(1) 3D simulation of Gjerdrum quick clay landslides.

(2) earthquake-induced submarine landslides.

S0502. Application of unsupervised learning methods for predictive modeling in geoengineering

Alla Sapronova (Graz University of Technology),
Paul Johannes Unterlass (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)
Room: R8 2023-08-08 11:00

To develop robust decision support systems in geoengineering, the analysis of data, mainly originating from signals processing and interpretations (e.g., data from drilling, lidar, or seismic exploration), shall be performed, and the predictive model shall be developed and validated. Geoengineering data is often characterized by large volumes, sparsity, and class imbalance. Combining these three characteristics is a primary constraint for applying machine learning (ML) for data-driven modeling: training ML models to forecast the underrepresented events would require data rebalance and sparsity reduction. However, the application of the most efficient rebalancing methods (like data oversampling or data synthesis) is limited in the case of big datasets, while other methods (e.g., random sampling) often result in information loss and, therefore, overall accuracy reduction.

This work demonstrates the application of unsupervised methods (PCA, K-means, and Spectral Clustering) for big data sub-setting to reduce the data volume yet minimize information loss. The produced subsets were rebalanced with various oversampling methods and then used to train supervised ML models. Finally, trained ML models were validated on the original (raw) data to forecast the rare (underrepresented) events.

This approach was applied to data from several engineering constructions (tunneling) sites, and various evaluation techniques, including confusion matrix based (e.g., F1Score), Log Loss, and Receiver Operating Characteristics (ROC) curve, were used to evaluate the accuracy of the predictive model.

S0503. Use and limitations of various metrics to assess the quality of extreme sparse datasets in geotechnics

Matthias Hahn (TU Graz & Montanuniversität Leoben),
Alla Sapronova (Graz University of Technology),
Marlene Villeneuve
Room: R8 2023-08-08 11:20

1 INTRODUCTION

Like the comparison of the round time in Formula 1 car racing, metrics show which machine learning algorithm best fits the data prediction requirements. The question is, are all metrics suitable for every algorithm and data set in geotechnics? This research gives a first recommendation of the stability of different metrics for regression and classification predictors for a wide variety of geotechnical data sets.

2 METHODOLOGY

Training and performance measurements of different algorithms on various data sets are used to show the stability of the metrics to eliminate the problem of subjectivity of parametrization. For this reason, a large number of different data sets are needed. Geotechnical data for machine learning are raw and often sparse, so to eliminate subjectivity the solution was to split every data set randomly into training and prediction subsets and repeat this procedure 300 times per data set. In this way the normalized standard deviation can be calculated, and the stability of the metric can be estimated. To make the metrics comparable, one probability distribution is used for every algorithm over all applied metrics. The stability of a metric is rated by Kolmogorov-Smirnov p-value > 0.05 and normalized standard deviation < 0.5.

3 CONCLUSIONS

The used methodology shows, in a limited way, how stable different metrics are, addicted to the algorithm. A recommendation for a small set of algorithms and metrics is possible. To validate these results, more data sets are needed. The best case would be, that the 300 times repeated training and performance measurement cycle can be substituted by a sufficiently large number of different data sets. Python provides more machine learning algorithms and metrics, which have not been examined in this research. If enough computational power is available, the algorithm-metric recommendation in this research can be done for all implemented algorithms and metrics in Python/Scikit-learn package.

S0504. Conditional WGAN-GP approach for geotechnical dataset augmentation

Paul Johannes Unterlass (Institute of Rock Mechanics and Tunnelling, Graz University of Technology),
Alla Sapronova (Graz University of Technology),
Georg Erharter (Graz University of Technology),
Thomas Marcher (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)
Room: R8 2023-08-08 11:40

In this work we propose a conditional generative adversarial network (CGAN) with gradient penalty (GP) based approach of generating synthetic tabular geotechnical data for dataset augmentation. In geotechnical datasets events of utmost importance, so called rare events (e.g., fault zones during tunnel excavation, sudden increase in displacements shown in geodetic measurement data, etc.) are heavily underrepresented. Conditional GANs present a promising approach to to generate synthetic geotehnical data based on a condition (e.g., underepresented class data) that shows the same characteristics as the original data, but still presents unique samples with no connection to the technical content of the original data.

A Wasserstein-GAN algorithm with Gradient Penalty (GP) for training stabilisation is used to generate the synthetic data and augment the existing dataset. The demands on the synthetic data are of a dualistic nature: on the one hand, the data has to be suffciently dissimilar to the original data, so that it does not simply create copies f the existing data (demand for originality). On the other hand, it has to show the same patterns and follow the same rules as the original data, so that it can be used as if it were real data (demand for conformity). The Conditional WGAN-GP model describes how a synthetic dataset is generated, in terms of a probabilistic model based on real data. By sampling from this model, we are able to generate new, unique and conditioned synthetic and realistic data.

We show that both imposed demands on the newly generated data are fulfilled and an existing data set's rare events can be sufficiently augmented. Thus, enabling balancing of unevenly distributed classes and enhancing the overall size of a dataset. The authors see big potential for subsequent analysis on enhanced and balanced datasets, which might lead to more accurate predictions of rare events in geotechnical engineering.

S0505. Balancing the datasets: an overview and evaluation of machine learning approaches

Alla Sapronova (Graz University of Technology),
Paul J. Unterlass (Institute of Rock Mechanics and Tunnelling, Graz University of Technology),
T. Dickmann (Amberg Technologies AG, Switzerland),
J. Hecht- Méndez (Amberg Technologies AG, Switzerland),
Thomas Marcher (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)
Room: R8 2023-08-08 12:00

Data-driven models that support decisions in geoengineering usually employ predictive analytics to forecast the probability of extreme events like water inrush, support failure, or similar. The development of reliable methods showing accurate predictions depends on the data quality. However, one of the main difficulties in predictive modeling with geotechnical data is the extreme sparseness and imbalance of datasets (i.e., some classes or types of events have much more observations than others). Therefore, applying "data hungry" intelligent algorithms (e.g., machine learning) in geoengineering meets constraints: when learning from underrepresented samples, the AI algorithms show low accuracy for the underrepresented classes. The solution that allows AI applications to imbalanced datasets usually comprises three procedures: balancing the data set, identifying data subsets exhibiting similar patterns, and identifying inputs that lead to better accuracy at predicting most (or a subset) of the target events. This work provides a short overview of the most common approaches and methods for each procedure.

After evaluating the methods for data preprocessing, we employed selected approaches to predict a rock mass class ahead of the tunnel face. For the application, we used available geological information and seismic datasets from Gotthard Base Tunnel (GBT), Switzerland, and Rogfast Tunnel (ROG), Norway, provided by Amberg Technologies AG.

S0506. Decision Support Based on Inventory Data

Vaibhav Shringi (Graz University of Technology, Institute of Rock Mechanics and Tunnelling, Graz,),
Paul Johannes Unterlass (Institute of Rock Mechanics and Tunnelling, Graz University of Technology),
Alla Sapronova (Graz University of Technology),
Thomas Marcher (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)
Room: R8 2023-08-08 14:00

The decisions on the tunnel support in a New Austrian Tunneling Method (conventional drill and blast excavation) for the approaching advance length (round length) are made based upon expert opinions. These opinions depend on the conditions observed on site, tunnel deformations, stresses reached in the lining, etc. All the measurements obtained on-site are recorded and stored for future purposes. Usually, the records created in the early- to mid-digitalization era, are stored as papers or files containing unstructured information (e.g., scanned documents). Such documents may include, e.g., handwritten notes, engineering sketches, printed plots, and photos. Therefore these records cannot be processed easily by a computer or a human.

This paper will discuss a project that aims to find a way to digitalize the valuable data from past engineering projects stored in papers. Even though every tunneling project is different regarding geomechanical boundary conditions, it can be observed that the decisions are often 'reinvented' on site, and there were similar project conditions in the past. Obtaining additional records would help automatize the decision-making process, by searching for a past tunneling project under a similar constellation of observed conditions. Upon successfully digitalizing the data, it would present us with a vast amount of data for already completed tunneling projects and the excavation decisions made against its conditions. Analytical approaches aiming to predict rare events might get extra valuable data for models' validation as extra documents with historical observations will become available in an analysis-ready structural form.

With access to the proposed digitized database of archived projects, it may also be possible to answer the question of whether the wealth of experience of a nominated expert can be aided by automatized scripts and learning models.

S0508. Rare event detection with Autoencoders: application in civil engineering

Alla Sapronova (Graz University of Technology),
Paul Johannes Unterlass (Institute of Rock Mechanics and Tunnelling, Graz University of Technology),
Thomas Marcher (Institute of Rock Mechanics and Tunnelling, Graz University of Technology)
Room: R8 2023-08-08 14:20

Rare events can significantly impact civil engineering, making models for rare event prediction critical for risk management. Unsupervised methods for rare event detection, such as autoencoders, are becoming increasingly popular due to their ability to identify patterns and anomalies without prior knowledge or training data.

DNN autoencoders are characterized by having multiple hidden layers, allowing them to capture high-level abstract features from large and complex datasets. They are well-suited for applications where a high level of complexity is expected and can provide accurate predictions when trained properly. However, their complexity can also make them computationally expensive and difficult to train.

Non-DNN autoencoders, on the other hand, have a simpler architecture with only one or two hidden layers. They are better suited for smaller datasets where the complexity of DNN autoencoders is unnecessary. Non-DNN autoencoders can still be effective at encoding and decoding data, but they may not be able to capture the same level of complexity as DNN autoencoders.

In this work, we apply autoencoders to rare event detection in TBM (Tunnel Boring Machines) and MWD (measurement while drilling) datasets. By leveraging the inherent features in the data, autoencoders can detect events that deviate significantly from the norm and provide early warning signals for potential failures or damages. The choice of which type of autoencoder to use on a specific application and dataset, as well as the desired level of prediction accuracy, are discussed.

S06 Advances in computational analysis of geochemical survey data for geochemical anomaly detection

Behnam Sadeghi (CSIRO),
Peter Atkinson (Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YR,),
Jennifer McKinley (Queen's University Belfast),
Mark Cooper (Geological Survey of Northern Ireland, Dundonald House, Belfast, UK)

Computational analysis has been developed in various fields of geosciences, especially mineral exploration and environmental studies. Such analyses, including data exploration and visualization, geostatistics, uncertainty modeling, and machine learning (ML) methods, integrated with remote sensing, and geochemical data, can provide greater identification and understanding of potential geochemical anomalies or contaminations on, or under, the ground. In this session, we welcome abstracts related to geochemical anomaly detection and characterization with remote (and, thus, a large number) geochemical surveys covering the following topics: (1) spatial resolution choice, sampling design, image analysis, deep learning, higher-order representations, spatial object representations, (2) mathematical/statistical anomaly classification methods to define geochemical anomalies in 2D to 5D such as multifractal modeling, ML, simulations, and spatial uncertainty evaluations, and (3) any relevant applicable data analysis, particularly geospatial and compositional data analysis (CoDa). The main focus of this session is to integrate the mentioned methods to come up with highly efficient and reliable representative models and exploration decisions with the lowest possible uncertainties and risks.

S0601. Spatial outlier detection using the spatially smoothed MRCD

Patricia Puchhammer (TU Wien),
Peter Filzmoser (Institute of Statistics and Mathematical Methods in Economics, TU Wien)
Room: R5 2023-08-07 14:00

Many methods are available for multivariate outlier detection but until now only a hand full are developed for spatial data where there might be observations differing from their neighbors, so-called local outliers. There are methods based on a pairwise Mahalanobis distance approach, however the type of the covariance matrices used is not yet agreed upon and covers only a global covariance (Filzmoser et al. (2013)) and a very local covariance structure (Ernst and Haesbroeck (2016)), more precisely one covariance estimation per observation.

To bridge the gap between the global and local approach by providing a refined covariance structure we develop spatially smoothed robust and regular covariance matrices based on the MRCD estimator (Boudt et al. (2020)) for pre-defined neighborhoods. As well known from the MCD literature, a subset of observations, the so-called H-set, is obtained by optimizing an objective function. In our case we obtain a set of optimal H-sets from minimizing an objective function which is based on a linear spatial smoothing design of local covariance matrices.

A heuristic algorithm based on the notion of a C-step is developed to find the optimal set of H-sets which also shows stable convergence properties in general. We demonstrate the applicability of the new covariance estimators and the importance of a compromise between locality and globality for local outlier detection with simulated and real world data including measurements of Austrian weather stations and further compare the performance with other state-of-the-art methods from statistics and machine learning.

This project has received funding by the European Commission within the Horizon 2021 programme under grant agreement ID 101057741.

References.

Filzmoser, P., Ruiz-Gazen, A., and Thomas-Agnan, C. (2013). Identification of local multivariate outliers. Statistical Papers, 55, 29–47.

Boudt, K., Rousseeuw, P. J., Vanduffel, S., and Verdonck, T. (2020). The minimum regularized covariance determinant estimator. Statistics and Computing, 30, 113–128.

Ernst, M. and Haesbroeck, G. (2016). Comparison of local outlier detection techniques in spatial multivariate data. Data Mining and Knowledge Discovery, 31, 371–399.

S0602. 3D U-Spatial Statistics applet: A Case Study for Epithermal Gold Deposit in Turkey

Gunes Ertunc (Hacettepe University),
Ali Imer (Department of Geological Engineering, Middle East Technical University)
Room: R5 2023-08-07 14:20

Geochemical anomalies detected around mineral deposits are linked various to ore-forming processes, and these anomalies play an important role in mineral exploration. The term background is equivalent with the absence of an anomaly. In order to evaluate the anomalies, geochemical background value should be assessed carefully.

There are two major methods used for assessing background concentrations: (1) Geochemical and (2) Statistical. The first method, also known as empirical method, refers to studies of samples not affected by industrial activities or to relatively sterile sites. These empirical studies are challenged by subjective sample selection criteria, high costs, and heavy laboratory workload. The second method, statistical methods, have been used not only for assessing background concentrations, but also for the separation of geochemical anomalies from geochemical background. Majority of the statistical approaches focus on outlier detection and omit that are considered to be anthropogenically influenced.

This study focuses on the statistical approaches for geochemical anomaly detection. Not limited to the comprehensive exploratory data analysis tool and outlier detection, the applet compiled by authors offers 3D U-spatial statistics application. In 2D mode, target area suggestion by indicator kriging is also available. This program is tested in one of the epithermal gold deposits located in Turkey, and outputs are presented in detail.

S0603. Soil geochemistry data analysis of hydrogen and other gases along the San Andreas Fault

Yashee Mathur (Stanford University),
Victor Awosiji (Stanford University),
Tapan Mukerji (Stanford University),
Kenneth E. Peters (Stanford University)
Room: R5 2023-08-07 14:40

Natural hydrogen has generated great interest as a potential clean and renewable energy source. To understand the occurrence of natural hydrogen due to serpentinization and faulting, 103 shallow soil gas samples 1-meter deep were acquired near the San Andreas Fault in Jasper Ridge and Portola Valley, California. The gas samples were analyzed for concentrations of hydrogen, helium, carbon dioxide, light hydrocarbons, and fixed gases. Gridding and contouring the hydrogen gas and carbon dioxide shows three concentration areas in the Jasper Ridge region which overlie faults. Statistical analysis, hierarchical clustering (HCA), principal component analysis (PCA), and alternating least squares analysis (ALS) were used to group samples, reveal their spatial distribution, and understand possible sources of the gases.

We observed high concentrations of hydrogen up to 20.32 ppmv and 17.3 ppmv in Jasper Ridge and Portola Valley respectively, which are approximately 30-35 times greater than the atmospheric concentration of hydrogen which is close to 0.55 ppmv. We also encountered outlier samples with anomalous quantities of methane and ethane. We performed HCA and PCA to establish five clusters in Jasper Ridge and six clusters in Portola Valley. Most of the samples having high hydrogen concentrations fall on or near faults. Elevated concentration of carbon dioxide has resulted from aerobic microbial degradation of organic activity and elevated concentrations of light hydrocarbons likely resulted from thermal cracking of organic matter. There is also the possibility of fault-related gases. Repeatability analysis shows an initial increase in soil hydrogen values and eventual equilibrium after 24 hours likely due to perturbation in the soil owing to disturbances during sample collection.

S0605. The off-axis plume–ridge interaction model: Confirmation from the mineral chemistry of Cretaceous basalts of the Ontong Java Plateau

Jing Chen (Sun Yat-sen University),
Shuang-Shuang Chen (Sun Yat-sen University),
Donald B. Dingwell (Ludwig- Maximilians- Universität),
Qiuming Cheng (China University of Geosciences/Sun Yat-sen University)
Room: R5 2023-08-07 15:00

The Ontong Java Plateau (OJP) is the largest known igneous province on Earth. The proposal that its origins involved off-axis plume–ridge interactions remains highly controversial. Here we present new in-situ major (EPMA), trace element (LA-ICP-MS) and Sr isotopic (LA-MC-ICP-MS) characterization of clinopyroxene and plagioclase phases present in OJP tholeiitic basalts from Sites 1183, 1186, 1185, 1187. Site 1187 and Upper Site 1185 have yielded olivine tholeiites containing olivine, Ca-rich augite, Ca-deficient pigeonite, and plagioclase, inferred to be derived from higher degrees of partial melting of dominantly peridotitic mantle and reflecting relatively low crystallization temperatures. Sites 1183 and 1186 and Lower Site 1185 yielded tholeiitic basalts consisting mainly of plagioclase and Ca-rich augite without olivine, interpreted to originate from lower degrees of partial melting of pyroxenite- and peridotite-derived melts with relatively high clinopyroxene and plagioclase crystallization temperatures. Site 1187 and Upper Site 1185 plagioclase, clinopyroxene, and equilibrium melts have much lower REE contents but slightly higher in-situ Sr isotopic ratios compared to those phases at Sites 1183 and 1186 as well as Lower Site 1185. Sites 1186 and 1187 plagioclases and clinopyroxenes display normal zoned textures which are attributed to the process of fractional crystallization and extensive magma mixing beneath the heterogeneous mantle source. In-situ Sr isotopic disequilibrium between plagioclase phenocrysts and plagioclase clusters further indicates an enriched compositional heterogeneity. Due to the strong tensile forces expected at the spreading ridge, the OJP off-axis mantle plume is inferred to be drawn towards the ridge axis, leading to shifts in mineral compositions, degrees of partial melting, and crystallization temperatures inferred for the volcanics of Sites 1183, 1186, 1185, and 1187 as a function of distance from the spreading ridge. The enriched invariability of in-situ Sr isotopic compositions is likely due to the relatively long half-life of the radiogenic elements. These mineral chemistries are consistent with off-axis ridge–plume interaction.

S07 Practical Aspects and Applied Studies in Geochemical Exploration and Mapping with Compositional Logratio Techniques

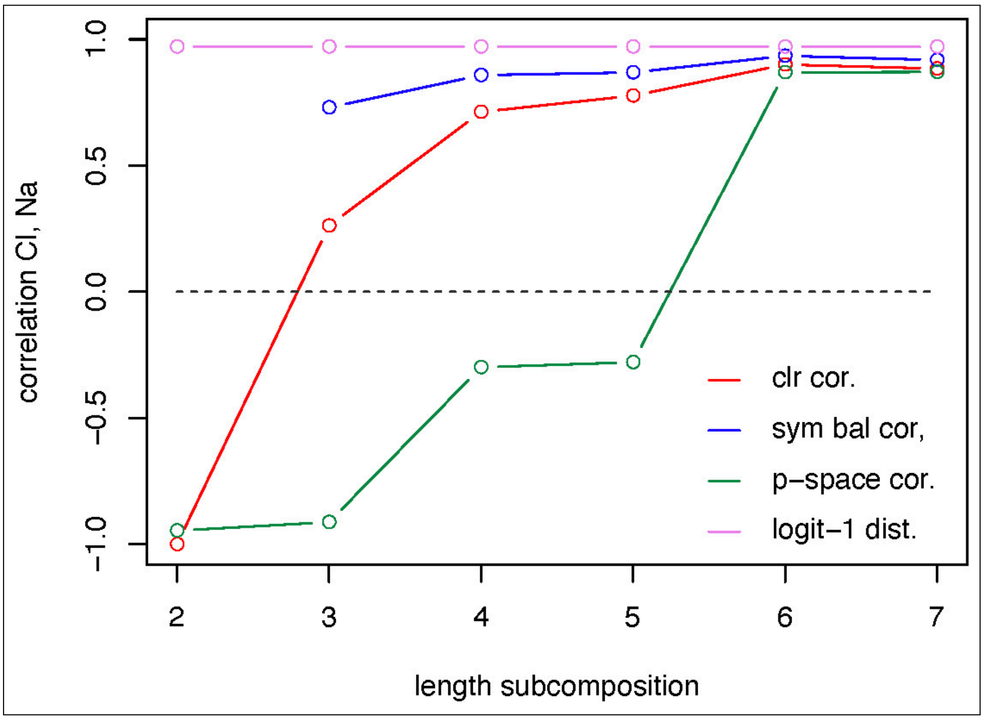
Jennifer McKinley (Queen's University Belfast),
Karel Hron (Palacky University, Olomouc)

This session will offer a practical forum of discussion for people concerned with the statistical treatment, modelling and interpolation of compositional data in geochemical applications, particularly focused on geochemical exploration and mapping. The session will invite papers related to the problems of geochemical mapping, and the compositional topics. Geochemical survey data have both compositional and spatial character. Thus, there are two spaces for these data; the geochemical and the geospatial domains. The multivariate relationships of the elements typically reveal patterns associated with processes determined by mineralogy, weathering, comminution and mass transport. Typically, these processes display a continuous spatial character in both 2 and 3 dimensions and time. Many methods are available for process discovery such as principal components or minimum/maximum autocorrelation factor analysis. The evaluation of geochemical data can be carried out using the approach of process discovery and process validation. In the context of Process Discovery, different methods use different measures of association. Some multi-element associations are common amongst all of the metrics, whilst other measures are unique to the metric. These metrics use both linear and non-linear mappings to define the associations. Typically, the approach is unsupervised with the intent to define multi-element associations that reflect mineralogy associated with rock types. When the procedure of process discovery identifies patterns that represent geologic phenomena, these patterns can be tested using classification methods and is termed process validation. Methods such as discriminant analysis operate in linear and non-linear modes to carry out prediction. Neural networks, logistic regression, random forests, support vector machines use various measures of association for classification using heuristics such as decision trees and multivariate measures of distance. This approach is usually described as a supervised way of testing the classes defined through process discovery or by pre-defined categorical information assigned with each geochemical composition. This session invites contributions where multivariate geoscience data (geochemistry, geophysics, geodetic, geologic) can be processed within a compositional framework that includes an evaluation of the data, or results of processing the data, in a geospatial context.

S0701. A subcompositionally coherent proportionality index of parts

Vera Pawlowsky-Glahn (University of Girona),
Juan Jose Egozcue (Technical University of Catalonia)
Room: R8 2023-08-09 14:00

An open challenge in Compositional Data Analysis is the measurement of the strength of dependence of parts in a composition. To find a subcompositionally coherent alternative to the usual correlation coefficient, recognised as spurious by Karl Pearson at the end of the 19th century, has shown to be a difficult task with the usual tools. Understanding that spurious results are due to a semantic incoherence is a first step towards establishing an alternative to relate compositional parts in a statistically consistent way. The second step is to define a coherent system of functions with respect to a subcomposition and to analyse the space of parts. This leads to understand why measures like covariance and correlation depend on the subcomposition considered, while measures like distance between parts are independent of the same. These steps allow finally to define a novel index of proportionality between parts which depends only on the parts of interest and not on the rest of the composition or on the number of parts in the composition.



S0702. Application of compositional data analysis to biogeochemical exploration for disclosing the lithology-derived signatures in plants from its physiological artifacts

Lucija Dujmović (HZDR, HIF),
Solveig Pospiech (HZDR),
Raimon Tolosana-Delgado (Helmholtz Institute Freiberg),
K. Gerald van den Boogaart (Helmholtz Institute Freiberg, Helmholtz-Zentrum Dresden-Rossendorf Modelling and),
Jörg Matschullat (Interdisciplinary Environmental Research Centre, TU Bergakademie Freiberg),
Maarit Middleton (Geological Survey of Finland GTK)
Room: R8 2023-08-09 14:20

In the early stages of mineral exploration, biogeochemical data of plants can provide a cost-effective and efficient insight into the underlying lithology. Plants accumulate certain element groups differently in their tissue types - specific to the physiological function of plant organs. This leads to substantial variability in elements of interest between tissue types (needles, twigs, bark) of plant species, not only in terms of their concentration ranges but also in terms of their element ratios. This work aims to investigate the multivariate relative relationships of elements and element groups to reveal patterns in biogeochemical data related to bedrock lithology and geochemistry. The ultimate goal would be to discover element ratios such that the contribution of plant species and tissue types (noise) could be minimized by appropriate normalization of the plant data.

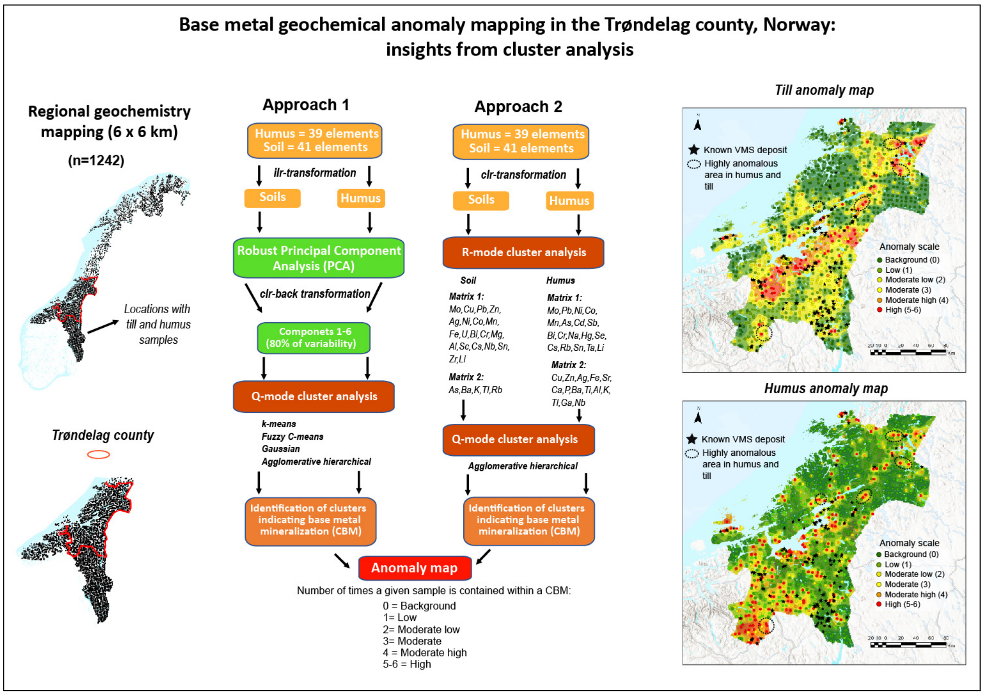
To reach the aim of the study we investigate plant geochemical data by using compositional data analysis techniques. In this case study, a biogeochemical data set collected in Northern Finland was used. The data consist of geochemical analyses of several species acquired on top of mineralized rocks of several types and their surrounding barren rocks. The biogeochemical data were normalized via perturbation by the inverse of the center, i.e., by removing the mean. The center was calculated by taking the compositional mean of the samples collected over the barren sites for each species and tissue type group to represent the baseline contribution by the physiological effects of a species/tissue type. For each sample, the centering was applied according to its species and tissue type separately by applying the corresponding mean. Application examples for the biogeochemical data show that the approach is successful in removing the plant's inherent signature so that only the remnant variability of the pattern is left for interpretation. The results of the data exploratory methods show that the variability after the normalization corresponds to the geological features such as underlying mineral deposits. This approach helps to reveal biogeochemical signatures originating from bedrock geology.

S0703. Base metal geochemical anomaly mapping in the Trøndelag county, Norway: insights from cluster analysis

Pedro Acosta-Gongora (NGU),
Y Wang (Geological Survey of Norway),
C Haase (Geological Survey of Norway),
K Saalman (Geological Survey of Norway),
I Henderson (Geological Survey of Norway),
T Bjerkgård (Geological Survey of Norway)
Room: R8 2023-08-09 14:40

The Trøndelag county in Norway is host to several Cu-Pb-Zn occurrences (e.g., Røros) associated to the emplacement of volcanogenic massive sulfide deposits (VMS). From 2014 to 2017, the Norwegian Geological Survey carried out a regional (6 km x 6km) geochemical mapping (major and trace elements) of till (41 elements) and humus (39 elements) across the entire county (n= 1280). Despite its relatively low sampling scale, the two-media high dimensional datasets offer an excellent opportunity to carry out unsupervised-learning methods to identify areas with high Cu-Zn-Pb potential.

Two approaches were carried out on the till and humus samples separately. The first approach consists of the implementation of robust principal component (PCA) analysis on isometric log-transformed data. The scores for the first six principal components (80% variability) were inputted as variables for Q-mode clustering (i.e, clustering of the samples) using k-means, fuzzy C-means, agglomerative hierarchical (agnes), and gaussian mixture model algorithms. For the second approach, the data were center-log transformed and R-mode clustering (i.e., clustering of the variables) was carried out for each dataset. Relevant groups of variables (herein matrices) consisting of elements signalling base metal mineralization were identified and separately clustered in Q-mode using agnes. In both approaches, the optimal number of clusters when performing Q-mode clustering was set at three, where clusters consistent with base metal mineralization (CBM) were identified for every clustering iteration. In total, each soil and humus sample were classified using 6 different clustering iterations. Samples that were not classified as CBM in any of these iterations are considered background (value of 0), whereas those labelled as CBM more than 4 times (values of 5-6) were deemed as the most anomalous ones. In general, the most anomalous soil samples have a better spatial correlation with known VMS deposits than humus samples. However, it was noted that in SE-Trøndelag (e.g., Løkken and Røros districts), samples from both media achieve moderate to low anomaly values indicating the presence of a distinct and most likely underrepresented (in the data) VMS mineralization type. Despite the later, we suggest that common prospective areas identified in both media should be then considered as the best exploration targets.


Base metal geochemical anomaly mapping in the Trøndelag county, Norway:
insights from cluster analysis
Regional geochemistry mapping (6 x 6 km)
(n=1242)
Locations with till and humus samples
Trøndelag county
Approach 1
Humus = 39 elements
Soil = 41 elements
ilr-transformation
Soils
Humus
Robust Principal Component Analysis (PCA)
clr-back transformation
Componets 1-6
(80% of variability)
Q-mode cluster analysis
k-means
Fuzzy C-means
Gaussian
Agglomerative hierarchical
Identification of clusters indicating base metal mineralization (CBM)
Approach 2
Humus = 39 elements
Soil = 41 elements
clr-transformation
Soils
Humus
R-mode cluster analysis
Soil
Matrix 1:
Mo,Cu,Pb,Zn,
Ag,Ni,Co,Mn,
Fe,U,Bi,Cr,Mg,
Al,Sc,Cs,Nb,Sn,
Zr,Li
Matrix 2:
As,Ba,K,Tl,Rb
Humus
Matrix 1:
Mo,Pb,Ni,Co,
Mn,As,Cd,Sb,
Bi,Cr,Na,Hg,Se,
Cs,Rb,Sn,Ta,Li
Matrix 2:
Cu,Zn,Ag,Fe,Sr,
Ca,P,Ba,Ti,Al,K,
Tl,Ga,Nb
Q-mode cluster analysis
Agglomerative hierarchical
Identification of clusters indicating base metal mineralization (CBM)
Anomaly map
Number of times a given sample is contained within a CBM:
0 = Background
1= Low
2= Moderate low
3= Moderate
4 = Moderate high
5-6 = High
Till anomaly map
Known VMS deposit
Highly anomalous area in humus and till
Anomaly scale
Background (0)
Low (1)
Moderate low (2)
Moderate (3)
Moderate high (4)
High (5-6)
Humus anomaly map
Known VMS deposit
Highly anomalous area in humus and till
Anomaly scale
Background (0)
Low (1)
Moderate low (2)
Moderate (3)
Moderate high (4)
High (5-6)

S0704. Compositionally relevant post hoc tests of compositional linear models

Solveig Pospiech (HZDR),
K. Gerald van den Boogaart (HZDR, HIF)
Room: R8 2023-08-09 15:00

The geochemistry of surficial Earth materials, e.g. (weathered) rocks, soils, plants, peatlands, waterbodies or snow, is influenced by many processes and environmental parameters. Often, however, it is unknown which of the observed environmental variables, like soil pH, precipitation, slope, underlying bedrock type, to name but a few, have a significant influence on the geochemical pattern of the respective material. For this purpose, one can test hypotheses about the influence of variables on the (ilr or alr transformed) composition, e.g. by means of linear regression. These tests can either be rejected or accepted.

If the result of the statistical analysis shows that the variable(s) tested could have an influence on the overall composition of the surficial material, the question immediately arises for geochemists as to which of the elements or element groups change in relation to this variable. In the case of non-compositional and univariate data, there are the classical post-hoc tests, e.g. tests according to Scheffé or Bonferroni, that allow to check which of certain interpretable subhypothesis are responsible for significant result. However, as far as the authors are aware, there are no such tests developed for multivariate situations and geochemically or compositionally relevant sub-hypotheses, such as those that consider sub-compositions or balances of specific elements.

The contribution provides a systematic compilation of compositionally and geochemically potentially relevant subhypotheses and corresponding post-hoc tests. A particular challenge is that the number of compositional hypotheses is much larger than the number of subhypotheses (all subcompositions and balances) than in a contrast's situation (all pairs). A Bonferroni approach is therefore impractical. We thus extend Scheffé's principle of post-hoc testing to the situation where any number of additional subhypotheses can be tested simultaneously without the need for additional p-value corrections.

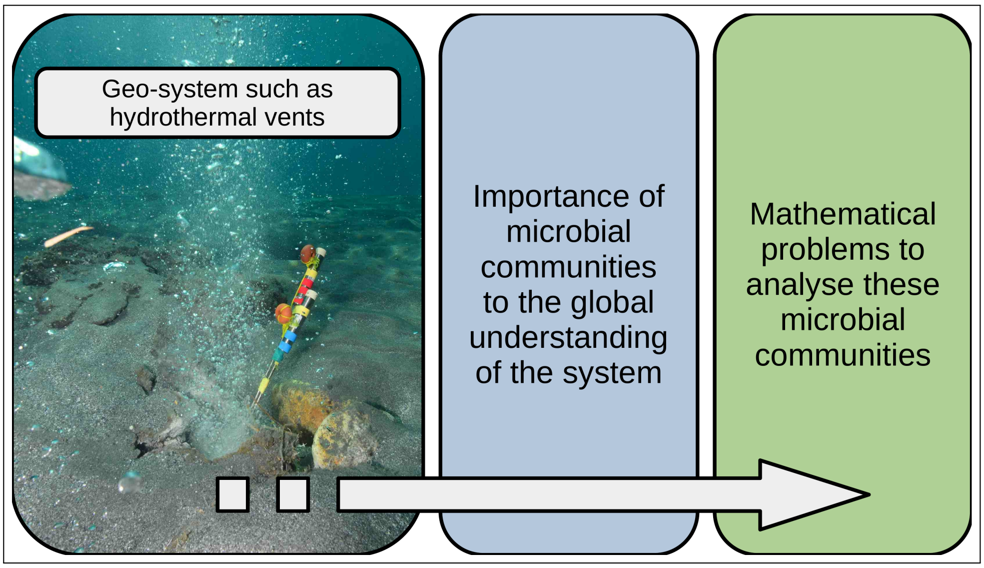
Geochemists in this system of post-hoc tests can then either test geochemically interpretable relevant sub-hypotheses or test the model for the least explanatory subhypotheses and interpret these results based on a hierarchy of implied hypotheses. The application of the method will be demonstrated on a snow data set, which was collected for exploration purposes on a Co-Au deposit.

S0705. Geomicrobiological study of a hydrothermal vent field using compositional methods

Sven Le Moine Bauer (University of Bergen - Center for Deep Sea Research),
Vera Pawlowsky-Glahn (University of Girona),
Juan José Egozcue (Technical University of Catalonia),
Steffen Jørgensen (Center for Deep Sea Research, Department of Earth Science, University of Bergen,)
Room: R8 2023-08-10 10:00

The past decades have seen a strong development of earth system thinking, an inherently multidisciplinary take on geo-sciences aimed at understanding interactions between various components of Earth. In this regard, investigating the biosphere has been shown to be fundamental to the understanding of many systems, as microbes are involved in virtually all geochemical cycles, therefore regulating local and global processes and fluxes. Nevertheless, the microbiological component is rarely taken into account in geo-studies. Therefore, this presentation will first describe the benefits of implementing microbiological analyses in geosciences.

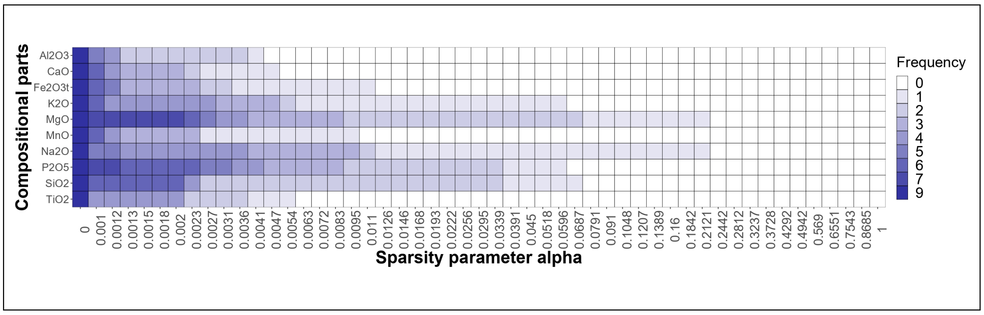
Then, focus will be given to our recent geomicrobiological study of the shallow hydrothermal vent field of Paleochori Bay (Milos, Greece) where we acknowledge the compositionality of the data produced through the sequencing analysis of microbial communities. While the use of compositional data analysis methods in microbiology has been given and increasing attention in the past decade, the current tools are mainly aimed at tackling interrogations from medical studies. Therefore, the conclusion will present the typical research questions fundamental to geomicrobiology and define the challenges and needs related to compositional data analysis tools for this field.



S0706. Identifying important pairwise logratios in geochemical data with sparse principal component analysis

Karel Hron (Palacky University, Olomouc),
Viktorie Nesrstová (Palacky University),
Ines Wilms (Maastricht University),
Peter Filzmoser (Institute of Statistics and Mathematical Methods in Economics, TU Wien)
Room: R8 2023-08-10 10:20

Compositional data are characterized by the fact that their elemental information is contained in simple pairwise logratios of the parts that constitute the composition. While pairwise logratios are typically easy to interpret, the number of possible pairs to consider quickly becomes (too) large even for medium-sized compositions, which might hinder interpretability in further multivariate analyses. Sparse principal component analysis (PCA) is therefore an appealing dimension reduction technique to identify few, important pairwise logratios (respectively parts contained in them) from the total candidate set. As PCA on centred logratios is consistent with a PCA based on all pairwise logratios, sparse PCA is applied to the possibly high-dimensional matrix of all pairwise logratios, and the L1 penalty serves as a tradeoff between maximizing explained variance and reducing the set of pairwise logratios. The performance of the procedure is demonstrated both with simulated and empirical geochemical data. In our empirical analyses, we propose three visualization tools showing (i) the balance between sparsity and explained variability, (ii) stability of the pairwise logratios, and (iii) importance of the original compositional parts to aid practitioners with their model interpretation.

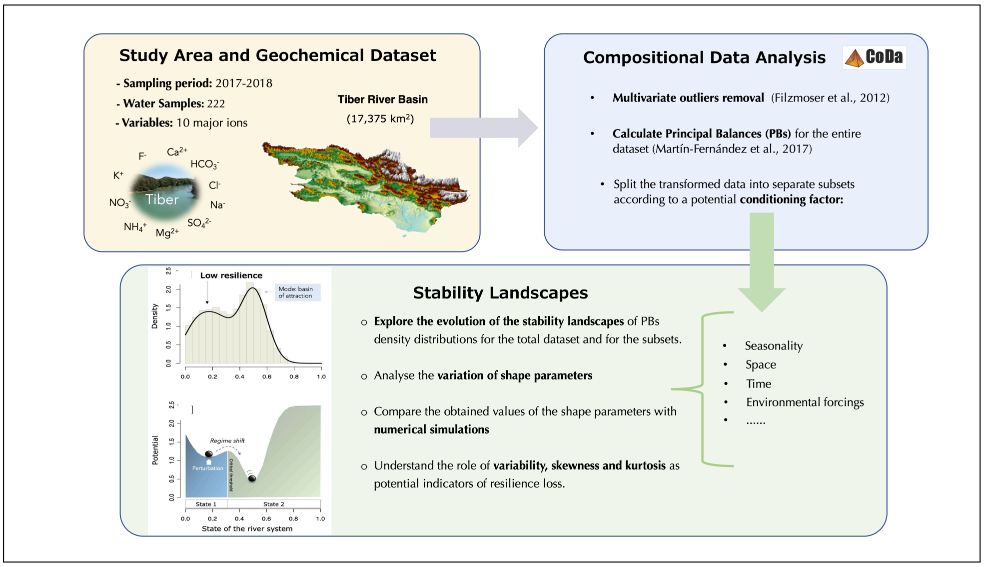


S0707. River chemistry behaviour in response to perturbations revealed by density distributions of Principal Balances

Caterina Gozzi (University of Florence),
Antonella Buccianti ((1) University of Florence, Department of Earth Sciences, Firenze, Italy; (2) NB)
Room: R8 2023-08-10 10:40

Density distributions can be used to infer a potential landscape describing the stability properties of a system. The reversed density can approximate the shape of a stability landscape with one or multiple domains of attraction (alternative states) depending on the number of modes. Hence, the variation in the shape of a basin of attraction (i.e. slope, depth and width) can reflect system changes over time/space under the effect of environmental perturbations. Variability, skewness and kurtosis describe the shape of a distribution. However, their role as potential early warning indicators to identify the loss of resilience in a dynamic system is still debated. This study investigates the possibility of quantifying the resilience to changes of river water chemistry by analyzing the shape of the density distributions of Principal Balances (PBs) calculated using Compositional Data Analysis.

The proposed approach was tested in different hydro-geological and morphological contexts of the Tiber River Basin in central Italy. PBs were calculated for a geochemical dataset of 222 water samples, resulting in balances dominated by anthropic- and geological-driven processes following the decreasing order of variability. The Inverse Distance Weighting interpolation method was used as a descriptive tool to visualize the distribution of each PBs at a catchment-wide scale and support a geochemical interpretation. Successively, the transformed data were split into subsets in order to highlight seasonal and spatial changes within the subbasins. The evolution of the measures describing the shape of each distribution was analyzed for the different subsets and the results were compared with numerical simulations. The outcomes suggest that variability seems to reveal the most favorable response of Pbs to perturbations, i.e. flickering between alternative states or, alternatively, unexpected critical transitions. These results coupled with the analysis of skewness and kurtosis for the different seasons reveal that drought periods can most likely favor regime shifts compared to flood ones. This evidence might have significant implications when quantifying watersheds' evolution to existing climatic and human pressures and the effect of these variations on biodiversity.


Study Area and Geochemical Dataset
- Sampling period: 2017-2018
- Water Samples: 222
- Variables: 10 major ions
Tiber River Basin
(17,375 km²)
Tiber
Compositional Data Analysis
CoDa
Multivariate outliers removal (Filzmoser et al., 2012)
Calculate Principal Balances (PBs) for the entire dataset (Martín-Fernández et al., 2017)
Split the transformed data into separate subsets according to a potential conditioning factor:
Stability Landscapes
Explore the evolution of the stability landscapes of PBs density distributions for the total dataset and for the subsets.
Analyse the variation of shape parameters
Compare the obtained values of the shape parameters with numerical simulations
Understand the role of variability, skewness and kurtosis as potential indicators of resilience loss.
Seasonality
Space
Time
Environmental forcings
......
Low resilience
Density
Potential
Regime shift
State 1
State 2
State of the river system

S0708. Tectono-geochemical anomaly identification and concealed ore deposit prediction based on compositional balance analysis

Yue Liu (China University of Geosciences)
Room: R8 2023-08-10 11:00

Tectono-geochemical exploration always involves the study of tectono-geochemical primary halos. In the present study, compositional balance analysis (CoBA) was introduced to recognize tectono-geochemical anomalies and infer blind ore deposits in the deep based on superimposed characteristics of tectono-geochemical primary halos. A case study for concealed Pb-Zn exploration based on tectono-geochemical data in the Huangzhuguan ore cluster, Gansu Province (China), was employed to demonstrate the potential application of the CoBA method. First, an optimal element combination (i.e., Pb-Zn -Cd) related to near-ore halo was determined by means of CoBA method and receiver operating characteristic curve analysis. Then, according to the element zonation sequences of primary halo of hydrothermal deposit, the front halo element combination (i.e., As-Sb-Hg) and the tail halo element combination (i.e., W-Sn-Mo-Bi) were selected for compositional balance analysis. Finally, the superimposed characteristics of three tectono-geochemical primary halos were interpreted for blind Pb-Zn deposit exploration in the Huangzhuguan ore cluster.

S0709. Mapping Potential Targets for Gold Mineralization Based on Score Balances of Soil Geochemical Data in the Amapari Region, Amapá State, Brazil

Luis Braga (Federal University of Rio de Janeiro),
claudio porto (FederalUniversity of Rio de Janeiro),
Joao Casado (Federal University of Rio de Janeiro)
Room: R8 2023-08-10 11:20

Geochemical exploration in deeply weathered tropical terrains is often hampered by a strong modification of the primary geochemical signals, on the surface, making it difficult to locate potential targets for gold mineralization relying on geochemical assays for Au in soil samples. Moreover, geochemical results for gold in soil samples maybe misleading due to the difficulty in obtaining representative samples coupled with analytical problems and unstructured variability such as the nugget effect. This problem is particularly relevant in the Amapari region, Amapá state, Brazil where humid tropical conditions contributed to develop regolith profiles that may be several tens of meters thick topped by lateritic crusts or highly leached colluvial soils. Therefore, it is important to examine carefully the metal distribution near the potential gold target in order to extract geochemical vectors to Au mineralization. In this study we utilize an innovative methodology based on score balances and data mining.

The aim is to select potential gold targets within a gridded area with a favourable metavolcano-sedimentary geological setting. Although no data mining package is used, the essential steps of the SEMMA (Sample, Explore, Modify, Model, Access) paradigm are followed in this work. The data is composed of a vast soil geochemical survey conducted in the region. A training step based on mineralized and non mineralized samples is essential to select the best variables to be adopted in a balance formula whose score will be calculated over a validation area. The mathematical background is given by the compositional analysis theory. Most calculations were done with functions of the package “Compositions” in and “Selbal”. The training samples were used to generate a numerical model to be used for discriminating between mineralized and non mineralized multivariate geochemical patterns.

The proposed method was successfully applied to highlight a mineralized trend in the Amapari region, which was not previously detected considering only Au grade or any other potential pathfinder metal alone. Given similar geological conditions of very thick overburden coupled with the lack of a distinct signal by the assayed gold grades, the method may be of use in other regions and even for other valuable metals.

S08 AI-driven mineral prospectivity mapping

Renguang Zuo (China University of Geosciences),
Emmanuel John M. Carranza (University of the Free State),
Yihui Xiong (China University of Geosciences)

Mineral prospectivity mapping as a computer-based approach to delineate targeted areas for a specific type of mineral deposits has changed from being knowledge driven to today's AI-driven. There are increasing applications of artificial intelligence (AI) in mineral exploration. The session welcomes the following submissions: (1) novel AI algorithms and application of AI for processing geological, geochemical, geophysical, remote sensing and drilling data; (2) novel AI algorithms and application of AI for identifying and integrating prospecting features in support of mineral exploration.

S0801. Exploration targeting of conduit-type orthomagmatic Ni-Cu-PGE-Co deposits in northern Finland.

Malcolm Aranha (Oulu mining school),
Shenghong Yang (Oulu mining school),
Elena Kozlovskaya (Oulu mining school),
Pertti Sarala (Oulu mining school),
Hanna Silvennoinen (Sodankylä Geophysical Observatory)
Room: R9 2023-08-07 10:00

Metals such as nickel (Ni), cobalt (Co), and platinum group metals (PGMs) have a wide range of significant uses in modern, green industries. However, their current supply levels are insufficient to meet the projected demand. Therefore, these metals are considered "critical raw materials" and are of strategic importance. Local production of these materials is critical due to their significance in industrial value chains and green sectors. The SEMACRET project is dedicated to promoting sustainable exploration of (critical) raw materials within the EU to ensure a consistent supply during the green energy transition.

This article describes a computer-based prospectivity analysis of conduit-type, sulphide-rich Ni-Cu-(PGM)-(Co) deposits in northern Finland using Fuzzy Inference Systems (FIS), a supervised, knowledge-based symbolic artificial intelligence technique; and Self Organising Maps (SOM), a neural network-based unsupervised machine learning clustering algorithm.

The supervised approach relies on a generalised conceptual mineral systems model for conduit-type sulphide deposits, which is used to identify the targeting criteria. The major components of the mineral systems include (1) Primitive, mantle-derived, magnesium-rich rocks (ultramafic and mafic rocks), which are the primary source of metal; (2) trans-lithospheric faults and chonoliths that serve as pathways of magma, locally enhanced by varying compressional/extensional tectonic regimes; and (3) dilatational zones of high, fracture-related permeability and localised structures that act as physical traps; along with abundant sulphur-rich crustal rocks that can provide the external sulphur required to form immiscible sulphide melts. The targeting criteria were represented by their spatial proxies in the form of GIS predictor maps, derived using spatial analyses and geoprocessing tools. These predictor maps were integrated in a multi-stage FIS structured on the mineral systems model to obtain the prospectivity map.

To complement the results of the knowledge-driven analysis and eliminate any subjective bias in the manual processing and filtering involved in the generation of predictor maps, SOM was applied to unprocessed gridded geophysical, geochemical and topographic datasets to cluster features representative of the mineralisation processes in an automated and unsupervised manner. The results of the two approaches were comparable.

This work is co-funded by the European Union and UKRI (SEMACRET, GA101057741).

S0802. Automatic Detection of Geological Structures from Map Data

David Oakley (Université de Pau et des Pays de l'Adour),
Thierry Coowar (BRGM),
Christelle Loiselet (Bureau de Recherches Géologiques et Minières),
Jean-Paul Callot (Université de Pau et des Pays de l'Adour)
Room: R9 2023-08-07 10:20

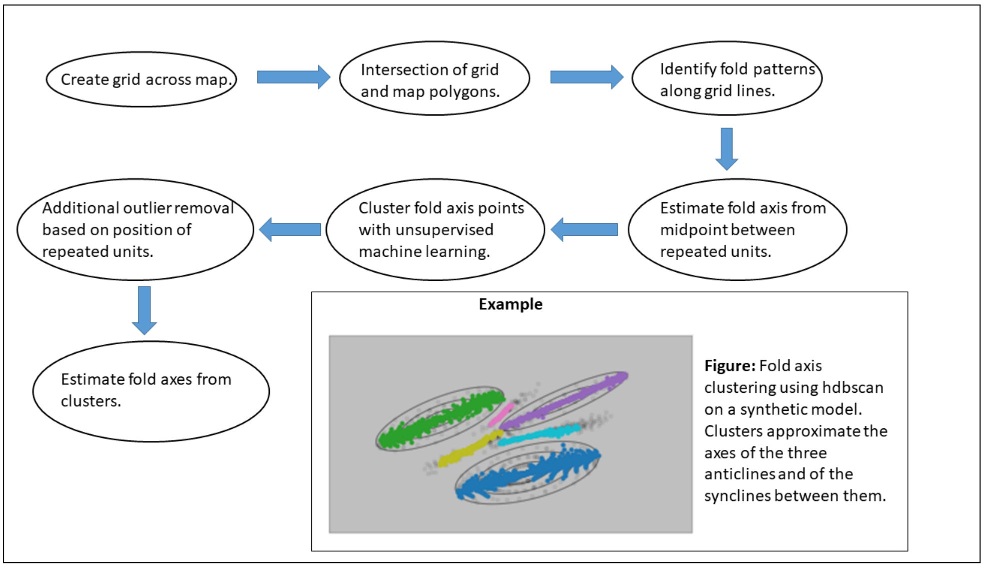
Identification and interpretation of geological structures is essential to subsurface utilization, such as in resource exploration and geological storage of wastes, fuels or energy. Interpretation is a traditionally human process, which is subjective and often time-consuming. Modern advances in data science and machine learning raise the possibility that geological data could be interpreted automatically, providing for efficient and reproducible interpretations of large data volumes. Translating the understanding of a trained geologist to an algorithm that can be implemented by a computer is, however, not straightforward. In this project, we are developing and testing a method to detect geological folds from vector map data using machine learning. Patterns of geological units are first extracted along a grid of lines crossing the map, and map patterns indicative of folds are identified using regular expression matching. Potential fold axes are identified from each match and are clustered together to identify distinct folds in the map using unsupervised machine learning algorithms, for which we consider the dbscan1 and hdbscan2 algorithms as well as a self-organizing map clustering algorithm3. The method has been tested on both synthetic and real geological map data. Initial results show success in identifying folds, but also show some variation depending on the choice of algorithm parameters that control cluster size and spacing. Ongoing work aims to test the method on a wider range of synthetic and natural fold geometries, incorporate additional data to constrain the interpretation, and extend the approach to structures other than folds. A particular emphasis of planned future work is on salt structures, which can present map patterns similar to folds.

References:

1. Pedregosa et al., 2011, Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825-2830.

2. L. McInnes, J. Healy, and S. Astels, 2017. hdbscan: Hierarchical density based clustering. Journal of Open Source Software, The Open Journal 2:11.

3. Gorzałczany, M.B. and Rudzinski, 2018. Generalized self-organizing maps for automatic determination of the number of clusters and their multiprototypes in cluster analysis. IEEE Transactions on Neural Networks and Learning Systems, 29:7, 2833-2845.


Create grid across map.
Intersection of grid and map polygons.
Identify fold patterns along grid lines.
Estimate fold axis from midpoint between repeated units.
Cluster fold axis points with unsupervised machine learning.
Additional outlier removal based on position of repeated units.
Estimate fold axes from clusters.
Example
Figure: Fold axis clustering using hdbscan on a synthetic model. Clusters approximate the axes of the three anticlines and of the synclines between them.

S0803. A hybrid method combines direct sampling and convolutional neural network for lithological mapping based on geochemical survey data

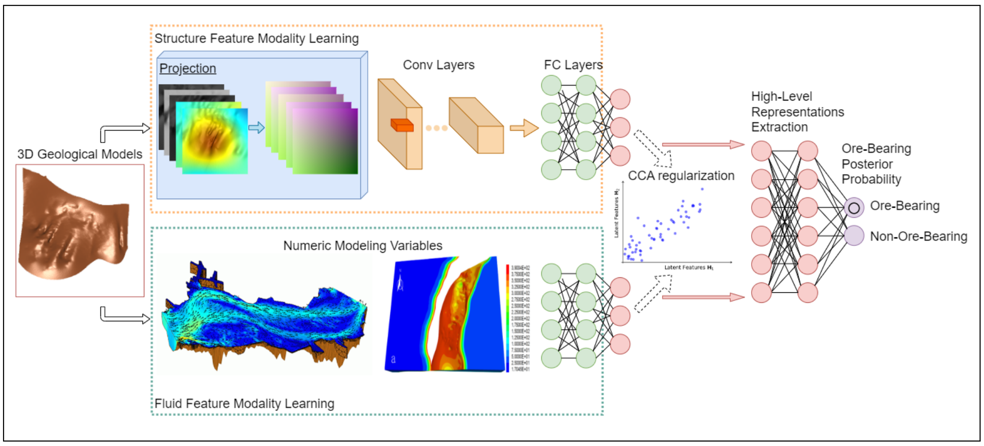
Ziye Wang (State Key Laboratory of Geological Processes and Mineral Resources, CUG),
Renguang Zuo (China University of Geosciences)
Room: R9 2023-08-07 10:40

Lithological mapping refers to the identification, classification, and interpretation of rock types to support Earth observation applications, involving geological surveys and mineral exploration. Numerous studies have demonstrated the potential of geochemical survey data for lithological mapping based on machine learning algorithms. These processes are generally implemented by interpolating discrete geochemical survey point data into raster maps and comparing chemical compositions and spatial distribution to a reference. However, performing lithological mapping based on geochemical survey data faces several challenges. First, geochemical sampling is usually spatially discontinuous, and the limited sampling density cannot adequately support quantitative discrimination of lithologic units. Second, supervised machine learning algorithms are restricted by insufficient observations. Third, most machine learning algorithms are typical pixel-wise models that do not account for the spatial features and correlations of neighboring samples. In this study, a novel hybrid method that combines direct sampling (DS) multiple-point statistical technique and convolutional neural network (CNN) algorithm was proposed to addresses the above challenges. Specifically, the DS technique is designed to produce spatially continuous and sufficient samples by reconstructing unsampled locations from sparse geochemical survey data; and CNN facilitates automatic lithological feature learning and classification based on multi-level convolutional operations that consider the spatial information within neighboring samples. The proposed method is illustrated by a case study mapping lithological units in the Daqiao district, western China, and compared with the deterministic interpolation approach visually and quantitatively. Most lithological units were correctly identified with an overall accuracy of 94%, providing feasible and practical insight into local-scale and high-precision geological mapping using stream sediment geochemical survey data.

S0804. Multimodal Fusion Modeling for 3D Mineral Prospectivity Prediction: Incorporating 3D geological models and dynamic numeric models with Canonical Correlation Analysis

Yang Zheng (School of Earth Sciences and Engineering, Sun Yat-sen University, Zhuhai),
Hao Deng (School of Geosciences and Info-Physics, Central South University, Changsha 41008)
Room: R9 2023-08-07 11:00

Three-dimensional (3D) mineral prospectivity modeling (MPM) involves data of multiple modalities, which poses a significant challenge due to the diverse nature, source and heterogeneity of the data. While state-of-the-art methods integrate these multimodal data and model their associations to mineralization by techniques such as machine learning, fewer concerns are drawn to fuse the gap between the multimodal data and exploit correspondence of multiple modalities for an in-depth understanding of mineral prospectivity. We propose a novel 3D MPM method that learns aligned features of different modalities to predict mineral prospectivity. Two modalities of data, i.e., 3D geological models that represent the 3D geological architecture, and numeric modeling variables that represent ore-forming dynamic conditions, are joined and aligned to indicate mineralization. A deep neural network is devised to effectively combine multiple sources of information and make more comprehensive and robust predictions about the distribution of mineral deposits. Specifically, We integrate 3D geological models and numerical modeling variables to indicate mineralization and maximize the correlation between the two modalities through Canonical Correlation Analysis (CCA) regularization. We evaluate the proposed method using a case study of the world-class Jiaojia gold deposit in east China, demonstrating its effectiveness in accurately predicting mineral prospectivity. Our method outperforms state-of-the-art techniques, highlighting its potential for future mineral exploration and resource management applications.



S0805. Multi-fidelity geological knowledge constrained neural network for mineral prospecting mapping

Zhang Chunjie (China University of Geosciences),
Renguang Zuo (China University of Geosciences)
Room: R9 2023-08-07 11:20

Deep learning algorithms have become increasingly prevalent in the field of mineral prospecting mapping. However, they face several challenges since geological exploration big data are derived from multiple sources linked to different spatial scales, spatial resolutions, and data formats. Furthermore, models lack the necessary geological knowledge constraints, relying solely on data-driven approaches that lead to suboptimal interpretability, generalization, and transferability. To address these challenges, we propose a novel model with two main advantages. Firstly, the multi-fidelity neural network can leverage multiple types of data from various sources, such as satellite images, geological and geophysical data, to extract deep-level and coupled mineral prospecting information, thereby significantly improving the accuracy and efficiency of mineral exploration. Secondly, the spatial distribution law of mineral deposits is integrated into the training process of the model, providing crucial constraints to the model's learning process. By constructing a hybrid model that is driven by both knowledge and data, we can increase the model's interpretability and generalization ability, thereby enabling more effective and efficient mineral prospecting mapping.

S0806. Machine-readable expression of unstructured geological information and intelligent prediction of mineralization associated anomaly areas in Pangxidong District, Western Guangdong, South China

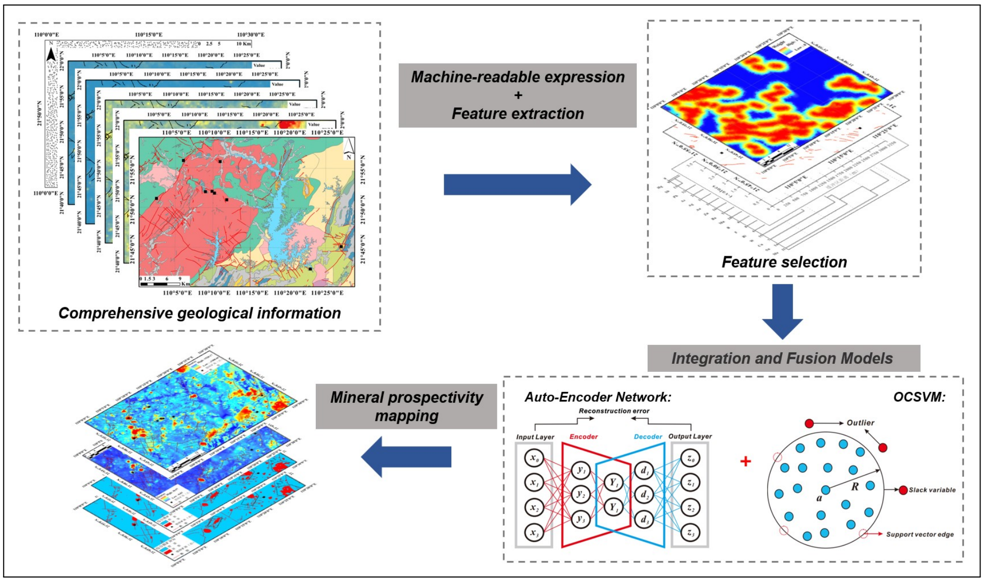
Kunyi Wang (Sun yat-sen University),
Qiuming Cheng (School of Earth Science and Engineering, Sun Yat-sen University),
Yongzhang Zhou (School of Earth Science and Engineering, Sun Yat-sen University)
Room: R9 2023-08-07 11:40

With the deepening of research and the improvement of analysis and testing technology, the data obtained by geological surveys has greatly increased in depth and breadth. For the study of mineralization prediction, the rapid growth of geoscience data has brought new opportunities and challenges. It is an important research trend to apply Big data mining and Machine Learning algorithms to mineralization prediction research, but the stratum, lithology, and faults in the geological data are unstructured data and cannot be directly applied, they need to be processed by machine-readable expression. In this study, the unstructured geological information such as stratum, lithology, and faults of the Pangxidong district in Western Guangdong Province were converted by machine-readable method. It also applies two anomaly detection algorithms such as One-Class Support Vector Machine and Auto-Encode Network to mine geochemical test data of stream sediments and comprehensive geological information such as stratum and fault. The excavated results are used to extract deep prospecting information and ore-forming anomalous characteristics, and finally realize intelligent delineation of ore-forming anomalous regions.

The main conclusions and understanding obtained in this study are as follows:

(i) Through the combined application of One-Hot Encoder and the weighted variable method in spatially weighted principal component analysis, the structural transformation of unstructured geological information such as stratum, lithology, and fault can be effectively achieved. In this way, the geological information it contains can be retained to the greatest extent, and it can be better applied to machine learning algorithm processing.

(ii)Compared with traditional prospecting geochemical methods, the analysis method in this study can process and integrate multi-source geological ore-forming information, and the final ore-forming anomalous recognition results are in good agreement with the spatial distribution of actual deposits. Applying this method to prospecting potential areas without discovered deposits can improve the efficiency of prospecting work and the possibility of discovering deposits.



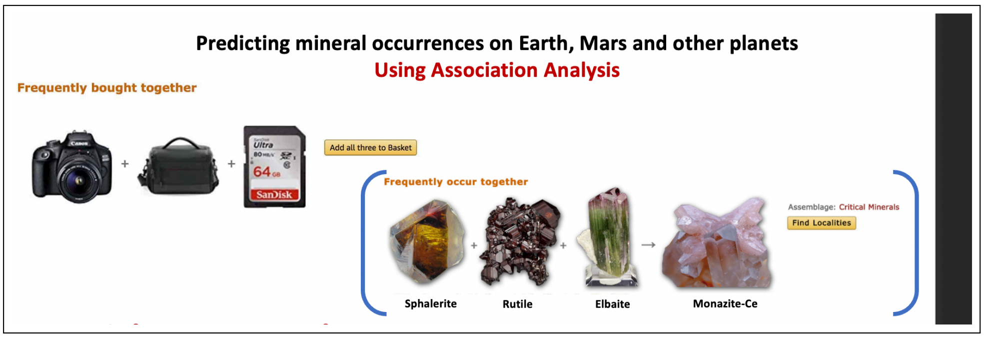
Comprehensive geological information
Machine-readable expression
+
Feature extraction
Feature selection
Integration and Fusion Models
Auto-Encoder Network:
Reconstruction error
Input Layer
Encoder
Decoder
Output Layer
OCSVM:
Outlier
Slack variable
Support vector edge
Mineral prospectivity mapping

S0807. Predicting unknown mineral occurrences using mineral association analysis and improving our understanding of mineralogy.

Anirudh Prabhu (Carnegie Institution for Science),
Shaunna M. Morrison (Earth and Planets Laboratory, Carnegie Institution for Science),
Ahmed Eleish (Rensselaer Polytechnic Institute),
Peter Fox (Rensselaer Polytechnic Institute),
Joshua J. Golden (Earth and Planets Laboratory, Carnegie Institution for Science),
Robert T. Downs (University of Arizona),
Samuel Perry (University of Notre Dame),
Peter C. Burns (University of Notre Dame),
Jolyon Ralph (Mindat.org),
Robert M. Hazen (Carnegie Institution for Science)
Room: R9 2023-08-07 11:42

Minerals are time capsules that store and provide information about the evolution of Earth and other planetary bodies, because they are the oldest surviving materials from the formation of our solar system. In addition to being a cornerstone of geoscience research, minerals also have economic, industrial and commercial importance in many sectors of society. One of the fundamental questions in mineralogy and geosciences in general is “Where to find minerals?”. The complex and intertwined nature of natural systems makes it hard to predict the occurrences of minerals. However, with increase in the volume and accuracy of mineral data and rise of mineral informatics, data science and analytics methods can be developed to answer this fundamental question in mineralogy.

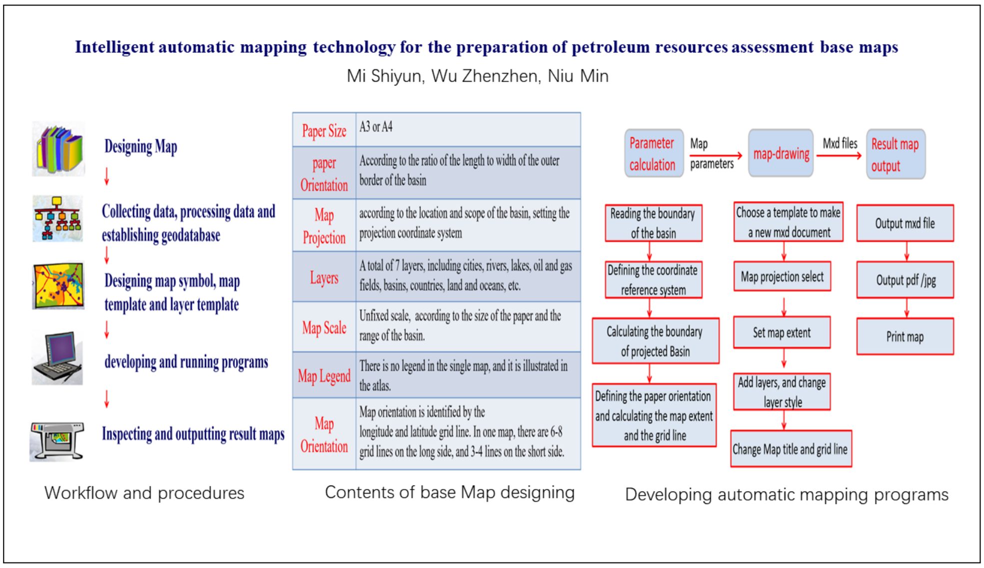
In this contribution, we present “mineral association analysis”, a method to: 1) Predict the mineral inventory for any existing locality. 2) Predict previous unknown localities for any given mineral. Mineral association analysis is a machine learning method that uses association rule learning to find interesting patterns based on mineral occurrence data. Using mineral association analysis, we have been able to predict locations of critical minerals, such as minerals with Li- and Th-bearing phases, predict the mineral inventory of mars analogue sites, and even understand how mineralization and mineral associations changed through deep time.



S0808. Intelligent automatic mapping technology for the preparation of petroleum resources assessment base maps

Shiyun Mi (RIPED of PetroChina),
Zhenzhen Wu (Research Institute of Petroleum Exploration & Development, PetroChina),
Min Niu (Research Institute of Petroleum Exploration & Development, Petrochina),
Qian Zhang (PetroChina)
Room: R9 2023-08-07 14:00

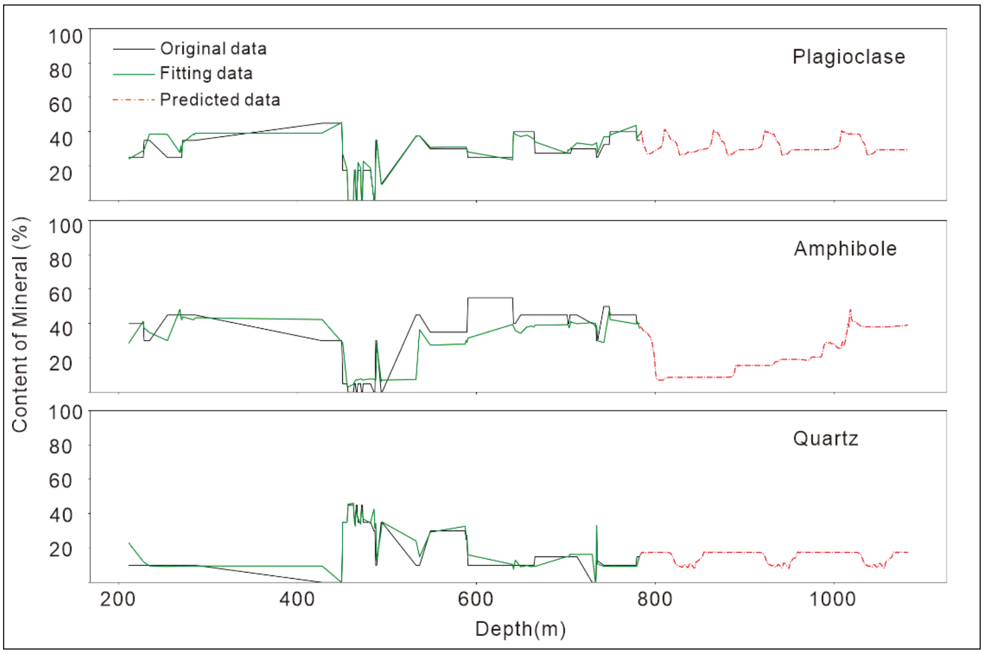
In the research of petroleum resources assessment, it is necessary to draw different types of maps. Geologists usually had to spend a lot of time to plot base maps of assessment area which are relatively constant. In order to improve mapping efficiency and standardize the mapping, an intelligent automatic mapping technology is developed based on the ArcGIS toolkit. It can automatically get spatial data, and apply the appropriate projection method according to the evaluation area, and select corresponding spatial data automatically according to different scales. The technology contains the following workflow and procedures: (1) designing map, (2) collecting data, processing data and establishing geodatabase, (3) designing map symbol, map template and layer template, (4) developing and running automatic mapping programs, (5) inspecting and outputting result maps. By using this technology, it takes only 90 minutes to generate base maps for 468 basins all over the world, also it can automatically modify the map by updating the source data.



S0809. Application of LSTM in Deep Ore Prospecting: A Case Study In Laowan Gold Deposit in Henan Province, China

Lingling Yuan (Institute of Geology, Chinese Academy of Geological Sciences),
Shoutao Jiao (Natural Resources Comprehensive Survey Command Center, China Geological Survey,),
Yueqin Zhu (National Institute of Natural Hazards, Ministry of Emergency Management of the P),
Rongmei Liu (Natural Resources Comprehensive Survey Command Center, China Geological Survey,),
Peng Chai (Institute of Geology, Chinese Academy of Geological Sciences (CAGS), Beijing 100),
Maojian Chen (School of Computer and Communication Engineering, University of Science and Tech)
Room: R9 2023-08-07 14:20

With the increasing demand for economic metals, it is an inevitable trend for geologists to develop diverse ways to understand the deep dynamic process and find deposits in existing mines (Lv et al., 2015). Machine learning and deep learning have been applied in ore prospecting increasingly. long short term memory (LSTM) is a widely used Recurrent Neural Network (RNN) and applied in speech recognition, text modeling, translation and other fields (Graves et al., 2013; Guo et al., 2018; Zhang et al., 2019). This study, based on the data of 35 drill holes in the Laowan gold deposit which is a large-sized deposit located in Henan Province, China, present the relationship between the content of quartz, plagioclase, amphibole and the depth of drill hole. For example, the drill hole 8-23, around 800m deep, wad tested in 3 groups, which were 600m, 650m, 700m for training, and 200m, 150m, 100m for predicting respectively (G1, G2 and G3). The accuracy for amphibole, plagioclase and quartz were 0.875, 0.840 and 0.920, and MSE were 0.0765 0.1116 0.0584 separately. For G2 and G3, the values of accuracy were 0.900, 0.850, 0.850 and 0.867, 0.867, 0.867 for MSE were 0.0713, 0.0669, 0.0767 and 0.0560, 0.0973, 0.0437. The results suggest that the accuracy for different groups is almost equal, above 0.8, but as for the MSE, there is a decreasing with the increasing of training data. In addition, the prediction accuracy for quartz is higher than the other two minerals, which probably results from that quartz occurs at almost every depth of the drilling, in other words, more data of quartz for training improved accuracy of results. A prediction of 300m below the drilling hole was carried on based on the training model, which may be enlightening for deep ore prospecting. In the follow-up work, the training data will be added more geological factors, such as structural data, ore-forming ages and rock types, as well as the relationship between depth and mineral assemblage and metallogenic elements, to make new contributions to deep ore prospecting.


Original data
Fitting data
Predicted data
Plagioclase
Amphibole
Quartz
Content of Mineral (%)
Depth(m)
100
80
60
40
20
200
400
600
800
1000

S09 Accelerating Geoscientific AI with Synthetic Data

Peter Tilke,
Hakan Basarir (NTNU)

Data is crucial for implementing, training and applying AIML solutions for subsurface interpretation. Synthetic data can generate labeled realistic data efficiently and at scale for AIML applications. It is non-identifiable so privacy and data security concerns don't apply. Can be used to train and test AIML models, for exploratory data analysis, to validate assumptions, demonstrate results that can be obtained with AIML models, and reject models producing poor results without the cost of acquiring and incorporating real data. Can be used for transfer learning.

In this session we would like to share experiences with generating synthetic geoscientific data (e.g., forward modeling) for the purpose of accelerating AIML workflows. Building a synthetic data library of geologic analogs represents a level of complexity which involves creating, managing and accessing vasts amounts of data efficiently for a variety of use cases.

S0901. Forward Model Based Synthetic Geological Data for AI/ML workflows

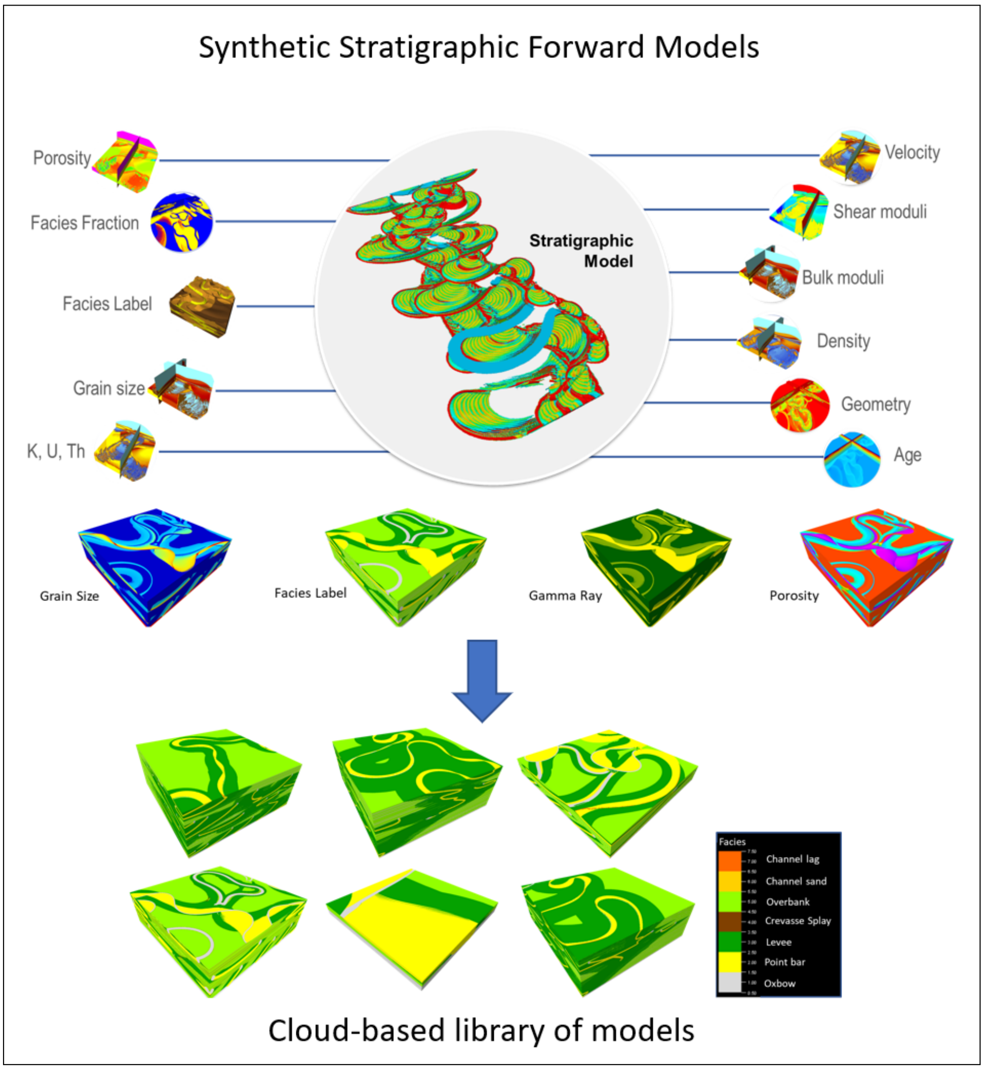
Peter Tilke,
Lingchen Zhu (SLB Schlumberger-Doll Research),
Marie Etchebes (SLB),
Marie LeFranc (SLB)
Room: R8 2023-08-10 13:40

Incorporating stratigraphic processes and geological constraints into 3D modeling helps to eliminate the generation of false representations of the subsurface, ensuring more accurate forecasting of reservoir production, CO2 sequestration capacity, and fluid migration for geothermal energy production.

However, the traditional methods for incorporating these elements using interpreted well logs, seismic data, and 3D stratigraphic modeling tools typically involve repetitive, manual, disconnected, and time-intensive steps.

Employing AI and ML algorithms to automate the steps in the interpretation process is a desirable approach; however, it requires a substantial amount of multi-scale training datasets to be successful. Further, using only actual interpreted well logs, 3D seismic data, or 3D facies models to train the AI/ML engines is not feasible as a large, diverse, and accurately labeled dataset, free from sampling and user bias, is required and unavailable.

To address these challenges, we present the SGFM (Synthetic Geological Forward Modeler) platform, which generates many synthetic geologic models at the reservoir scale. This paper focuses on applying the SGFM to fluvial depositional environments. A diverse set of model parameters covering various depositional settings results in an extensive library of synthetic "geologic analogs." Example model parameters include mean channel width and depth, aggradation, and migration rates. The resulting synthetic 3D models can be used directly to guide interpretation or utilized in various AI and ML-based semi-automated workflows, with examples provided in the paper.



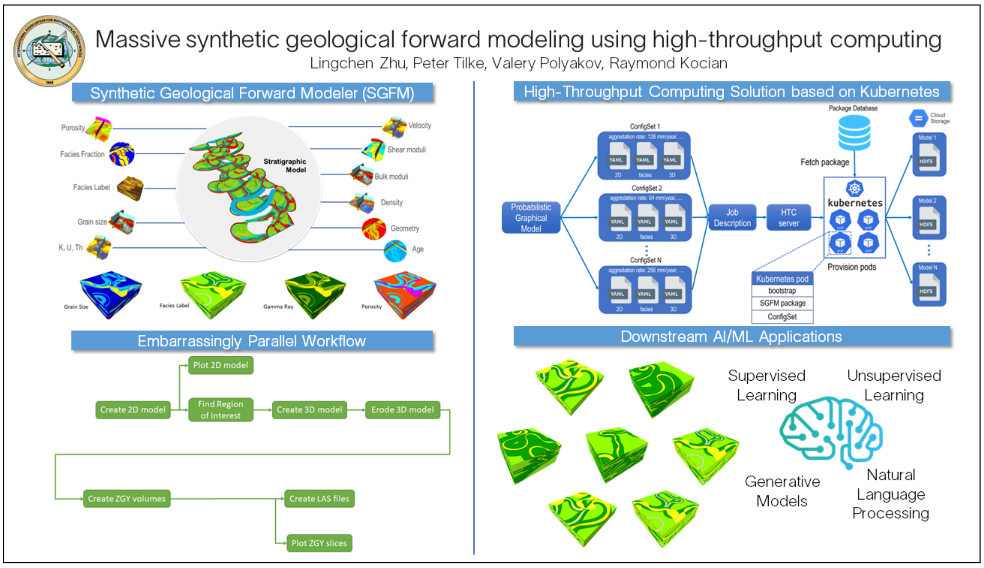
S0902. Massive synthetic geological forward modeling using high-throughput computing

Lingchen Zhu (SLB Schlumberger-Doll Research),
Peter Tilke (SLB),
Valery Polyakov (SLB Schlumberger-Doll Research),
Raymond Kocian (SLB Schlumberger-Doll Research)
Room: R8 2023-08-10 14:00

Generating synthetic geological model data at the reservoir scale has gained interest for many subsurface interpretation workflows, including production forecasting, CO2 sequestration capacity, fluid migration for geothermal energy production, etc. Synthetic data have many advantages compared to real data, including but not limited to minimal privacy concerns, error-free labels, managed features, and unbiased annotations. Most importantly, they are unlimited in number. In this abstract, we present a method to generate a large library of synthetic geological models using our developed Synthetic Geological Forward Modeler (SGFM) at the reservoir scale in a computationally parallel manner, leveraging a hybrid-cloud computing infrastructure that incorporates both on-premise and cloud-based computer clusters.

The overall workflow of SGFM is a Directed Acyclic Graph in which each node is an independent program module that has its own inputs from its predecessor(s) and outputs to its successor(s). However, generating large (in km scale) and thick (in hundred-meter scale) 3D models incurs considerable computational costs as some SGFM modules are inevitably slow. Therefore, to obtain a large amount of synthetic data to accelerate geoscientific AI applications and exploratory data analysis, scaling up computing resources becomes crucial. Fortunately, this computational problem is embarrassingly parallel since multiple 3D geological models and multiple bricks of one single 3D geological model can be created simultaneously with the same codes but using different configuration file sets. Therefore, we can leverage High-Throughput Computing (HTC) to significantly reduce computing time.

In this work, we leveraged an HTC framework based on Kubernetes to conduct parallel SGFM computing jobs on any available on-premises and cloud-based computing resources that we can provide. Results show that we can significantly reduce the wall-clock computing time from more than one year to 3 days if 512 models are created in such way. These large and thick 3D geological models constitute a model library that can be used to train various downstream AIML applications as well as to generate high-fidelity well logs and seismic data for subsurface analysis. To enable fit-for-basin data analysis, we also developed a probabilistic graphical model of geological parameters to generate unique configuration file sets for different geological scenarios.

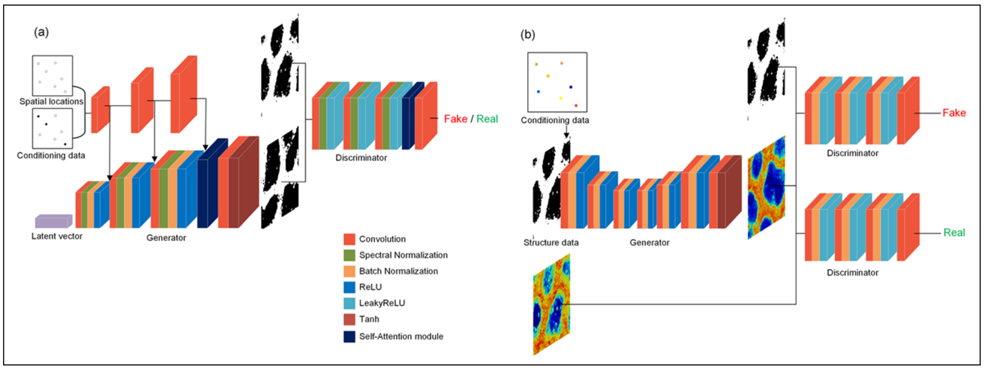


Massive synthetic geological forward modeling using high-throughput computing
Lingchen Zhu, Peter Tilke, Valery Polyakov, Raymond Kocian
Synthetic Geological Forward Modeler (SGFM)
Porosity
Facies Fraction
Facies Label
Grain size
K, U, Th
Stratigraphic Model
Velocity
Shear moduli
Bulk moduli
Density
Geometry
Age
Embarrassingly Parallel Workflow
Plot 2D model
Create 2D model
Find Region of Interest
Create 3D model
Erode 3D model
Create ZGY volumes
Create LAS files
Plot ZGY slices
High-Throughput Computing Solution based on Kubernetes
Probabilistic Graphical Model
ConfigSet 1
ConfigSet 2
ConfigSet N
Job Description
HTC server
Package Database
Fetch package
kubernetes
Provision pods
Kubernetes pod
bootstrap
SGFM package
ConfigSet
Downstream AI/ML Applications
Supervised Learning
Unsupervised Learning
Generative Models
Natural Language Processing

S0903. SA-RelayGANs: a novel framework for the characterization of complex hydrological realism based on GANs and self-attention mechanism

Zhesi Cui (China University of Geoscience (Wuhan)),
Qiyu Chen (School of Computer Science, China University of Geoscience (Wuhan)),
Gang Liu (Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences)
Room: R8 2023-08-10 14:20

Stochastic simulation and deep learning approaches are widely applied to the characterization of complex phenomena and structures in earth and environmental science. These approaches have been proven to be an extremely useful tool in the application to the characterization and simulation of hydrological phenomena. An important factor that hinders our understanding of hydrological realism is the characterization of heterogeneous structures with continuous attributes inside the structures (e.g., porosity, permeability, fluid saturation, etc.). Compared with categorical attributes, continuous attributes can carry more realistic characteristics but with highly increasing in computational consumption of their characterization and simulation. In this work, we propose a novel deep learning approach for the characterization of complex hydrological realism with continuous attributes based on generative adversarial networks (GANs) and self-attention mechanism, named SA-RelayGANs. To reduce the complexity of characterizing the heterogeneous hydrological structures with continuous attributes, we divide the modeling process into two stages: structure construction stage and attribute reconstruction stage. We firstly apply an improved GAN with self-attention mechanism to construct the heterogeneous structures with hard conditioning constrains. Then, another GAN with attributes enhancing term is used to reconstruct realizations based on the constructed structures and observations. SA-RelayGANs can successfully predict the heterogeneous structures with continuous attributes by using limited observations.



S0904. Biosignatures in Complex Molecular Mixtures-a Machine Learning Approach

Grethe Hystad (Purdue University Northwest),
H. James Cleaves II (Carnegie Institution for Science, Tokyo Institute of Technology, International S),
Anirudh Prabhu (Carnegie Institution for Science),
Michael L. Wong (Carnegie Institution for Science),
George D. Cody (Carnegie Institution for Science),
Sophia Economon (Johns Hopkins University),
Robert M. Hazen (Carnegie Institution for Science)
Room: R8 2023-08-10 14:40

What is the fundamental difference between the chemistry of life and the chemistry of the nonliving world? Our hypothesis is that the diversity and distribution of molecules in living systems are different than those in nonliving systems since biomolecules are selected through evolutionary processes. We suggest that there exist "molecular biosignatures" composed of sets of molecules that distinguish terrestrial biochemistry from abiotic chemistry and that even alien biology would produce interpretable molecular sets via pyrolysis that would differ from those formed through abiotic processes.

A varied collection of natural and synthetic organic molecular mixtures was analyzed using pyrolysis gas chromatography-mass spectrometry (py-GC-MS), which decomposes the samples into fragment ions for molecular identification. The GC-MS data are high dimensional and can be represented as a three-way array (samples $\times$ retention times $\times$ mass to charge ratio) with intensity peaks. We used machine learning and statistical learning theory to investigate classification rules to predict whether materials were of biotic or abiotic origin. Our goal was to determine collective features that allow differentiation between the two groups. In this talk we present preliminary results that suggest that organic pyrolysis products visibly fall into three clusters: organic material derived from abiotic sources (material from laboratory prebiotic chemistry simulations, meteoritic organics, etc.), living terrestrial matter (microbes, plants, fungi, etc.), and geologically processed biotic organics (including coal, oil shale, petroleum, etc.). We discuss why these tools will likely be useful for the search for evidence of life beyond Earth, verification of the earliest traces of life on Earth, and possibly the detection of the origin of life in lab simulants.

S10 Spatiotemporal Methods and Data Analysis

Dionysios Christopoulos (Technical University of Crete),
Sandra De Iaco (University of Salento),
Emmanouil Varouchakis (Technical University of Crete)

The development of spatiotemporal models is at the forefront of current research in Geostatistics and Machine Learning. Currently there is a need for models that can accurately capture the complex patterns observed in various dynamic processes and to reliably estimate uncertainties. Key modeling challenges for the analysis of modern spatiotemporal data involve the development of models suitable for large data sizes, heterogeneous data, multivariate dependence, multiple correlation scales and non-Gaussian probability distribution.

This session seeks contributions that will advance spatiotemporal Geostatistics by proposing novel concepts and methodologies, computational algorithms, and innovative applications or studies that focus on spatiotemporal data analysis. An indicative but non-exhaustive list of topics includes the development of novel space-time covariance functions (e.g., non-separable models, models on the sphere and manifolds, multivariate dependence, complex-valued models), covariance-free approaches (e.g., models based on stochastic partial differential equations and explicit precision operators), innovative simulation methods, computational advances for big space-time datasets. The session also invites contributions that focus on applications involving interesting and/or challenging spatiotemporal datasets, as well as approaches for non-Gaussian space-time data and multiscale models. Contributions that combine Geostatistics with current developments in Applied Mathematics (e.g., uncertainty quantification and multifidelity frameworks), and Machine Learning (e.g., hierarchical models, sparse and multiscale Gaussian process regression) are also welcome.

S1001. Spatiotemporal interpolation of large meteorological fields

Sofia D. Nerantzaki (University of Saskatchewan),
Dionysios Christopoulos (Technical University of Crete),
Simon Michael Papalexiou (Department of Civil Engineering University of Calgary, Calgary, Canada)
Room: R7 2023-08-08 10:00

Temperature and precipitation are the primary variables characterizing the Earth's climate system and affecting hydrological processes. Therefore, they are the main input variables in most water management and risk assessment studies. Real-world, large-scale hydrologic applications are essential for understanding the global climate system and water cycle and assess the natural hazards risk, and the water supply and demand in a changing environment. For large-scale studies, temperature and precipitation fields of large dimensions are needed. However, due the scarcity of meteorological stations and the underperformance of existing stations, the available observational fields have many missing values in space and time. Spatiotemporal stochastic simulations of such environmental flows can also be employed for water management applications. However, large-scale simulations of temperature and precipitation fields (RFs), are subject to size limitations due to computational speed and memory requirements.

We propose using the stochastic local interaction (SLI) model to achieve fast space-time interpolation which allows generating dense grids of large (simulated or observed) meteorological fields. The SLI method uses local interactions among near neighbors while kernel functions with local bandwidths define the interaction strength and the neighborhood size. The method can be applied to both lattice and non-lattice data. The locality of the interactions allows constructing sparse and explicit precision (inverse covariance) matrices for space-time dependence. SLI interpolation thus does not require matrix inversions and consequently has modest requirements for storage, enabling fast generation of large spatiotemporal fields.

We apply SLI to large-scale observational temperature data and simulated rainfall RFs. For the simulation of rainfall RFs, we use the Complete Stochastic Modeling Solution (CoSMoS) framework. CoSMoS preserves any non-Gaussian marginal distribution, any spatiotemporal correlation structure, the locally varying anisotropy, and the general advection given by velocity fields with locally varying speed and direction. Thus, we are able to simulate RFs at fine spatiotemporal scales, which exhibit spatiotemporal patterns and characteristic motions of realistic rainfall storms. This study provides a methodological framework for the interpolation of meteorological networks and the simulation of large moving meteorological fields which can then be used as input to large-scale hydrological studies.

S1002. Locally refined spline surfaces and volumes for representation and analysis of geo-spational points

Vibeke Skytt (SINTEF)
Room: R7 2023-08-08 10:20

Modern acquisition methods produce large geospatial point clouds. Approximating the points with a surface or a volume, as appropriate, reduces the data size of a recorded point cloud and provide a structure. Such a mathematical representation further gives an excellent basis for analysing the present data. Locally refined splines (LR B-splines) is a flexible surface or volume representation format that is able to represent the smooth component of given data with a specified accuracy. The surface is constructed by adaptive approximation facilitating a concentration of surface coefficients in areas with high local variation while being lean in smoother areas. An example combining bathymetry data from the seabed and a LIDAR point cloud from adjacent land into one surface will be discussed. Spatio-temporal data organized as a sequence of corresponding geospatial point clouds from different points in time can be represented by an LR B-spline volume. When differentiating this volume with respect to time, deformations are easily recognized.

S1003. Stochastic Local Interaction Models for Multivariate Datasets

Dionysios Christopoulos (Technical University of Crete)
Room: R7 2023-08-08 10:40

Multivariate datasets are typically modeled in terms of vector random fields (VRFs). The interpolation and simulation of VRFs rest on a matrix covariance function which is estimated from the data. Permissible covariance functions should ensure that the variance of any linear combination of component processes is non-negative. Using Bochner's theorem, it can be demonstrated that many scalar functions are non-negative definite. In the case of VRFs, covariance functions should satisfy Cramer's theorem which places more stringent conditions than Bochner's. Matrix covariance functions employ the separability hypothesis which uses the product of a marginal correlation function and a constant nonnegative definite matrix. This construction is convenient but not compatible with solutions of physical models. The popular linear model of co-regionalization (LCM) represents each vector component as a linear combination of latent, independent, univariate spatial processes. The LCM also poses significant limitations, since its smoothness is dominated by the roughest latent component.

For large datasets, covariance functions with infinite support can lead to computational difficulties resulting from the inversion of very large covariance matrices. Stochastic local interaction models (SLIs) can overcome the computational costs [1]. SLI models lead to sparse precision (inverse covariance) matrices which can be much faster than covariance-based methods (e.g., kriging). SLI predictive equations scale linearly with the size of the dataset, compared to the cubic dependence exhibited by covariance–based methods [2].

This presentation focuses on recent progress that extends SLI models to multivariate datasets. This is achieved by constructing energy functions that incorporate interactions between different vector components. The interactions are implemented by means of sparse precision matrices which involve several hyperparameters and parameters, depending on the number of vector components. The explicit SLI precision matrix implies that permissibility is trivially satisfied for non-negative model parameters. We test vector SLI with the Jura heavy metal dataset and compare its performance with literature results for co-kriging and machine learning algorithms.

References

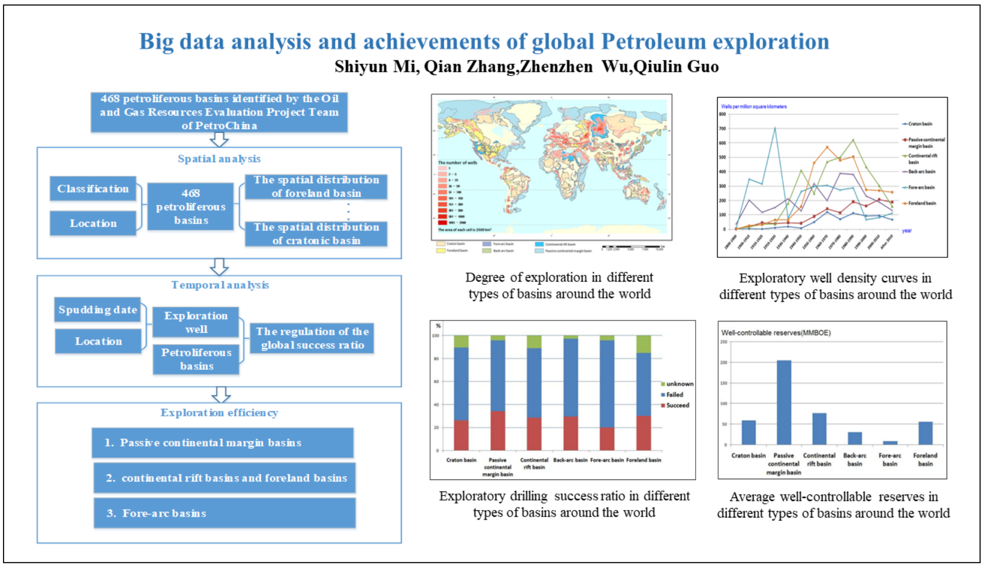
1. D. T. Hristopulos, Random Fields for Spatial Data Modeling; Springer, Netherlands, 2020.

2. Hristopulos, D.T.; Agou, V.D. Stochastic local interaction model with sparse precision matrix for space–time interpolation. Spatial Statistics, 2020, 40, 100403.

S1004. Big data analysis and achievements of global Petroleum exploration

Jingdu Yu (Petrochina),
Shiyun Mi (RIPED of PetroChina),
Qiulin Guo (PetroChina),
Qian Zhang (PetroChina),
Zhenzhen Wu (Research Institute of Petroleum Exploration & Development, PetroChina)
Room: R7 2023-08-08 11:00

Through applying the data mining analysis techniques of cluster analysis and correlation analysis, as well as spatial analysis and data visualization analysis techniques closely combined with GIS technology, based on IHS's global exploratory well data, global oil and gas field data, the data and its classification schemes of 468 petroliferous basins identified by the Oil and Gas Resources Evaluation Project Team of PetroChina, this paper conducted topological analysis, spatial overlay analysis and spatial statistical analysis research in the system of ArcGIS. For the first time, it completed the globally big data mining and analysis of various basin drilling exploration. Meanwhile, it applied small patches technique to analyze the heat of global oil and gas drilling exploration, obtained some new understandings of the global oil and gas exploration activity routines, which is considered to be significant in guiding the optimization of new exploration projects worldwide, and indicated the good application prospect of the big data analysis technology in the oil and gas industry. The new findings are as follows. (1) The prospects with high degree of exploration globally mainly concentrate in Tethyan tectonic domain. (2) In the history of petroleum industry, the active period was distinct with basin types. Specifically, it was in 1910-1920 for for-arc basins, 1950-1980 for foreland basins, and 1980-1990 for continental rift basins. (3) The global success ratio of exploratory wells is 28.8% on the average, and the success ratio does not differ greatly for various types of basins. Generally, the ratio is relatively higher for passive continental margin basins, but relatively lower for for-arc basins. (4) Through estimating the average well-controllable reserves to denote the exploration efficiency in a specific basin type, it is illustrated that the exploration efficiency was the highest in passive continental margin basins, followed by continental rift basins and foreland basins, and the lowest in fore-arc basins.


Big data analysis and achievements of global Petroleum exploration
Shiyun Mi, Qian Zhang,Zhenzhen Wu,Qiulin Guo
468 petroliferous basins identified by the Oil and Gas Resources Evaluation Project Team of PetroChina
Spatial analysis
Classification
Location
468 petroliferous basins
The spatial distribution of foreland basin
The spatial distribution of cratonic basin
Temporal analysis
Spudding date
Location
Exploration well
Petroliferous basins
The regulation of the global success ratio
Exploration efficiency
1. Passive continental margin basins
2. continental rift basins and foreland basins
3. Fore-arc basins
Degree of exploration in different types of basins around the world
Exploratory well density curves in different types of basins around the world
Exploratory drilling success ratio in different types of basins around the world
Average well-controllable reserves in different types of basins around the world
Well-controllable reserves(MMBOE)
unknown
Failed
Succeed
Craton basin
Passive continental margin basin
Continental rift basin
Back-arc basin
Fore-arc basin
Foreland basin

S1005. EOF-based analog selection for satellite-derived wind speed image time series reconstruction: A multiple-point statistical framework

Stylianos Hadjipetrou (Cyprus University of Technology),
Gregoire Mariethoz (University of Lausanne),
Phaedon Kyriakidis (Cyprus University of Technology)
Room: R7 2023-08-08 11:20

The rapid transition to renewable energy exploitation has been accompanied by a radical expansion in the offshore wind energy sector in recent years. A critical step in planning an offshore wind farm is the evaluation of the project feasibility via the collection of wind field data. Related studies require long-term datasets to evaluate the wind climate over the region of interest. Acquiring such data in the desired spatial and temporal scales, however, can be challenging, particularly in offshore regions where in-situ measurements are scarce.

Satellite-mounted SAR sensors provide fine spatial-resolution imagery over extensive offshore areas, forming a consistent record of wind speed estimates. Sentinel-1 A&B satellites have been in operation since 2014 and 2016, respectively, emitting radio waves toward the sea surface. Once processed, the backscattered intensity of sea-level surface waves provides reliable images of wind speed estimates. The resulting image time series, however, suffers from poor temporal resolution due to the satellites' revisit frequency. Despite several attempts to reconstruct satellite-derived image time series using spatial and/or temporal gap-filling methods, only a few studies have achieved the reconstruction of the fine-scale variability of the underlying fields and none within an offshore wind speed context.

This study addresses the above issue by employing a Multiple-point Statistical (MPS) framework to reconstruct Sentinel-1 A&B OCN Level-2 wind speed image time series over the offshore area of Cyprus. In contrast to previous related works, the selection of the co-registered auxiliary reanalysis data (Training Images-TIs) is based on Empirical Orthogonal Function (EOF) analysis preceded. Once physically-consistent analogs are selected, MPS are employed to simulate the missing satellite information. The performance of the proposed method, in terms of reproducing the spatial and temporal variability of the reference data, is evaluated via the use of cross-validation and statistical metrics. Overall, the results demonstrate the effectiveness of the proposed method in accurately reconstructing the wind speed image time series over the offshore area of Cyprus and can be generalized to other regions, scales or variables.

S1006. A multivariate space-time geostatistical approach for modeling and predicting agrometeorological variables

Claudia Cappello (University of Salento),
Sandra De Iaco (University of Salento),
Monica Palma (University of Salento, Lecce, Italy),
Klaus Nordhausen (University of Jyäskylä, Jyäskylä, Finland)
Room: R7 2023-08-08 11:40

In environmental studies one of the main interests is the analysis of the spatio-temporal relationships among several variables and the prediction of one variable of interest in unsampled locations (in space and in time).

Multivariate Geostatistics offer appreciable techniques and tools to estimate the spatio-temporal covariance function and define an apt model which provides reliable spatio-temporal predictions of the primary variable under study.

This paper aims at presenting a spatio-temporal multivariate geostatistical modeling approach, based on the joint diagonalization of the empirical covariance matrix. In particular, the proposed procedure consists of a) identifying the spatio-temporal covariance structure of the independent components b) testing the main features, regarding second order properties which characterize each basic independent component and c) choosing appropriate classes of covariance models for the basic components in space-time. Hence this procedure helps to consider a reduced number of independent variables (lower than the number of observed variables) and to model separately the spatio-temporal evolution of these independent.

A case study concerning weekly averages of environmental variables, such as evapotranspiration, minimum and maximum humidity, maximum temperature and precipitations, recorded at some stations in Veneto Region (Italy), for 23-year span is thoroughly discussed. A space-time linear coregionalization model (ST-LCM) with properly chosen models for the retained latent components will be fitted to the matrix-valued covariance function estimated for five relevant agrometereological variables. Finally, the fitted ST-LCM will be used to predict evapotranspiration levels. The results of this multivariate study might be of interest for their reflections in agriculture and for contrasting the drought emergency in Veneto Region.

S1007. Signal processing and reconstruction methods with data gaps in geophysics (applications to Mont Terri URL)

Rachid Ababou (IMFT: Institut de Mecanique des Fluides de Toulouse),
Jean-Michel Matray (IRSN: Institut de Radioprotection et de Sûreté Nucléaire, France)
Room: R7 2023-08-08 14:00

This work presents statistical and deterministic methods for reconstructing geophysical signals, with applications to pore pressure and rock deformation signals collected at the Swiss Mont Terri URL (Underground Research Laboratory), BGR experiment MBA-2 Phase 25. These signals have strong nonlinear trends and significant data gaps. Other signals, synthetic or measured, are also used to test possible extensions of the methods.

We present the in-situ pressure/deformation signals, including positions of sensors in the excavated gallery, raw signals plots (pore pressures P(t) [Pa], rock deformations D(t) [mm], and the evolving excavation front X(t) [m]), pointing out significant data gaps.

We then present (A) a class of statistical reconstruction methods based on modified adaptive AR1 auto-regressive process, and (B) deterministic methods based on either polynomial interpolation or smoothing (best fits). All methods are implemented for the same reconstruction time step of 900s.

Advantages and drawbacks of methods (A) and (B) are discussed.

Concerning methods (A), some AR1-based statistical reconstructions seemed acceptable. The least satisfactory results were obtained for some pore pressure signals combining long data gaps with strongly nonlinear trends. These results seemed related to inadequate characterization of mean, variance, autocovariance structure of signals, given that: (i) the true signal is not AR1; (ii) it is not Gaussian; (ii) it is non-stationary to both 1rst order and 2nd order; and (iii) the data gaps are too many or too long for a proper characterization of the signal.

Concerning deterministic methods (B), PCHIP interpolation emerges as the preferred method for highly variable time step signals (PCHIP: "Piecewise Cubic Hermite Interpolation Polynomials), compared to other interpolations (e.g. Spline interpolation) and to best fit methods (e.g. piecewise cubic Spline fit). Lagrange interpolation was useless due to Runge phenomena near data gaps.

Finally, other signals, real and synthetic, are used to assess extensions of these methods. Thus, we implement Choleski factorization for generating random processes based on empirically estimated autocovariance functions with variable timesteps. Future progress may depend on combining deterministic/statistical processing, new methods for separating nonlinear trends from residuals, and designing better estimation of signal structure with data gaps.

S1008. Exploring seasonal patterns in baseflow using functional data analysis

Kathryn Leeming (British Geological Survey),
John P Bloomfield (British Geological Survey),
Gemma Coxon (British Geological Survey),
Yanchen Zheng (British Geological Survey)
Room: R7 2023-08-08 14:20

We demonstrate the application of functional data analysis to understand annual seasonal shapes in baseflow using the CAMELS-GB dataset. The general seasonal shape in baseflow is similar across time and space, however the timing of the annual peak and trough differs with some catchments exhibiting later Spring peaks in baseflow. It is important to understand the timing of the peaks and troughs in baseflow for consideration of flood and low-flow events.

Our work applies the funFEM method to form clusters of the annual baseflow curves, grouping catchments together with similar seasonal distribution of baseflow. This approach results in geographically and geologically coherent groups, allowing for further exploration of the reasons behind the different seasonal patterns. We also split the original data into time blocks to explore whether the seasonal patterns have changed over time, and which other variables may have led to such changes.

This application of functional data analysis gives an additional way to consider the seasonal patterns found in many environmental datasets, by treating them as curves rather than as discrete seasonal observations.

S1009. Adaptive ocean sampling using surrogate models with spatially varying anisotropy in three dimensions

Martin Outzen Berild (NTNU),
Geir-Arne Fuglstad (Norwegian University of Science and Technology)
Room: R7 2023-08-08 14:40

Mapping coastal biomes and ocean masses such as polar melting water or river plumes are of growing interest. Currents, winds, and freshwater runoff make ocean variables such as salinity and temperature dynamic, and therefore, simple statistical models can be unreasonable for monitoring such complex environments. The stochastic partial differential equation (SPDE) approach has been shown to be effective in predictions on large-scale spatial and spatio-temporal models and allows for spatially varying effects. Using the SPDE approach, we construct Gaussian random field (GRF) prior models for ocean variables outside Trondheim, Norway, as surrogate models for a complex numerical ocean model. We compare different models specified through the SPDEs in prediction accuracy on field measurements collected with an autonomous underwater vehicle (AUV).
In a separate field deployment, the viability of an adaptive sampling strategy in mapping the river plume outside of Trondheim is evaluated. This approach extends the methods of adaptive oceanographic sampling by considering the three spatial dimensions of the ocean and by using a more realistic model accounting for spatial variability in salinity. The adaptive sampling method uses a myopic strategy and focuses on separating fresh and saline water masses where the AUVs objective is to reduce the variance of this excursion set. The three-dimensional GRF model enables fast on-board updating of the predicted salinity field and associated uncertainties from the in-situ AUV data. Results show that the AUV conducts effective sampling of the front of the river plume.

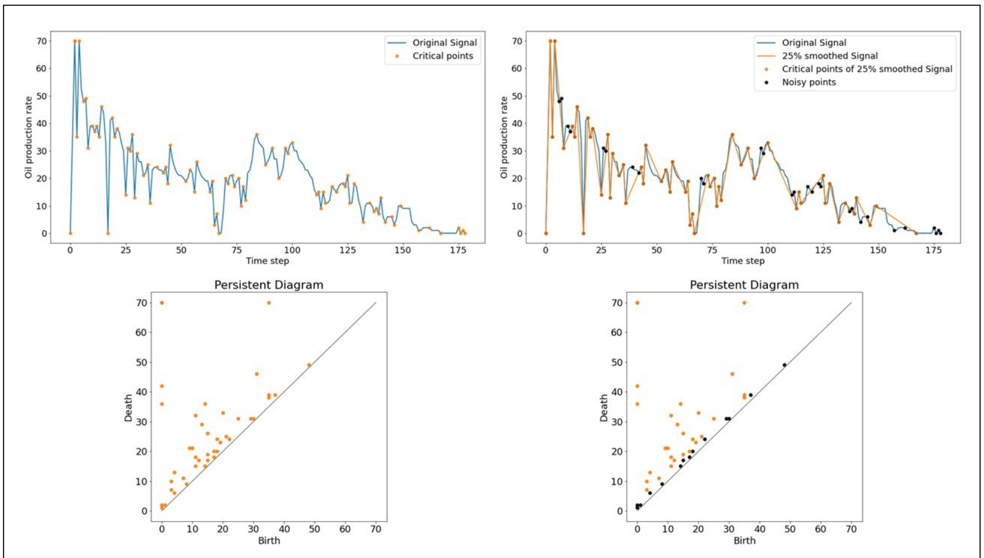
S1010. Extracting critical production data for assisted history matching problem using topological data analysis

Anna Vetkina (National Research Tomsk Polytechnic University),
Gleb Shishaev (Heriot-Watt University)
Room: R7 2023-08-08 15:00

In this study, Topological Data Analysis (TDA) of time series describing production data is provided to identify and use only the critical production data, or data of high importance (i.e. data records with a high degree of influence) in the process of Assisted History Matching (AHM) of reservoir models. Production data, such as liquid rates, water cut, and bottom hole pressure, that are typically recorded monthly or sometimes daily, and represented as time series with irregular erratic fluctuations of small amplitudes, make the AHM problem pretty complicated to address, due to being computational cost and the extensive time required. Thus, these small peaks and fluctuations are assumed to be inappropriate data for AHM. As a result, the present study focuses on the effective use of data by showing that extracting only critical (important) points for being applied to AHM process is much more efficient and preferable than using very detailed historical data in the AHM process.

Topological analysis of production data, using ideas from Persistent homology theory, allows extracting the most important points of the data. Zero-dimensional homology explores topological features of a one-dimensional function and, hence, allows analyzing time series. Applying TDA to the production data, the critical points in the dataset are first distinguished from the noises, and further are considered in the parsed history in AHM. The latter is implemented by setting weighting factors in the objective function at AHM.

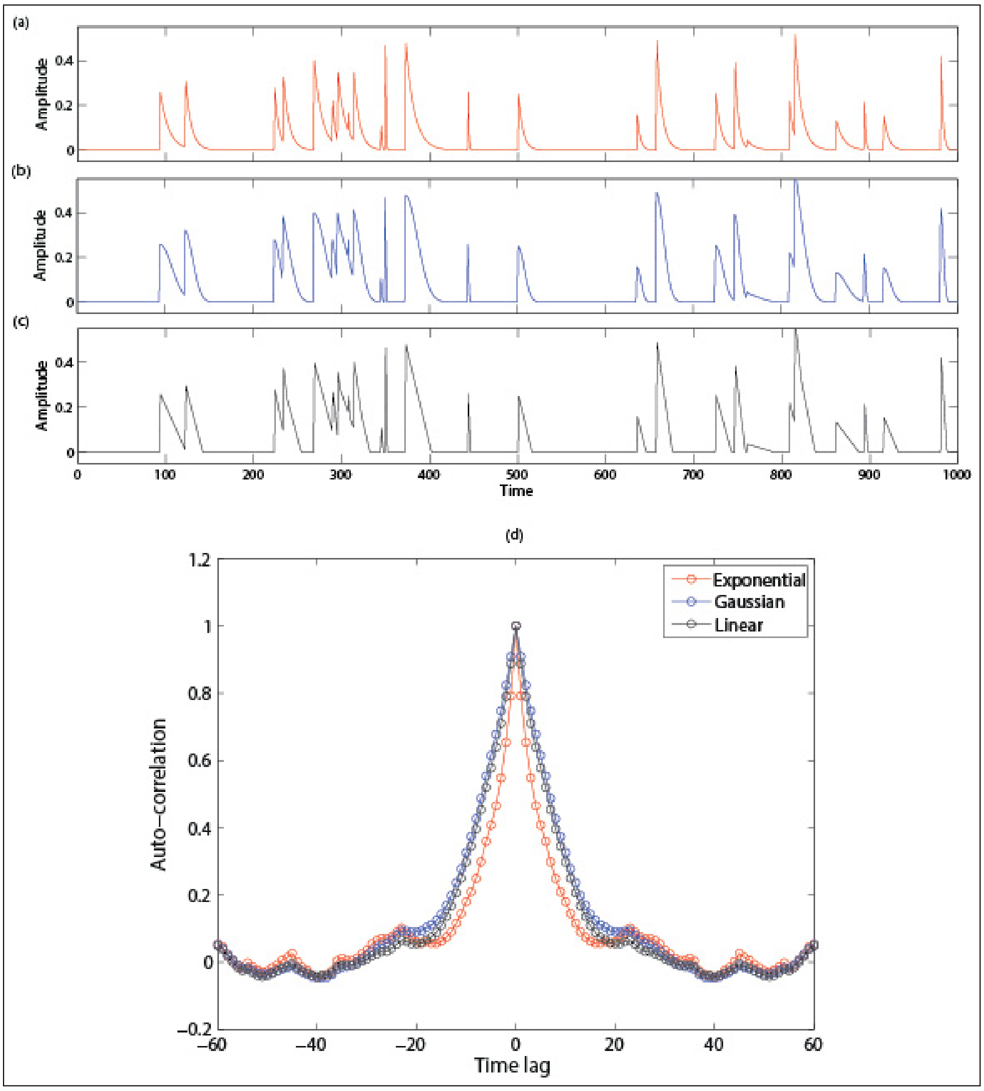
A case study was performed on a simulation model of a sector of a real field, which has 50 producing wells over an area of 25 km2 with a development history of over 15 years. The findings revealed that the quality of history-matched models and their forecasts depend on the critical records of production data rather than the full production entire excessive dataset. Therefore, the less influential data records are recommended to be eliminated from datasets to reach a computationally efficient AHM.



S1011. Cautionary remarks on the cross- and auto-correlation analysis of non-Gaussian and non-linear time series

Sung Yong Kim (KAIST)
Room: R7 2023-08-08 15:20

Correlation analysis has been widely used for time domain analysis to find relevance between two variables in geosciences. In this talk, we examine the auto-correlations of the self-similar time series to identify spurious decorrelation structures in terms of the number of independent pulses and the shape of decay patterns as a part of time domain analysis of non-linear time series in geosciences. The self-similar time series is defined as a continuous time series having similar shapes of disturbance or amplitudes of which statistics are non-Gaussian, such as records of river flows, rainfall, wind speed, the concentration of Chlorophyll, and inertial amplitudes in geosciences. The decorrelated structure and decorrelation scales estimated from the correlation analysis of self-similar time series are not directly relevant to the shapes and averaged decay scales of the self-similar time series. Thus, the decorrelated structure and decorrelation scales should be estimated with individual pulses of the self-similar time series instead of cross- and auto-correlation analysis. We present an evaluation of the cross- and auto-correlations of the modeled self-similar time series and discuss the relevant cautionary remarks.



(a)
(b)
(c)
Amplitude
Time
(d)
Exponential
Gaussian
Linear
Auto–correlation
Time lag

S1012. Exploring Residential Built-Up Form Typologies in Delhi: A Grid-Based Clustering Approach towards Sustainable Urbanization

Aviral Marwal (University of Cambridge)
Room: R7 2023-08-09 10:00

Urbanization in developing nations is a multifaceted issue that presents opportunities and difficulties for sustainable living. Beyond its relationship to economic growth, it's crucial to have a more thorough comprehension of urbanization to tackle its various challenges. This can be accomplished by characterizing and assessing urban form, which significantly affects socio-economic and environmental well-being. The knowledge of urban forms in cities in developing countries is limited in academic literature, largely due to the difficulty in analyzing the extensive diversity of micro-scale urban form features. This study aims at uncovering different residential built-up form typologies in Delhi using a grid-based clustering algorithm and evaluate their impact on sustainable urbanization. The k-means clustering algorithm measures and visualizes variations in urban form by dividing residential areas into 100x100 meter grid cells and assigning attributes of accessibility, built-up density, and street design. The results reveal the presence of six unique built-up form typologies in Delhi, with only 19% of residential areas classified as sustainable. The study highlights the need for interventions in unsustainable areas to achieve sustainable urbanization, including providing services, improving street design, creating open spaces, and enhancing pedestrian-friendly street networks. These findings not only add to the limited literature on urban morphology in developing cities but also provide a deeper understanding of the pace of urbanization.

S1013. Wavelet-based local structural complexity analysis of Lancaster-Sarmanov subordinated random fields

José M. Angulo (Department of Statistics and O.R., University of Granada), María D. Ruiz-Medina (Department of Statistics and O.R., University of Granada), Ana E. Madrid (Department of Statistics and O.R., University of Granada, Spain)
Room: R7 2023-08-09 10:20

A crucial problem in Geostatistics consists of suitable modeling of the underlying small-scale spatial dependence structure when only discrete observations are available. The present work provides an approach based on mutual information and wavelet analysis, beyond the Gaussian and linear scenarios; even in the latter case, this modeling problem arises in determining suitable conditions for the existence of a well-posed solution to spatial linear extrapolation in the mean-square sense.

More specifically, generalized mutual information measures are applied to studying asymptotic local dependence properties of stationary spatial processes subordinated to the Lancaster-Sarmanov (LS) random field class. Multiscale structural complexity is analyzed, with the main focus on situations where the underlying random field may display a spatial fractal behavior.

In the context of risk assessment of threshold exceedances, the special case of Minkowski functionals subordinated to Gaussian and Gamma-correlated LS random fields is studied in some more detail. The results derived are illustrated and interpreted from the perspective of microscale geometrical diversity analysis of sample paths, particularly regarding the eventual occurrence of anomalous local behaviour that may arise in diverse geophysical phenomena.

Extension to the spatiotemporal case, following the functional approach derived in Angulo and Ruiz-Medina (2022, 2023), is finally discussed.

(Acknowledgements: Work partially supported by grants PGC2018-099549-B-I00 (M.D. Ruiz-Medina) and PID2021-128077NB-I00 (J.M. Angulo and A.E. Madrid), funded by MCIN / AEI/10.13039/501100011033 / ERDF A way of making Europe, EU, and grant CEX2020-001105-M funded by MCIN / AEI/10.13039/501100011033.)

References

Angulo, J.M., Esquivel, F.J., Madrid, A.E., and Alonso, F.J. (2021) Information and complexity analysis of spatial data. Spatial Statistics 42, 100462.

Angulo, J.M., and Ruiz-Medina, M.D. (2022) Macroscale structural complexity analysis of subordinated spatiotemporal random fields. arXiv2212.12209v1.

Angulo, J.M., and Ruiz-Medina, M.D. (2023) Infinite-Dimensional Divergence Information Analysis. In: Trends in Mathematical, Information and Data Sciences, N. Balakrishnan, M.A. Gil, N. Martín, D. Morales and M.C. Pardo (eds.), 147-157. Studies in Systems, Decision and Control 445. Springer.

S1014. Spatiotemporal Signal Decomposition for CO2 Sequestration and Monitoring

Jose L. Hernandez Mejia (The University of Texas at Austin),
Michael J. Pyrcz (University of Texas at Austin)
Room: R7 2023-08-09 10:40

Carbon Capture and Storage (CCS) consist on capturing CO2 from anthropogenic sources and injecting CO2 into the geological formations with the goal of permanent storage; therefore monitoring the subsurface CO2 migration with a variety of techniques is essential to ensure that it remains safely sequestered is critical. Pressure sensing is cost-effective, and efficient for monitoring large pore networks. However, in the presence of multiple CO2 injectors and monitoring wells, the pressure response in the monitoring wells can be influenced by the multiple CO2 injector wells making it complicated to monitor individual CO2 plumes from associated injector wells.

In this work, we propose a spatiotemporal method for bottomhole pressure signal decomposition to identify single injector signal contributions into the monitoring well. First, using a physics-constrained Voronoi decomposition the time and spatial boundaries at the time when multiple adjacent injection wells initiate interference with the monitor well are identified. Then, using a deep neural network (DNN) the different signal contributions when more than one injection well is interfering with the monitor well are decomposed. The DNN is trained on pressure time series obtained from high-fidelity reservoir simulations under multiple subsurface realizations. Finally, a physics-constrained dynamic time warping (PCDTW) is implemented to find the optimal alignment between the individual signal contributions in the monitor well and the CO2 injector signals. The total cost of the optimal warping path between the signals is used to identify the injection source of the pressure signal in the monitor well. The lag time obtained from the signal matching using PCDTW is then used to characterize the CO2 plume movement in the geological formation for improved spatiotemporal monitoring. This workflow enhances subsurface CO2 storage monitoring over time to minimize environmental risk.

S1015. Cointegration Analysis of Factors Affecting Carbon Emission Based on Time Series Data

Xin Xie

Room: R7 2023-08-09 11:00

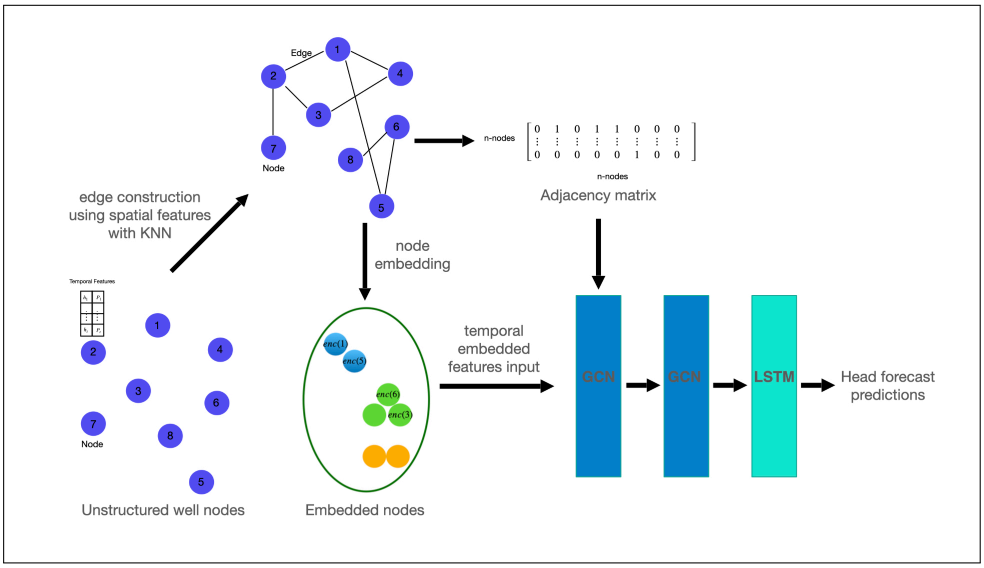
The integrated carbon budget management framework plays a positive role in guiding and promoting the global climate governance process. According to the Global CO2 Emission data, the work compares and analyzes the daily anthropogenic carbon emission data from 12 countries from January 1st, 2019 to August 31st, 2022, including Power, Industry, Ground Transport, Residential, Domestic Aviation, International Aviation, and total carbon emission data. The basic statistics analysis, R/S time series analysis and a computational simulation model based on five possible future scenarios, put forward by IPCC, including SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 are used to analyze and predict the evolution of carbon and evaluate the future scenarios of carbon emissions. The basic statistical analysis shows that the country with the largest daily CO2 emissions are 38 times of the smallest region. There is a certain correlation between the 6 anthropogenic emissions sources of Power, Industry, Ground Transport, Residential, Domestic Aviation and International Aviation in different regions. The overall variation of CO2 emissions in every country is small, with a coefficient of variation less than 0.5, a skewness close to 0, and a kurtosis less than 5. R/S time series analysis shows that Hurst index is greater than 0.5, which shows that all regions have strong sustainable carbon emissions, indicating that the world's carbon emissions may also depend more on the energy consumption structure for a long time. In addition, carbon emission simulation suggests the growth rate of total carbon emissions will reach the maximum around 2050, following the scenario of SSP5-8.5, and will increase slowly in the later period with decreasing growth rate, which may be attributed to the rapid growth of the global economy and is driven by the exploitation of fossil fuels and land use. This prediction result implies continuous and rapid warming will last for at least several decades, and the work will provide new insights into global warming study.

S1016. Deep Neural Network Workflow for Spatio-Temporal Groundwater Head Forecast

Xiao Xia Liang (Institut National de la Recherche Scientifique)
Room: R7 2023-08-09 11:20

In the recent years, the advancement of machine learning, specifically, deep neural networks, has many researchers utilizing Recurrent Neural Networks (RNN) in the field of hydrologeology to forecast head levels in wells. While these RNN frameworks work well with temporal data like timeseries head measurements, it does not consider the parameters' spatial distribution in and around wells, e.g. hydraulic conductivity. Currently, there are no published frameworks that consider the spatio-temporal aspect of groundwater flow and its' forecasting. Here, we present a Graph Neural Network framework that incorporates spatial and temporal data for spatio-temporal forecasting of groundwater flow.

The framework takes spatially unstructured pumping and observation wells as nodes. Each node contains edge features (spatial data) and node features (temporal data). Spatial features such as latitude, longitude and hydraulic conductivity are used for the edge construction. These edges, which creates the spatial relationship between the nodes with the k-nearest neighbours' (KNN) algorithm, where the KNN generated edges are hydrogeological links between the nodes. The results of the edge connections are a n-nodes by n-nodes matrix which is known as the adjacency matrix. Temporal data associated with the nodes goes through a manifold embedment for dimension reduction. The embedded temporal data and the adjacency matrix are feed into two layers of Graph Convolutional Network (GCN) then through a layer of Long Short Term Memory (LSTM). The forecast is made after the LSTM layer. To train our framework, we used simulated data from a conceptual hydrogeological model developed for hydrogeological assessments near Montreal, Québec, Canada. The results from our GNN model forecasts compare well with the results generated using the numerical simulations. However, once trained, the neural network forecasting takes minutes instead of days for the numerical model. Our next step is to re-train the framework that is trained with simulated data with field data, and to adapt the framework to be inductive.



S1017. Prediction of Total Phosphorus Concentration based on Multi-Modal Deep Learning for Four Major Rivers of the Republic of Korea

Min Kim (Severe Storm Research Center, Ewha Womans University),
Hye Won Lee (Center for Climate/Environment Change Prediction Research, Ewha Womans Universit),
Jung Hyun Choi (Department of Environmental Science and Engineering, Ewha Womans University),
Baehyun Min (Ewha Womans University)
Room: R7 2023-08-09 11:40

This study utilizes a multi-modal deep learning approach to predict the concentration of total phosphorus (TP) in the Republic of Korea's four major rivers (i.e., Han, Nakdong, Kuem, and Yeongsan rivers). TP is known to cause eutrophication, which can adversely affect aquatic ecosystems. Land use and meteorological data from 748 stations were collected between March 2018 and November 2022, excluding winter months. The input features include coordinate-based land-use data, monthly rainfall, temperature, wind speed data, and the mean and standard deviation of TP concentration data from March 2013 to November 2017. The coordinate-based land-use data are classified into seven groups (i.e., urban or built-up, agriculture, forest land, rangeland, wetlands, barren, and water). The multi-modal deep learning structure is employed to train the input features effectively. The model's architecture consists of embedding, convolutional, and bidirectional LSTM layers. The model's average performance over 45 months is characterized by an outstanding R2-value of around 0.71 (train-set of 0.73, validation-set of 0.71, and test-set of 0.70). Based on the numerical results, this multi-modal deep learning model can be utilized for TP concentration monitoring across Korea.

S11 Knowledge graph, knowledge-infused machine learning, and neuro-symbolic AI in geosciences

Xiaogang Ma (University of Idaho),
Anirudh Prabhu (Carnegie Institution for Science),
Xiang Que (University of Idaho),
Chengbin Wang (China University of Geosciences, Wuhan)

Knowledge graph, as an umbrella topic for works on ontologies, controlled vocabularies, and knowledge engineering, has received significant implementation in various disciplines of geoscience. Very recently, in the field of artificial intelligence (AI) and data science, the topics of knowledge-infused machine learning and neuro-symbolic AI arise as cutting-edge areas, which aim to combine the rule and logic-based approaches with the state-of-art machine learning methods. Knowledge graph is a representative method in the rule and logic-based approaches, but many previous studies were in the domain of data curation and integration, and there was limited work on infusing knowledge graphs with computational processes such as spatial statistics and machine learning. The new focus on knowledge-infused machine learning and neuro-symbolic AI provides a venue for researchers to review and reflect the role of knowledge graph in data-intensive geoscience studies. This session welcomes submissions on perspectives, methods, and case studies of those topics mentioned above.

S1101. Accelerating computation of non-Euclidean distances in implicit models with pathfinding heuristics

Luiz Gustavo Rasera (University of British Columbia (UBC)),
Jeff Boisvert (University of Alberta),
Shaun Barker (University of British Columbia)
Room: R5 2023-08-08 14:00

Geological structures such as folds, veins, and channel systems have non-linear geometries, rendering the Euclidean norm unsuitable for measuring spatial autocorrelation between data points and grid nodes along these structures. While multidimensional scaling can transform non-Euclidean spaces into a Euclidean metric, recent studies have shown that it can introduce biases in distance estimates and subsequent predictions. Pathfinding algorithms, such as Dijkstra's algorithm, are effective at computing shortest path distances in complex non-Euclidean spaces. However, Dijkstra's algorithm can be computationally expensive when dealing with dense graphs and sparse conditioning data, which are often encountered in geological modeling. To address this issue, we investigate the application of the A* search algorithm, a heuristics-based pathfinding method, for computing non-Euclidean distances in the estimation of locally varying anisotropy implicit models. A* search relies on heuristics to achieve better performance, making it particularly well-suited for estimation based on dense graphs with sparse data, as well as on search neighborhoods with a limited number of data points. We demonstrate the approach with illustrative examples and evaluate different heuristic functions.

S1102. Machine Learning on Zircon Trace Elements Chemistry Reveals Onset of Plate Tectonics since Hadean

Guoxiong Chen (China University of Geosciences)
Room: R5 2023-08-08 14:20

The tectonic affiliation and magma compositions that formed Earth's earliest crusts remains hotly debated. Previous efforts on this goal rely heavily on low-dimensional geochemical approaches developed from Phanerozoic samples, which are inadequate for capturing systematic differences and ignore secular changes in zircon composition. Here we apply high-dimensional machine learning (ML) approaches with zircon chemistry big data (spanning 19 elements over 4.0 billion years) to characterize zircons crystallized in some typical tectonic settings (e.g., arcs, plume-related hotspots, and rifts) and from either igneous (I-) or sedimentary (S-) type magmas. The proposed semi-supervised ML method, from a non-uniformitarian perspective, identifies the tectonic settings and granitoid types of given zircons (from Archean to Phanerozoic) at a higher prediction accuracy of >89% compared to 66-82% for traditional discriminant diagrams (e.g., U/Yb vs Y and REE+Y vs P). The ML-based discriminator depends on the systematic differences in zircon chemistry, notably for significant differences in U, Th, and Er for tectonic settings, while P, Hf, and Eu/Eu* for I- and S-type magmas. Applying the trained ML models to Hadean zircons from Jack Hills suggest that these zircons were mainly crystallized in continental arcs forming magmas with contributions partly from sediment melting, supporting the operation of plate tectonics since the Hadean.

S1103. Knowledge-guided visual exploratory analysis of open mineral data

Xiaogang Ma (University of Idaho),
Xiang Que (University of Idaho),
Jiyin Zhang (University of Idaho, USA),
Jingyi Huang (University of Idaho, USA),
Weilin Chen (University of Idaho, USA)
Room: R5 2023-08-08 14:40

The campaign of open data and open science has led to the establishment of many open geoscience data portals in recent years. In the field of mineralogy, the Mindat database (mindat.org) is one of the largest resources sharing records of mineral species, their chemical and physical attributes, their occurrences on Earth, as well as the geological, geochemical, and petrological background information. Recently, the OpenMindat research project is building an Application Programming Interface (API) to provide machine accessibility to a large portion of the existing records on Mindat. Using the OpenMinda API we have recently built several applications to conduct visual exploratory data analyses of co-relationships in the mineral data, such as those between elements and mineral species, and between localities and mineral species. Our intention of the data exploration is not to make breakthroughs in mineralogy. Instead, we treat data exploration as a way to tackle the big geoscience data, and to find clues and initial patterns in the data that can lead to more discussions and initiate hypothesis building, and we have successfully collaborated with mineralogists and geologists to implement several use cases in our OpenMindat data exploration activities. This presentation will report the recent progress. The work is supported by the National Science Foundation in USA (#1835717 and #2126315).

S1104. Loop Iteration-Based Different-Grained Geological Entity Recognition: A Case of Porphyry Copper Deposit

Chengbin Wang (China University of Geosciences, Wuhan),
Yu Xiong (China University of Geosciences)
Room: R5 2023-08-08 15:00

Unstructured geological literature is a crucial carrier of geologists' study results and hides geological knowledge for the geologist to understand the geological processes. Information mining of geological domain text data utilizing natural language and other technologies may increase the secondary development and usage of geological text large data and encourage the discovery of new domain knowledge and the advancement of geological research. The quality of geological knowledge extraction and knowledge graph creation in unstructured geological texts is determined by named entity recognition, which is one of the fundamental research contents of text data mining in the geological domain. This paper proposes a loop iteration-based method to extract the geological entities from geological literature. We used Bi-LSTM+CRF neural network to build the basic model of named entity recognition and use the loop iteration method to obtain three levels of entity labels step by step. The F-score value of our model is 84.48%, which is 13.47% higher than that of the traditional model.

S1105. Modelling and Simulating the Built-up expansion in West Delhi through an integrated Machine Learning and ABM based approach.

Aviral Marwal (University of Cambridge)
Room: R5 2023-08-08 15:20

The expansion of built-up areas is a multifaceted phenomenon that is influenced by a variety of factors that vary across both space and time. With rapid urbanization in developing countries, it has become essential to monitor the expansion of built-up areas for the purposes of effective urban planning. This study employs ML-ABM-based modeling and simulation techniques to predict the expansion of built-up areas in the West Delhi region of India in 2041. The study initially performs land use classification using the random forest classifier in Google Earth Engine, utilizing Landsat imagery from 2001, 2011, and 2021. To simulate future land use imagery, the study employs the Multi-Layer Perceptron (MLP) Neural Net algorithm in Terrset Software. The model is trained and tested using the 2001 and 2011 land use images along with six independent variables relating to built-up density and proximity to amenities. The tested model demonstrates a high accuracy rate of 87.8%, which is then combined with the Cellular Automata - Markov Chain model to simulate the land use of 2021. The resulting simulated image is validated using the reference land use map of 2021, with the validation accuracy, as measured by the kappa coefficient, being 95.6%. These results demonstrate a high level of accuracy in the simulation. The model is then employed to predict land use in 2041 under a business-as-usual approach. The study subsequently employs an agent-based model to demonstrate how the predicted land use of 2041 can be modified under two different scenarios: prioritizing growth near the district center and prioritizing growth near rural areas. The predicted land use of 2041 demonstrates distinct built-up density distributions under these two objectives, highlighting the usefulness of ABM in modifying the ML predictions based on desired policy objectives. The results of this study can be utilized for effective resource distribution planning in different locations of the study area.

S12 Novel geophysical data integration methods to assess natural resources

Katherine Silversides (DARE Centre, University of Sydney),
Guillaume Pirot (The University of Western Australia),
Jeremie Giraud (GeoRessources Lab, University of Lorraine)

The session focuses on applying novel methods to geophysical data integration for the exploration and characterization of natural resources including mineral deposits, water, hydrogen and geothermal energy. Proposed techniques may include multi-disciplinary integration approaches and applications of machine learning to geophysical integration problems.

S1201. How good can we classify geological units from airborne geophysics dataset using multiple-point statistics?

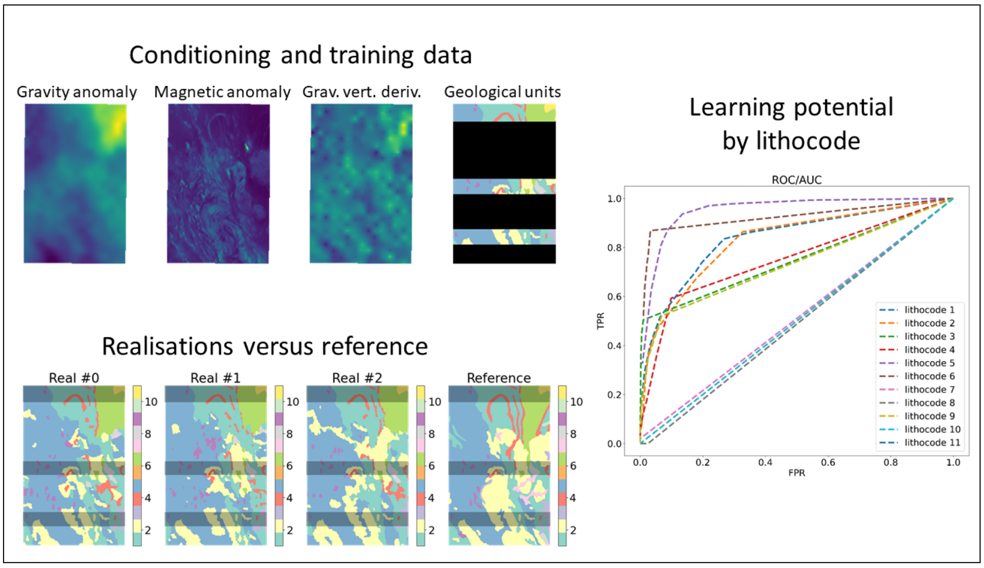
Guillaume Pirot (The University of Western Australia),
Leonardo Luiz Portes dos Santos (The University of Western Australia),
Mark W Jessell (The University of Western Australia, School of Earth Sciences, Crawley, WA 6009)
Room: R5 2023-08-07 10:00

Geological mapping and interpretation rely mainly on the availability of accessible or visible outcrops and becomes challenging in areas with high surface cover, and very limited or inexistent outcrop or exploration drillhole data. One possibility to mitigate these limitations is to consider geophysical datasets. In particular, airborne geophysics provide a coverage over large areas, and the increasing accessibility of such datasets (e.g. https://portal.ga.gov.au/) should not be ignored if it can improve geological characterization.

Here we propose to assess the potential of multiple-point statistics to identify geological units from geophysical anomaly and data maps. Indeed, Mariethoz et al. (2010) showed the potential of using multiple-point statistics for the reconstruction of incomplete datasets or images (https://doi.org/10.1007/s11004-010-9270-0). The method assumes that a prior geological map can be defined in relationship with corresponding geophysical images as a training image. This prior can eventually be defined from analogs if not enough data is available at the site of interest.

The method is illustrated on an example close to the Mount Isa in north-west Queensland, Australia. The quality of the classification is assessed against classical metrics such as accuracy, precision, recall, F1-score, Matthews correlation coefficient or Receiver Operator Characteristic curves.

We acknowledge the support from the ARC-funded Loop: Enabling Stochastic 3D Geological Modelling consortia (LP170100985) and the Mineral Exploration Cooperative Research Centre. This is MinEx CRC Document 2023/**.



S1202. Comparing clustering methods for detecting geological structure and domains from airborne geophysics

Katherine Silversides (DARE Centre, University of Sydney),
Mark Lindsay (CSIRO Mineral Resources, Australian Resources Research Centre),
Mark Jessell (The University of Western Australia, School of Earth Sciences, Crawley, WA 6009),
Arianne Ford (Geoscience Australia, Canberra, ACT 2609)
Room: R5 2023-08-07 10:20

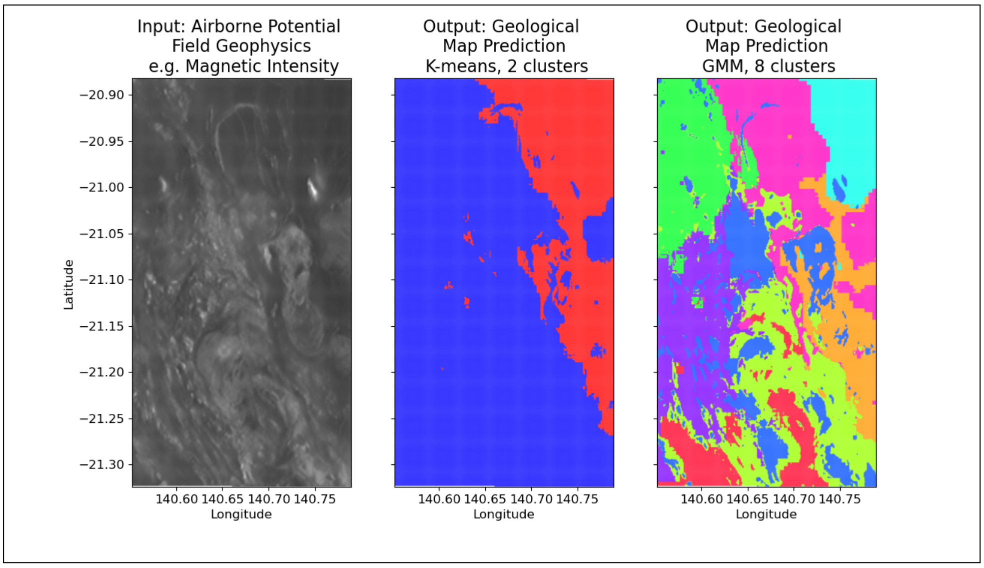
Airborne geophysical methods provide a practical method of exploring geology over large areas. However, in areas with high surface cover and little to no outcrop or previous exploration to correlate signal to rocks, results can be particularly hard to interpret. This project aimed to aid interpretation by estimating the number of different geological units, and their approximate location, using only airborne potential field geophysics.

The clustering methods used included both standard clustering methods (fuzzy c-means, self-organising maps and k-means) and Gaussian Mixture Models. Three validation metrics (Davies-Bouldin index, silhouette coefficient and Calinski-Harabasz index) were used to determine the optimum number of clusters for each method and to compare methods.

The test site is a well-studied area near Cloncurry, in north-west Queensland, Australia, where extensive surface and interpreted solid geology mapping was available for comparison to the clustering results. As the aim was to develop a method for regions without mapping, the geology was not used as an input feature. The potential field geophysical data used was reduced to pole total magnetic intensity, spherical cap Bouguer gravity anomaly, and the first vertical derivative of the gravity, along with the coordinates from publicly available datasets at Geoscience Australia.

The validation metrics frequently suggested a two-cluster result was optimal, which roughly splits the area along the Cloncurry Fault which divides the region into the Soldiers Cap Group to the east and the Staveley Formation to the west. Including the coordinates in the clustering increased the spatial consistency in the maps and often increased the optimal number of clusters chosen. However, it also produced a larger variation in the number of clusters chosen by the different validation metrics. K-means consistently outperformed the other traditional clustering methods.

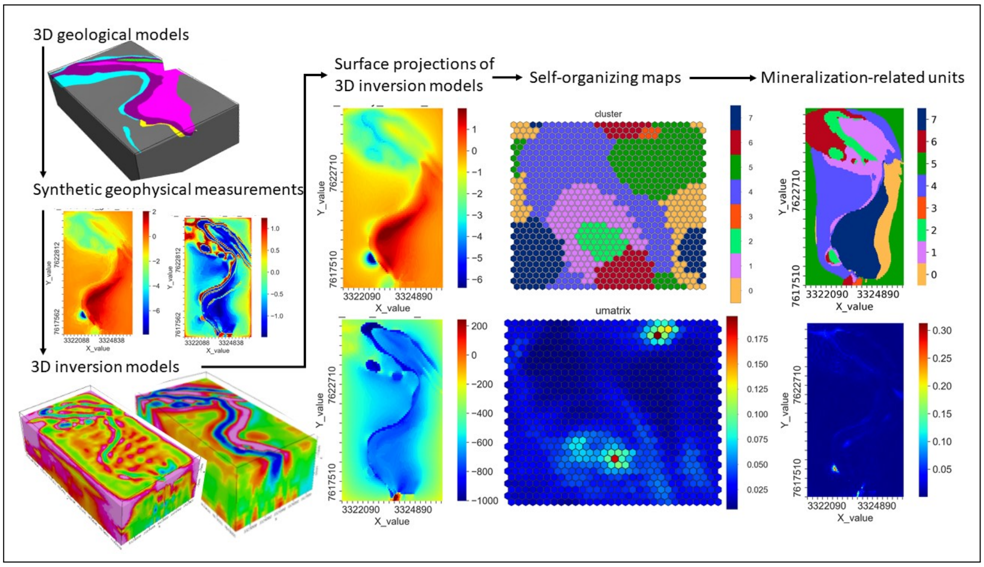
This study demonstrates a method that can be used in to aid in geological mapping in areas with laterally extensive cover. It provides the user with a range of estimates of the number of different rock types in the area, and may highlight features such as faults or distinct geological domains. The demonstrated approach provides an aid for geological and geophysical interpretation to guide more targeted mineral exploration.


Input: Airborne Potential Field Geophysics e.g. Magnetic Intensity
Output: Geological Map Prediction K-means, 2 clusters
Output: Geological Map Prediction GMM, 8 clusters
Latitude
−20.90
−20.95
−21.00
−21.05
−21.10
−21.15
−21.20
−21.25
−21.30
140.60 140.65 140.70 140.75
Longitude

S1203. Self-organizing maps for integration of 3D geophysical datasets: Applications to exploration of mineral deposits.

Bijal Chudasama (Geological Survey of Finland),
Noora Thurman (Geological Survey of Finland),
Johanna Torppa (Geological Survey of Finland)
Room: R5 2023-08-07 10:40

This study demonstrates integration of 3D geophysical data using self-organizing maps for exploration targeting of mineral deposits. The study is implemented using the geological models of the Hietakero area located in Enontekiö, northern Finland. The study area has geological settings favourable for hosting orthomagmatic mineral deposits such as Ni-Cu and PGE-deposits. The 3D geological models of the study area were used to simulate synthetic gravity and magnetic measurement data. Inversion modelling was implemented on the measurement data to generate 3D susceptibility and density models. Unsupervised clustering of surface projections of the inverted data was implemented using self-organizing maps (SOM) and subsequent k-means clustering. Results show that the ore bodies and mineralized units usually have anomalous responses in geophysical data in terms of susceptibility and density values. Hence these units are well clustered to similar SOM neurons and in subsequent k-means clusters. Although self-organizing maps method is unsupervised, labelled data of mineral deposits facilitate identification of SOM nodes and clusters representative of mineralization in the study area. The results indicate that SOM is efficient method for identifying mineralization-related patterns in multivariate data and it also facilitates better understanding of the information content of the dataset.



S1204. Interpretability and scalability of geophysics inversion in machine learning with Haralick texture features

Leonardo Luiz Portes dos Santos (The University of Western Australia),
Guillaume Pirot (The University of Western Australia),
Michel N Mamboukou (The University of Western Australia),
Jeremie Giraud (GeoRessources Lab, University of Lorraine),
Edward Cripps (The University of Western Australia),
Mark Jessell (The University of Western Australia, School of Earth Sciences, Crawley, WA 6009)
Room: R5 2023-08-07 11:00

Increasing computer power and advances in data analytic techniques have contributed to recent interest in how very large datasets of synthetic geological models, with their respective geophysical responses and metadata, could be used in real-world applications. For instance, they can be used to train and validate machine learning and geophysical inversion approaches and open a potential pathway to reducing the model space for interpreting complex natural images. Recent studies have applied complex and large machine-learning schemes for classification, parameter extraction and inversion. On the one hand they produce promising results. On the other hand, such approaches tend to penalise interpretability and scalability.

Here we explore the use of Haralick features to reduce the 40x40 pixel magnetic response images down to 13 features per model for a set of 50K 3D synthetic geological models composed of simple combinations of geological events (tilts, faults and folds). Results are two-fold. Firstly, we show that the Haralick feature representation provides clear and interpretable information about systematic structures within the magnetic images. This is done using t-Distributed Stochastic Neighbor Embedding (t-SNE), which shows clustering according to event history and can even discriminate some event characteristics, such as fault dip. Secondly, we train simple, lightweight machine-learning classifiers (support vector machine, k-nearest neighbours and random forest) to extract first-order information about the ordering of structural events of the underlying 3D geological model from the magnetic images. While performance is similar to those from very deep convolution neuron networks, the computational effort is smaller in order of magnitudes.

Based on these results we plan to expand the model suite to more geologically complex suites of models, and to combine gravity and magnetic images, so that we can apply this approach to more general geological settings.

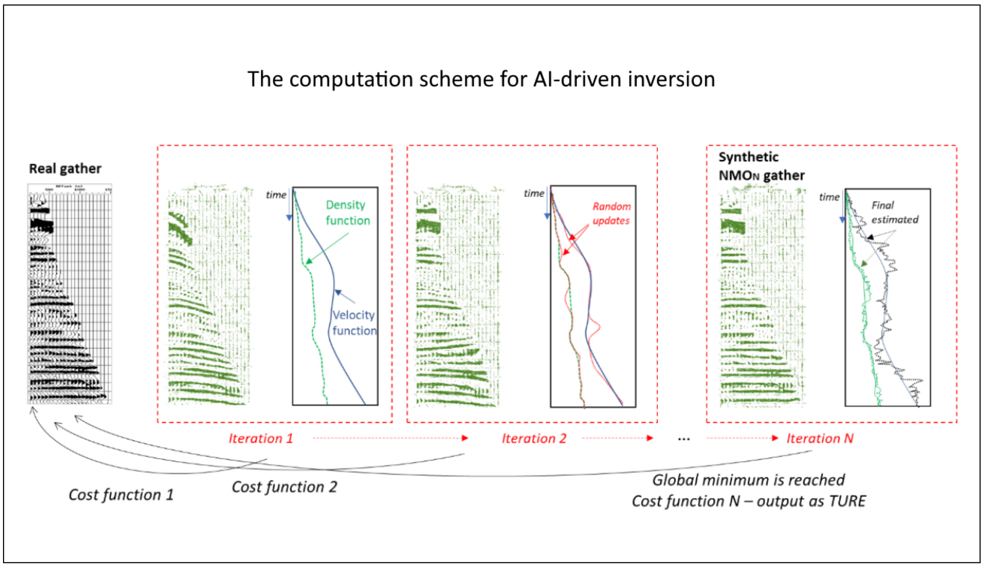
S1205. Non-regression pre-stack AI-based seismic inversion algorithm

Rune Øverås (Pre Stack Solutions - Geo AS),
Vita Kalashnikova,
Barbara Eva Klein (1)
Room: R5 2023-08-07 11:20

Rock properties can be estimated by performing a seismic inversion. Seismic pre-stack inversion approaches are based on the fundamental assumptions: linearisation of the Zoeppritz equation, estimation of impedances or elastic parameters and extraction of the linked velocity and density through regressions. The regression limitations come from an ill-posed problem that gives equivalent solutions for a large ensemble of rock properties. To reduce uncertainties in the regression approaches, statistical and advanced clustering techniques are commonly used. However, the environment does not obey regression laws everywhere.

We propose to overcome the regression limitations by solving the problem radically differently. Here we demonstrate that combining metaheuristic procedures, designed artificial constraints, and computation-efficient schemes reveals the unexampled possibility of estimating non-linked rock properties from the seismic data.

For each inversion iteration, we individually update the model parameters of Vp, Vs and Dn in a random scheme. Synthetic traces are calculated for each iteration, and we compute a global misfit between the synthetic and real data. In our approach, the P-wave velocity is constrained by kinematics. To avoid convergence in local minima, we use a global optimization scheme. The inversion is not angle limited by NMO stretch and linearization of the Zoeppriz equation; thus, we can use data up to refraction energy. This inversion approach requires many iterations and is, therefore, computer-intensive — different solutions can be used to overcome it. For post-stack inversion, when estimating acoustic impedance, the approach requires N iterations (n computations). For pre-stack inversion of an offset gather with K traces, also a factor F for calculation of the kinematics implies. Then searching for Vp, Vs, and Dn, the approach requires approximate N 3 iterations (f*k*n 3). We decrease the number of computations by choosing a hybrid scheme. First, we do a two-parameter search for Vp and Dn. Then, as an example, we fix a Vp parameter and perform a two-parameter search for Vs and Dn. This scheme could be repeated j times differently, reducing the computation cost to j*f*k*n 2. We demonstrate that this approach yields benefits in the rock properties estimation from seismic data.


The computation scheme for AI-driven inversion
Real gather
Synthetic
NMON gather
time
Density
function
Velocity
function
Random
updates
Final
estimated
Iteration 1
Iteration 2
...
Iteration N
Cost function 1
Cost function 2
Global minimum is reached
Cost function N – output as TURE

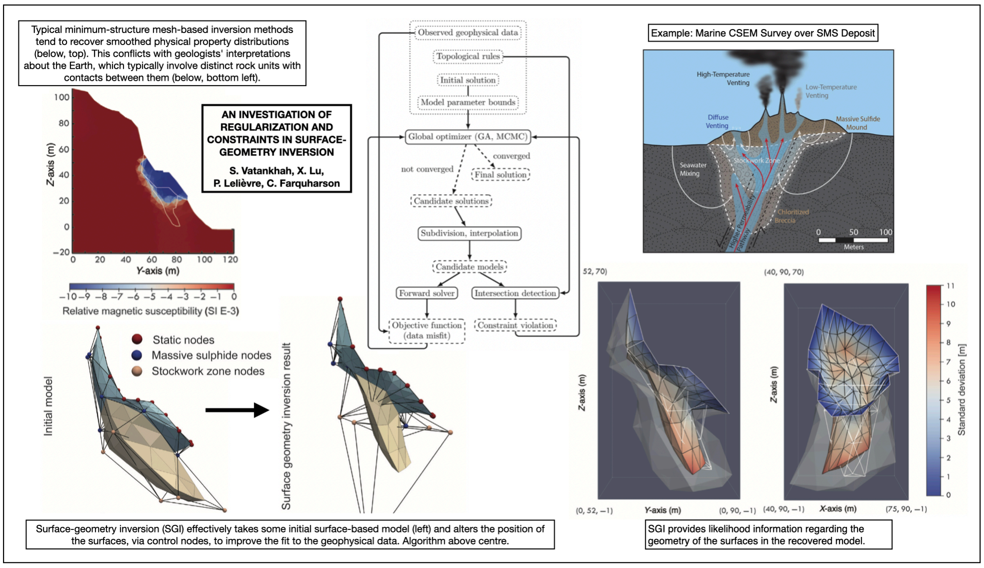
S1206. An investigation of regularization and constraints in surface-geometry inversion

Saeed Vatankhah (Mount Allison University),
Xushan Lu (Memorial University of Newfoundland),
Peter Lelievre (Mount Allison University),
Colin Farquharson (Memorial University of Newfoundland)
Room: R5 2023-08-07 11:40

Typical minimum-structure mesh-based inversion methods tend to recover smoothed physical property distributions. This conflicts with geologists' interpretations about the Earth, which typically involve distinct rock units with contacts between them. While focussing and other regularization methods can help, working on a mesh is always inconsistent with the way geologists think of the Earth. Hence, we have been developing a fundamentally different inversion approach that works directly with surface-based representations of the Earth volume of interest, where surfaces of tessellated triangles represent the interfaces between rock units. We call this approach surface-geometry inversion (SGI). SGI effectively takes some initial surface-based model and alters the position of the surfaces to improve the fit to the geophysical data.

It is important to know whether the solutions obtained from SGI are unique and stable and, if they are not, how to add regularization or constraints to make them so. Without a well-posed problem, any interpretations of the subsurface based on those solutions, and any exploration decisions based on those interpretations, are unreliable. Assessing the numerical characteristics of SGI problems is challenging because we use global heuristic optimization methods and stochastic sampling in their solution, they are severely nonlinear, and they lack explicit matrix operators and derivatives. A critical aspect is understanding when regularization/stabilization should be incorporated into the SGI optimization problem to create a well-posed problem. In this work, we make headway towards a better understanding of these issues.

The extent to which our SGI problems are well-posed likely depends on the number of data, the resolving capability of the data type and survey geometry, the signal-to-noise ratio, the tessellation of the surface-based models, the subdivision strategy, and the size and shape of the anomalous targets. These are difficult questions to address: general rules are difficult to develop with rigorous mathematics given the nature of our SGI problems and solution methods. Hence, in this preliminary stage we develop an empirical methodology to assess the numerical characteristics of some specific, common exploration scenarios.


Typical minimum-structure mesh-based inversion methods tend to recover smoothed physical property distributions (below, top). This conflicts with geologists' interpretations about the Earth, which typically involve distinct rock units with contacts between them (below, bottom left).
AN INVESTIGATION OF REGULARIZATION AND CONSTRAINTS IN SURFACE-GEOMETRY INVERSION
S. Vatankhah, X. Lu, P. Lelièvre, C. Farquharson
Z-axis (m)
Y-axis (m)
Relative magnetic susceptibility (SI E-3)
Static nodes
Massive sulphide nodes
Stockwork zone nodes
Initial model
Surface geometry inversion result
Observed geophysical data
Topological rules
Initial solution
Model parameter bounds
Global optimizer (GA, MCMC)
not converged
converged
Final solution
Candidate solutions
Subdivision, interpolation
Candidate models
Forward solver
Intersection detection
Objective function (data misfit)
Constraint violation
Example: Marine CSEM Survey over SMS Deposit
High-Temperature Venting
Low-Temperature Venting
Diffuse Venting
Massive Sulfide Mound
Stockwork Zone
Seawater Mixing
Chloritized Breccia
Higher Permeability Pathway
Meters
X-axis (m)
Standard deviation [m]
Surface-geometry inversion (SGI) effectively takes some initial surface-based model (left) and alters the position of the surfaces, via control nodes, to improve the fit to the geophysical data. Algorithm above centre.
SGI provides likelihood information regarding the geometry of the surfaces in the recovered model.

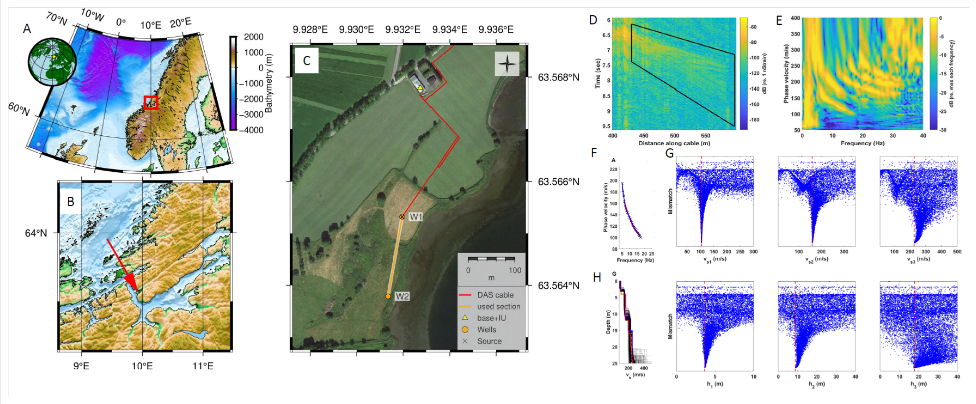
S1207. Quick clay monitoring using surface wave recorded on a distributed acoustic sensing array.

Robin Andre Rørstadbotnen (NTNU),
Hefeng HD Dong (Norwegian University of Science and Technology, Department of Electronic Systems),
Martin ML Landrø (Norwegian University of Science and Technology),
Kenneth KD Duffaut (Norwegian University of Science and Technology, Department of Geoscience and Pet),
Kevin KG Growe (Norwegian University of Science and Technology, Department of Electronic Systems),
Umedzhon Kakhkorov (NTNU),
Susann SW Wienecke (Alcatel Submarine Networks Norway AS; Tiller, Norway),
Joacim JJ Jacobsen (Alcatel Submarine Networks Norway AS; Tiller, Norway)
Room: R5 2023-08-07 12:00

Quick clay avalanche is one of the most devastating landslide types worldwide, and an early-warning system is in demand to mitigate the fatal consequences caused by such events. Distributed Acoustic Sensing (DAS) data were acquired on a dedicated fiber in an area containing quick clay deposits to address this (see Figure 1A-C). The data were recorded between July 2021 and February 2022 in Rissa, Norway, while a new road was constructed on the quick clay. Road construction can induce changes to the clay's mass balance which has previously been shown to trigger quick clay landslides. For this purpose, active and passive data were collected to test and compare various analysis methods. The active data contained sledgehammer blows on a metal plate. We focused on the Rayleigh waves generated by these sledgehammer shots and used their dispersive behavior to estimate shear-wave velocity depth profiles.

Furthermore, the environmental noise data is used first to enhance the Rayleigh waves in the background noise and then to estimate shear-wave velocity profiles by ambient noise interferometry. To estimate the shear-wave velocity profiles, two different inversion algorithms were used, one linearized based on singular-value-decomposition and one non-linear known as Adaptive Simplex Simulated Annealing. Figures 1D-G show the estimation procedure. In Figure 1D, data from a sledgehammer shot is shown, and Figure 1E depicts the associated dispersion image. Figures 1F-G show the non-linear inversion in practice. Thus, this study provides two possible data collection methods and two possible inversion methods for an early-warning system.

The obtained dispersion curves and the estimated shear-wave velocity profiles indicate minor time-lapse variation during the seven-month acquisition period. The study finds that the road construction work, and the extra load added to the quick clay, do not alter the quick clay's behavior and hence do not change the quick clay attributes within the monitoring period. Nevertheless, the results provide reference shear-wave velocity trends for the study area in undisturbed conditions. This work is the first step in developing the early-warning system for quick clay landslides using fiber optics.



A
B
C
D
E
F
G
H
70°N
60°N
64°N
10°W
0°
10°E
20°E
9°E
10°E
11°E
Bathymetry (m)
9.928°E
9.930°E
9.932°E
9.934°E
9.936°E
63.568°N
63.566°N
63.564°N
W1
W2
DAS cable
used section
base+IU
Wells
Source
Distance along cable (m)
Frequency (Hz)
Mismatch

S1214. Trans-dimensional 3D geometrical inversion: proof of concept and field application using gravity data from the Boulia region (Queensland, Australia)

Jeremie Giraud (GeoRessources Lab, University of Lorraine),
Guillaume Caumon (GeoRessources Lab, University of Lorraine),
Lachlan Grose (Monash University),
Vitaliy Ogarko (The University of Western Australia),
Julien Herrero (GeoRessources Lab, University of Lorraine),
Radu Stoica (University of Lorrain)
Room: R5 2023-08-07 10:40

We present and apply a trans-dimensional inversion method for 3D gravity inversion. This method builds on a multiple level set inversion approach using signed-distance functions to model the location of interfaces between rock units. To account for the unknown number of rock units that need to be modeled, we use a death and birth process where rock units can be inserted into or removed from an existing geological model. In addition, we also invert for the geometry of rock units and their densities. We test our method using a synthetic model before applying it to field data from the prospective Boulia region (Queensland, Australia) to image rock formations under sedimentary cover. In this field application, we start from an implicit geological model derived from the interpretation of 2D seismic lines, borehole data, and geological rules that presents a relatively strong gravity data misfit. Preliminary results indicate that up to 2 rock units that were not initially identified may need to be added to the model. We argue that this shows the ability of our method to infer the presence of unseen geological features such as inclusions or facies variations., Guillaume Caumon, Lachlan Grose, Vitaliy Ogarko, and Radu Stoica

S13 Data Integration

Jo Eidsvik (NTNU),
Thomas Mejer Hansen (Aarhus University),
Klaus Mosegaard (Niels Bohr Institute, University of Copenhagen)

Many geoscientific challenges require the integration of information from diverse sources, such as geology, geophysics, geochemistry, and hydrology. This calls for methods that allow a) quantifying information, and b) integration of information. Once all information has been combined, it needs to be made available to the end-users, e.g. the decision makers, in a way that allows taking an informed decision based on all available information. For this session, we invite contributions that deal with any of these aspects related to data integration in geoscience, with the goal of allowing decision-makers to take an informed decision. These contributions can for example be based on inverse problem theory, geostatistics and machine learning.

S1301. Probabilistic groundwater modelling for decision-makers with neural networks

Mathias Busk Dahl (Aarhus University),
Rasmus Bødker Madsen (Overfladenær Land- og Maringeologi),
Troels Norvin Vilhelmsen (Niras, Ceres Allé 3, Aarhus, Denmark),
Thomas Mejer Hansen (Aarhus University)
Room: R7 2023-08-07 10:00

Realistic groundwater models are essential for decision-makers in a variety of topics related to groundwater exploration and management. These models are applied together with flow modelling methods to analyze for example the effect of pumping and the propagation of pollution. In practice, most groundwater models are deterministic, representing a single set of hydrological parameters. Deterministic models offer simplicity and are computationally efficient to handle, but a single model of hydrological properties is not adequate to represent the actual uncertainty of the subsurface. In contrast, probabilistic-based groundwater models can represent the full uncertainty of the subsurface. By generating multiple examples of hydrological models from such a probabilistic model, it is possible to propagate the uncertainty of the subsurface hydrological parameters to uncertainty in the hydrological response simply by running flow modelling on each of the generated realizations. This will allow decision-makers to take decisions, based on all the information combined in the probabilistic groundwater model.

Two main challenges arise with such a probabilistic framework for decision-making. 1) One needs access to a probabilistic groundwater model. While this is not common, more focus has recently been put on the potential use of probabilistic models. 2) Flow modelling, even on a single 3D hydrological model, may be computationally demanding, and simple Monte Carlo propagation of uncertainty using multiple flow modelling runs on multiple realizations of a probabilistic groundwater model may not be feasible.

We propose to construct a neural network trained on computationally expensive flow modelling results for efficient computations of some features related to the probabilistic hydrological model. Once trained, this network will provide an efficient and practical tool for probabilistic groundwater management. We present a method for training the network and compare its predictions to results from deterministic and probabilistic numerical groundwater modelling results. Essentially, we show a way to incorporate the subsurface uncertainty of a groundwater model in model results within a reasonable timeframe for decision-makers.

S1302. Coupling an ensemble smoother with a truncated Gaussian model for aquifer characterization

Valeria Todaro (University of Parma),
Marco D'Oria (University of Parma),
Andrea Zanini (University of Parma),
J. Jaime Gómez-Hernández (Universitat Politecnica Valencia),
Maria Giovanna Tanda (University of Parma)
Room: R7 2023-08-07 10:20

Aquifers are an essential source of water supply for human consumption, agricultural irrigation, and industrial purposes. Understanding their complexity is crucial for sustainable groundwater management. However, estimating the properties of the aquifer and their spatial variability can be challenging, but it is necessary to accurately simulate groundwater flow and transport. In this study, an inverse approach is applied to infer the characteristics of a binary field using concentration data obtained from a tracer test carried out in a laboratory sandbox. The experimental device reproduces a vertical section of an unconfined aquifer. Glass beads of two different diameters are used to mimic the binary porous medium, while fluorescein sodium salt served as a tracer. Breakthrough curves recorded at different locations are used as observations to estimate the spatial distribution of two lithotypes and their main hydraulic and transport parameters. The inverse problem is solved by coupling the ensemble smoother with multiple data assimilation (ES-MDA), which is an ensemble Kalman-based method, with a truncated Gaussian model (ES-MDA-T). The advantage of the ES-MDA-T, with respect to the standard ES-MDA, is the possibility to simultaneously estimate different properties of the geological lythotypes and their spatial distribution with an acceptable amount of computational time. The proposed technique is tested using two methods: a fully parameterized scheme and a pilot point approach. In order to evaluate the capability of the methodology, a synthetic case that reproduces the sandbox experiment was initially developed. Once the procedure was validated, the optimal configuration obtained from the synthetic case was used for the experimental case study. The results show that the ES-MDA-T method replicates well the binary field and its aquifer properties even in the presence of considerable measurement errors. The pilot point approach takes less computation time than the fully parameterized scheme, while yielding comparable results.

S1303. Core And Cuttings Derived Mineralogy Logs for improved subsurface characterization

Torolf Wedberg,
Kristian B. Brandsegg (BD Coordinator, CapeOmega AS),
Craig Lindsay (Core Specialist Services, UK)
Room: R7 2023-08-07 10:40

Elemental data from X-Ray Fluorescence (XRF) analysis of core and cuttings has significant intrinsic value that is currently not being realized in its complete extent. XRF is often acquired using multi-sensor systems, where it is acquired in conjunction with a range of other data types which are used both individually and as specific data sets.

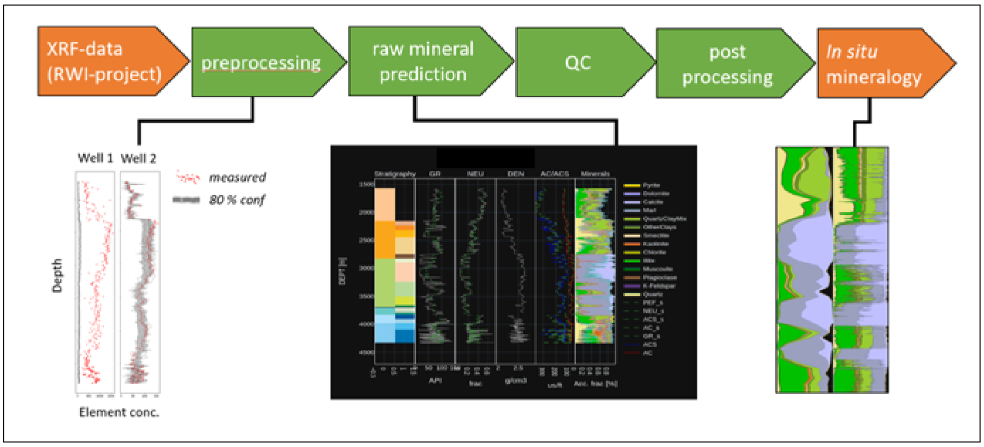
XRF elemental data can be converted to a mineral assemblage. Ideally this would be calibrated by a limited number of XRD analyses (sampling sites based upon the initial XRF log). So far, so nothing new as this approach has been available for some time.

Now we introduce a step change in ability to confidently generate a representative mineral assemblage. Artificial intelligence (AI) is used to generate synthetic well log responses (gamma, density etc.). based upon the predicted mineral assemblage. The actual well logs are then used, involving an iterative AI based refinement process, adjusting the predicted mineral assemblage to produce a best match between the actual and synthetic well logs. Thus, the predicted mineral assemblage is first calibrated from XRD and then optimized against the actual well logs, both using AI. The net result is a predicted mineral assemblage which has a much high confidence level than a standalone interpretation – a massive improvement on the standard approach.

The first quality assurance of the results through the methodology start when assessing the input data from both XRF and wireline to bring certainty its validity prior for mineralogy reassembly and in the post-prosessing. The final quality assurance is when applying the improved subsurface characterization and linking it with other data not used in the methodology. This ultimately delivers quicker and more reliable analytic results for better decisions.

An industry survey triggered an impressive response supporting our assertion that the availability of robust log of sub-surface mineralogy would have many applications and benefits across the sub-surface sector including oil and gas, CCS, mining, and geothermal.

Figure 1 shows the methodology flow from XRF-data to high resolution in situ mineralogy composition.



S1304. Comparing PINN and Numerical Solutions for Heterogeneous and Anisotropic Unconfined Aquifer Systems

Daniele Secci (Università degli Studi di Parma and Universitat Politècnica de València),
Vanessa A. Godoy (Technical University of Valencia),
J. Jaime Gómez-Hernández (Technical University of Valencia)
Room: R7 2023-08-07 11:00

Physics-Informed Neural Networks (PINNs) are a class of machine learning models that brings physical interpretability to classical neural networks (NNs) and can be adapted to solve many complex scientific and engineering problems, even in data-scarce regimes. Unconfined groundwater flow is a challenging problem that involves spatially and temporally varying boundary conditions and can only be easily solved with significant simplifications. In this work, we consider the groundwater flow partial differential equation in an unconfined aquifer, where the phreatic surface is a time-varying boundary condition of unknown geometry that should satisfy that the piezometric head equals the elevation. The equation is solved numerically using MODFLOW and using a PINN. The present study highlights that PINNs can effectively compute the piezometric values in an unconfined aquifer and the free surface in time. Finally, these results suggest that PINNs can serve as a valuable alternative to or supplement traditional numerical models in the simulation of unconfined aquifer flow problems and are an attractive option for hydrogeology modeling and simulation.

The corresponding author wishes to express his deep gratitude to the IAMG Student Affairs Committee for awarding him with the Computers and Geosciences Research Scholarship for the project “Physic-Informed Neural Networks (PINNs) for subsurface hydrology” that supported the completion of the current study.

This work was developed under the scope of the InTheMED project. InTheMED is part of the PRIMA programme supported by the European Union’s HORIZON 2020 research and innovation programme under grant agreement No 1923.

S1305. An Exponential Class of Ensemble based Optimization Algorithms

Mathias Methlie Nilsen (NORCE),
Andreas S. Stordal (NORCE),
Patrick N. Raanes (Norce),
Kjersti S. Eikrem (Norce),
Rolf J. Lorentzen (Norce)
Room: R7 2023-08-07 11:20

Stein's lemma gives a new interpretation of EnOpt (ensemble optimization). This interpretation enables us to study EnOpt in the context of an exponential class of mutation distributions. In this paper a generalized version of EnOpt (named GenOpt) is proposed with a dual update scheme where the control is updated using Stein's lemma and marginal mutation distributions are updated individually by the mutation approach. A Gaussian Copula is then used to construct a multivariate mutation distribution and also to model the correlation structure between the marginals. The correlation matrix in the copula is also adapted via the mutation gradient. It is also shown that using GenOpt with Beta distributions as marginals eliminates the truncation problem that often arises when applying EnOpt on bounded optimization problems. The performance of the proposed optimization algorithm is evaluated through several test cases, including a fictitious reservoir management case. The potential pros and cons of GenOpt is also discussed

S1306. Hydrogeophysical inversions using ensembled smoother with multiple data assimilation

Camilla Fagandini (University of Parma),
Valeria Todaro (University of Parma),
Claudia Escada (Instituto Superior Tecnico Lisboa),
Leonardo Azevedo (Instituto Superior Técnico, Lisboa),
Andrea Zanini (University of Parma),
Jaime Gómez-Hernández (Universitat Politecnica Valencia)
Room: R7 2023-08-07 11:40

Environmental pressure on groundwater systems has intensified during the past century because of the massive development of industrial and agricultural activities. The effective and safe management of the groundwater environment represents a significant challenge to modern society, requiring a detailed understanding of the systems involved. In hydrogeology, direct measurements of subsurface geology are often limited. Recently, more attention has been paid to inverse hydrogeophysics modelling for the spatial prediction of hydrogeological subsurface properties.

In this work, electrical resistivity tomography (ERT) data and pollutant concentrations measured sparsely at borehole locations were jointly used to predict the hydraulic conductivity field using the Ensemble Smoother with Multiple Data Assimilation (ES-MDA). A synthetic case that simulates a heterogeneous aquifer was developed to assess the efficacy of the approach. The ES-MDA is an iterative data assimilation approach that allows the estimation of unknown parameters using observed data and a forward model that relates model parameters and observations. One of the advantages of the ES-MDA is its capability to assimilate multiple data sources simultaneously. The hydraulic conductivity field is estimated using the ERT data and concentrations as observations in this case. The forward model is represented by laws that establish the relationship between observed data and parameters to be estimated. The method workflow begins with an initialization phase, in which an initial ensemble of parameter realizations is defined, followed by an iterative phase consisting of a forecast and update steps. During the forecast step, the forward model provides predictions corresponding to the available observations for each parameter realization. Then, the algorithm updates the ensemble of parameters based on the misfit between predictions and observations. In ES-MDA, all available observations are assimilated multiple times during the iterative process.

The results demonstrate the potential of ES-MDA for hydrogeophysical inversion using both ERT data and concentrations concurrently for subsurface characterization while accounting for the uncertainty of the predictions. Furthermore, the ES-MDA assimilates multiple data sources, which can significantly improve the accuracy of the estimated conductivity field. As a future development, it is planned to use data collected in a laboratory experiment under fully controlled conditions.

S1307. Consistency and Causality in Bayesian Inversion Paradigms

Klaus Mosegaard (Niels Bohr Institute, University of Copenhagen)
Room: R7 2023-08-07 14:00

Bayesian formulations of inverse problems are regarded as powerful tools for integration of data and prior information. However, the Bayesian inference paradigm suffers from a number of often overlooked difficulties.

The most notable difficulty is related to the use of probability densities on continuous manifolds. Investigations by Kolmogorov's (1933/1956) found that the classical definition of probability densities is inconsistent: Different parameterizations lead to different conditional probability densities (which are not related through the usual Jacobian transformation). This problem is persistently neglected in the scientific literature, especially in applications of Bayesian inversion. We will show that the problem threatens the objectivity of Bayesian computations, including Bayesian inversion, Bayes Factor computations, and trans-dimensional inversion.

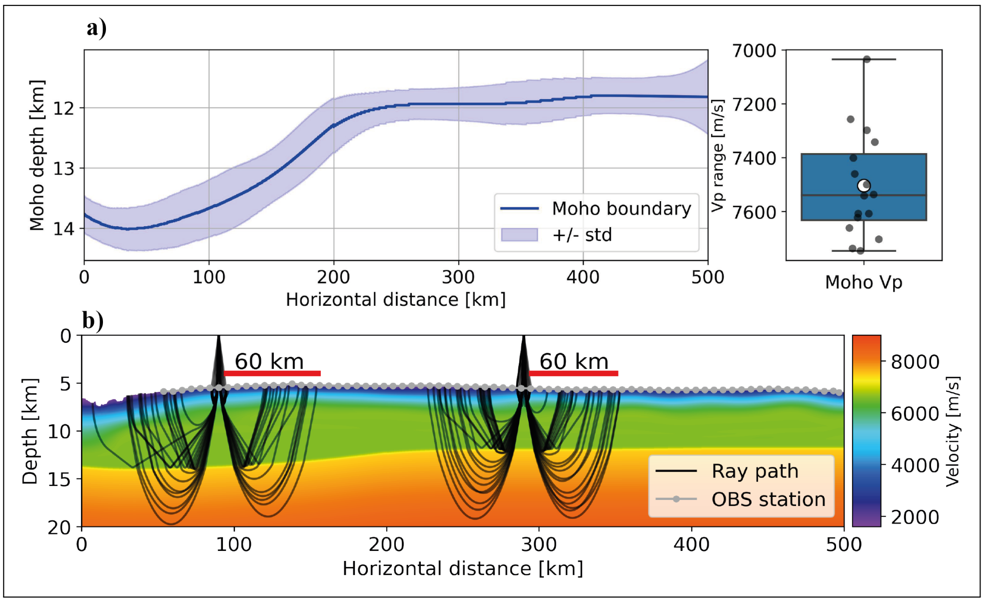
Other difficulties in Bayesian inversion derives from the fact that data uncertainties, and/or prior information are often unknown or poorly known. This has led to invention of the so-called hierarchical methods where parameters (known as hyper-parameters) are controlling the uncertainties. However, since data uncertainties and prior information on model parameters are supposed to be known 'a priori', but are calculated 'a posteriori', we face a violation of causality. We investigate the consequences of this acausality, and show how it challenges the validity of Bayesian computations.

S1308. Global inversion and parametrization for building tomographic velocity models

Umed Kakhkhorov (NTNU),
Børge Arntsen (Department of Geoscience and Petroleum, NTNU, Norway),
Wiktor Waldemar Weibull (University of Stavanger)
Room: R7 2023-08-07 14:20

Traveltime tomography is applied to investigate seismic structures of the Earth's subsurface. An accurate tomographic velocity model is important for a high-resolution waveform velocity building and its availability is one of the main components to mitigate the nonlinear inverse problem. We present a new methodology of obtaining velocity models for traveltime tomography studies. We found a way to get a highly accurate first-arrival traveltime tomography in combination with global optimization. The role of global optimization is twofold: to find initial solutions that are close to 'truth', and to guide tomographic inversion towards a geologically consistent model that explains the data. The main advantage of our workflow is a data-driven approach avoiding the use of a conventional layer-based parameterization and incorporation of manual interpretations into the velocity model (Fig. 1).

To date, a few geophysical studies have been focused on developing data-driven and a labour non-intensive regional tomographic velocity model building workflow. In our study, we present the tomographic velocity model building workflow as a combination of first-arrival traveltime tomography and global optimization. Global optimization allows searching for velocity parameters and depth to interfaces in the larger search area with a higher chance of convergence. After defining the geometry of main layers and general velocity trends, traveltime tomography with a bi-cubic B-spline model parameterization can be fitted to further update the velocity model. Our approach allows obtaining a highly accurate velocity model which can be used for seismic depth migration and as a starting model for a FWI seismic imaging. The workflow is developed and a regional Ocean Bottom Seismic data acquitted along a subduction zone.



S1309. Geophysical inversion for geological features.

Thomas Mejer Hansen (Aarhus University),
Erik Skovbjerg (Afdeling for Geofysik og sedimentære bassiner, Danmarks og Grønlands Geologiske),
Ingelise Møller (Afdeling for Grundvands- og Kvartærgeologisk Kortlægning, Danmarks og Grønlands),
Peter Sandersen (Afdeling for Grundvands- og Kvartærgeologisk Kortlægning, Danmarks og Grønlands),
Rasmus B Madsen (Overfladenær Land- og Maringeologi)
Room: R7 2023-08-07 14:40

In most instances, the purpose of geophysical inversion is to estimate subsurface geophysical properties, such as resistivity and velocity, from geophysical data. Although these properties can be somewhat interesting, they are typically not the primary focus for subsurface exploration and management. Decision-makers may be more interested in questions related to geophysical parameters, such as ”What is the thickness of Miocene Clay?” or ”What is the probability of a pollution plume’s existence?”. We will discuss methods that enable direct answers to such questions using geophysical data. Both approaches involve obtaining information about the parameters of interest related to geophysical properties.

First, we will demonstrate how a probabilistic model can describe geological knowledge, incorporating information about lithostratigraphy, sedimentology, depositional environment, and grain size. Next, we will explain how a statistical model can be developed to describe the relationship between geological and geophysical parameters, using well log analysis. We will also show how to compute the electromagnetic geophysical data response in this case. The combination of these efforts results in a joint probabilistic prior model.

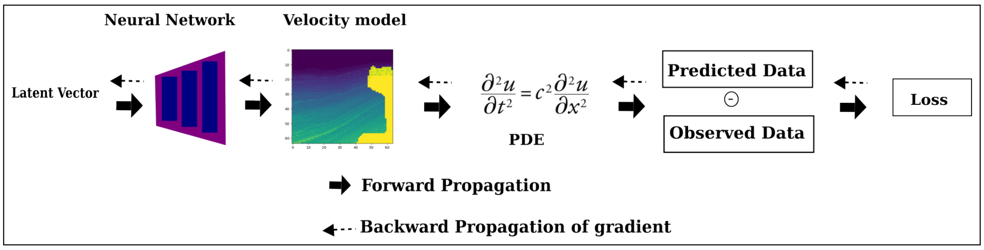
We will then demonstrate how information about any property or feature of the parameters defined in the prior model can be inferred directly from a posterior distribution based on observed electromagnetic data. This will be exemplified using both a sampling approach and a method that employs a neural network.

S1310. Generative Neural Network driven Full Waveform Inversion

Bhargav Boddupalli (NTNU),
Børge Arntsen (Department of Geoscience and Petroleum, NTNU, Norway),
Tor Andre Haugdahl (Department of Computer Science, NTNU, Norway),
Mrinal K Sen (Jackson School of Geosciences, The University of Texas at Austin, USA),
Martin Landrø (Norwegian University of Science and Technology)
Room: R7 2023-08-07 15:00

Full waveform inversion (FWI) has emerged as the most powerful subsurface imaging application for a wide range of geophysical studies. FWI is an ambitious physics-based data-fitting technique that aims to model seismic waveform in an iterative manner updating the starting model. In principle, FWI is capable of generating high-resolution models of subsurface parameters such as compressional and shear wave velocities, attenuation parameters, anisotropy parameters, etc. However, practical applications are mostly restricted to the inversion of compressional wave velocity models. This limitation is mainly because of its strong non-linearity, high computational cost, and ill-posedness; thus, making it prone to local minimum convergence during optimization and challenging to quantify uncertainties.

Deep neural networks' ability to learn spatial correlations and predict data faster makes them a promising way forward in overcoming the limitations. Although recent studies have demonstrated the benefits of integrating NN with FWI, in terms of regularization, reparameterisation and uncertainty estimation, not much work has been done in studying the architectural constraints on the neural networks for FWI. In the present work, we investigate the links between 2D elastic FWI physics and deep neural networks by experimenting with different neural network architectures using different velocity models.



S1311. A pseudo-Bayesian approach for non-linear seismic inversion

T. Tien Mai (NTNU),
Mina Spremic (NTNU),
Jo Eidsvik (NTNU)
Room: R7 2023-08-07 15:20

Given the observed seismic data $\mathbf{Y}_{obs}$, the main aim in seismic inversion is to determine the model parameters $\mathbf{m}$ such that (at least approximately)

$\mathbf{Y}_{obs} = \mathbf{F}(\mathbf{m})$,

where $\mathbf{F}$ is the forward map that is a (known or can be evaluated) non linear function. As an alternative to the Bayesian approach, where we often specify a Gaussian model for the residual

$Y_{obs} - \mathbf{F}(\mathbf{m})$,

we propose to replace the traditional evaluation of the data fit (likelihood) with a more comprehensive evaluation of a risk or a pseudo-likelihood. The benefits of the approach, termed pseudo-Bayesian approach, are:

- no stringent assumptions regarding the process that generated the data,
- reduce the cost when performing MCMC sampling from the (approximate) posterior

We consider the following Gibbs posterior, given a prior distribution $\pi(\mathbf{m})$,

$$\rho_\lambda(\mathbf{m}|\mathbf{Y}_{obs}) \propto \exp\left[-\lambda\, d(\mathbf{m}, \mathbf{Y}_{obs})\right]\, \pi(\mathbf{m})$$

where $\lambda > 0$ is a tuning parameter, $d(\cdot,\cdot)$ is a metric to measure the divergence between the observation $\mathbf{Y}_{obs}$ and a hypothesis forward model $\mathbf{F}(\mathbf{m})$. A fast distance metric evaluation is trained from simulations, and this enables effective pseudo-Bayesian analysis.

A pseudo-Bayesian approach for non-linear seismic inversion

The Tien Mai, Mina Spremic, Jo Eidsvik

Department of Mathematical Sciences,
Norwegian University of Science and Technology, Trondheim, Norway.
Email: the.t.mai@ntnu.no

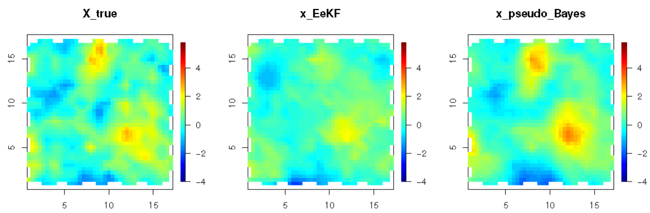
Given the observed seismic data $\boldsymbol{Y}_{obs}$, the main aim in seismic inversion is to determine the model parameters $\boldsymbol{m}$ such that (at least approximately) $\boldsymbol{Y}_{obs} = \boldsymbol{F}(\boldsymbol{m})$, where $\boldsymbol{F}$ is the forward map that is a (known or can be evaluated) non linear function. As an alternative to the Bayesian approach, where we often specify a Gaussian model for the residual $Y_{obs} - \boldsymbol{F}(\boldsymbol{m})$, we propose to replace the traditional evaluation of the data fit (likelihood) with a more comprehensive evaluation of a risk or a pseudo-likelihood. The benefits of the approach, termed pseudo-Bayesian approach, are:

- no stringent assumptions regarding the process that generated the data,
- reduce the cost when performing MCMC sampling from the (approximate) posterior

We consider the following Gibbs posterior, given a prior distribution $\pi(\boldsymbol{m})$,

$$\rho_\lambda(\boldsymbol{m}|\boldsymbol{Y}_{obs}) \propto \exp[-\lambda d(\boldsymbol{m}, \boldsymbol{Y}_{obs})]\, \pi(\boldsymbol{m})$$

where $\lambda > 0$ is a tuning parameter, $d(\cdot,\cdot)$ is a metric to measure the divergence between the observation $\boldsymbol{Y}_{obs}$ and a hypothesis forward model $\boldsymbol{F}(\boldsymbol{m})$. A fast distance metric evaluation is trained from simulations, and this enables effective pseudo-Bayesian analysis.



1

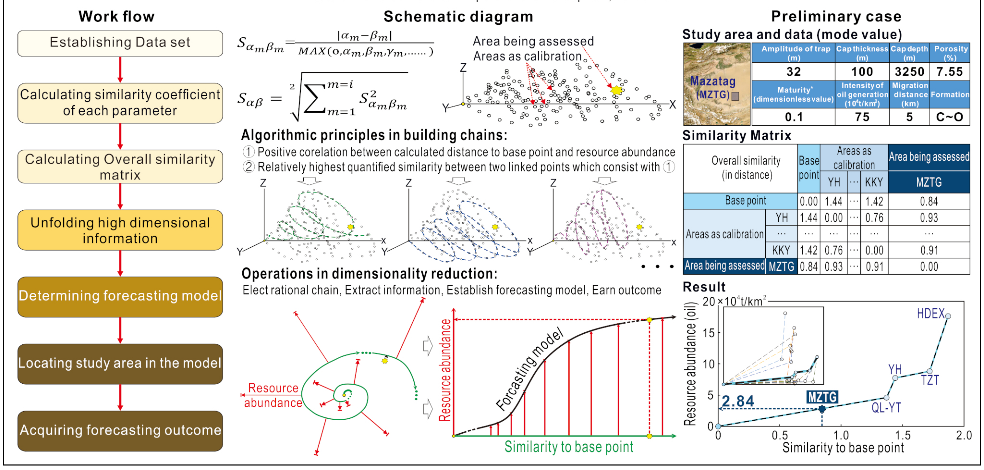
S1312. An improved method in hydrocarbon resource abundance forecast through analogism based on isometric mapping theory and its preliminary application

Jingdu Yu (Petrochina),
Qiulin Guo (PetroChina),
Shiyun Mi (RIPED of PetroChina),
Qian Zhang (PetroChina)
Room: R7 2023-08-07 15:22

Analogism is significantly applied in oil and gas resource assessment. By using quantified information encompassed in series of petroleum geological parameters, unknown hydrocarbon abundance of the area being assessed can be forecasted according to the known geological conditions and resource abundance of areas as calibration. For applying linear model is difficult to be universally effective due to actual complexity in oil and gas exploration, a method based on similarity calculation and inspired by isometric mapping theory which views the high dimensional space as a composition of twining rolls is proposed for oil and gas resource abundance analogism. The Process is generally divided into 5 stages: (1) Optimizing the composition of parameters and establishing the theoretical base point; (2) Calculating the similarity coefficient of each parameter among the data set and aggregating the coefficients into a matrix through distance calculation to represent the overall similarity; (3) Plotting the scatter diagram between resource abundance as Y axis and overall similarity to the base point as X axis, sequentially linking the points starting from the base point, and establishing a series of chains as the digital operation process for unfolding the information under two algorithmic principles: Positive correlation between the calculated similarity to base point and the corresponding resource abundance, The relatively highest quantified similarity between the two linked points which consist with ; (4) Determining the rational chain belonged to the area being assessed to be as the forecast model according to calculated results of the area's overall similarity to the areas as calibration; (5) Comparing the computed similarities from the area being assessed and areas as calibration which are applied in the model to base point respectively, and determining the location of area being assessed to acquire the forecasting outcome. A preliminary application based on actual resource assessment is accordingly established. High quantitative degree, low dependency on subjective experience and beneficial applicability to nonlinear analysis are tried to be presented through this proposed method with the expectation about the favorable potential in promoting effectivity and objectivity of resource assessment and related forecast in hydrocarbon exploration.

An improved method in hydrocarbon resource abundance forecast through analogism based on isometric mapping theory and its preliminary application

Jingdu Yu, Qiulin Guo, Shiyun Mi, Qian Zhang
Research Institute of Petroleum Exploration and Development, PetroChina.



S1313. Cold and hot subduction systems in the Andes: Insights from machine learning calculations of continental and oceanic heat flow in South America

Junjie Ji (Sun Yat-sen University),
Qiuming Cheng (China University of Geosciences/Sun Yat-sen University),
Yang Zhang (School of Earth Science and Engineering, Sun Yat-Sen University)
Room: R7 2023-08-07 15:24

Previous heat flow maps did not consider the effect of volcanoes in the South American Andes on surface heat flow, leading to their low results. We collected data from several databases and used a machine learning approach to calculate a new surface heat flow map for South America and its surrounding oceans, we compared the heat flow profiles and seismic location profiles between the continental arcs and the corresponding subduction slab, showing that different heat flows as initial conditions of the subducted slab have an indirect effect on the thermal activity with the continental arc. From the seismic and heat flow data, the subduction of the colder slab segment can extend to deeper depths due to less serpentinization. Compared to the hot plate segment it may be easier to form a double seismic zone at intermediate and deep depths. Two flat slab subductions in the South American Andes can effectively reduce continental surface heat flow, facilitating the formation of several large porphyry copper deposits.

S14 Geologic Forward Modeling

Daniel Tetzlaff (Westchase Software Corp.),
Scott Bowman (Petrodynamics),
Sergio Courtade (Schlumberger),
Gerard Massonnat (Total Energies)

Methods, algorithms, and practical applications of models that simulate physical geologic processes through time, including sedimentation, stratigraphy, compaction, tectonics, diagenesis, heat flow, and hydrocarbon maturation and migration. Also, hybrid techniques that use geostatistics, machine learning, and artificial intelligence in combination with forward modeling to improve data matching and enhance predictability away from data.

S1401. Stratigraphic forward modeling software package for research and education

Daniel Tetzlaff (Westchase Software Corp.)
Room: R5 2023-08-08 10:00

We have developed an open-source software package ("SedSimple") for stratigraphic forward modeling. It is aimed for use in research and education. The package uses ASCII files for input and output to facilitate understanding, debugging and preparation of multiple data sets though simple user-developed scripts. A free graphic program is available to assist in editing input and rendering results in 3D plus geologic time animation. The open-source concept offers the ability to understand the algorithm in detail, and modify or extend it as needed.

Several algorithms are available in the program to model various sedimentary environments and processes, including alluvial, fluvial, turbiditic, carbonatic, deep-sea clastic, wave action, vertical tectonics and sea-level changes. These processes can be run concurrently. Furthermore, the effect of some processes can be made to affect others (for example wave action can affect carbonate growth, while also causing erosion and redistribution of the deposits).

The algorithms are fully deterministic. As most forward models of natural environments, the results cannot be readily conditioned to data. However, the model is much faster than similar commercial packages of the same resolution. Therefore, we foresee that one of the main uses of the package will be to study workflows that require multiple realizations to obtain statistical predictions, such as Monte Carlo methods, combination with geostatistical methods, artificial intelligence and machine learning that require multiple inputs of geologically realistic models.

S1403. Forward modeling and uncertainty analysis of depositional parameters to support the construction of uncertainty scenarios for probabilisticI modeling in a pre-salt area of the Santos Basin

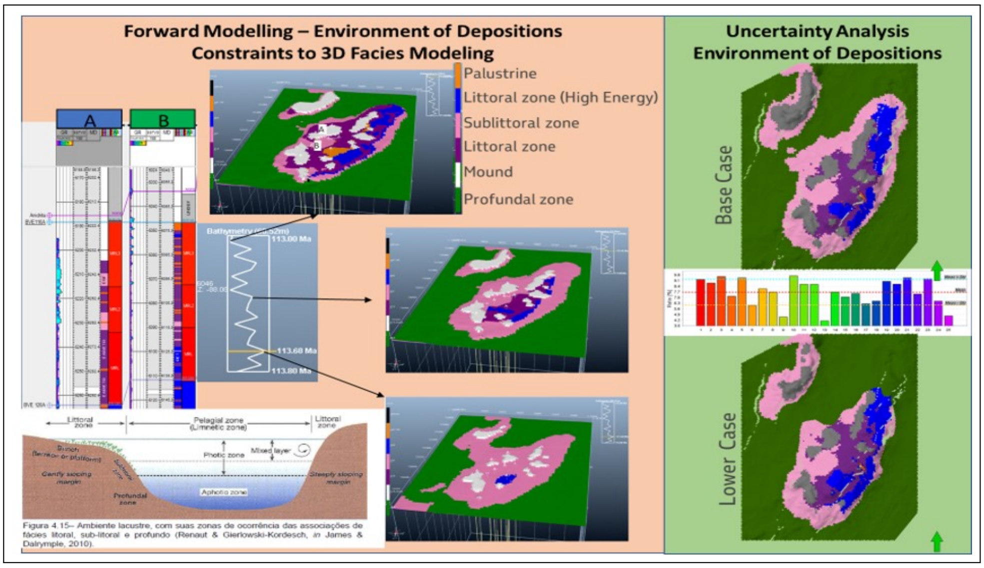
Renato Kramberger Carvalho (UFOP),
Olinto Gomes de Souza Junior (Petrobras),
Caroline Rangel Sales (Petrobras),
Mellissa Combas Baiense (Petrobras),
Desiree Liechoscki de Paula Faria (Petrobras)
Room: R5 2023-08-08 10:20

The area of interest discovered in the early 2010s reached carbonate reservoirs in the sag and rift section below the salt layer. The characterization and distribution of facies in Pre-salt structures is still a challenge, especially in the context of the study area, whose data acquisition is quite impaired due to severe losses during well drilling. To overcome this challenge, a process model was built to support probabilistic facies modeling and to represent its variations and depositional elements.

The workflow to forward modeling was developed through the construction of a model to represent each subzone of the field, controlling the facies distribution through stratigraphic, depositional and structural processes. However, since it is a model, there is a great uncertainty in the inputs used to control the processes that condition the deposition of carbonates.

A new methodology is proposed to incorporate an uncertainty analysis in the modeling workflow. An analysis of the uncertainties of the processes that control the deposition within the Dionisos software was carried out, changing some elements, such as subsidence and productivity of the facies that form the mounds. After carrying out this analysis, the P90 model was selected, considering the productivity of the mounds. Finally, this model was incorporated in the construction of the deterministic pessimistic scenario of the 3D geological modeling.

The work seeks, in this way, to emphasize the importance of analyzing the uncertainties inherent in the forward modelling that impact on other geological properties. A better estimation of reservoir uncertainties allows a more robust identification of risks, in addition to adequate mitigation through data acquisition. It leads to a better design decision, allowing increases in reserves and field recovery factor.

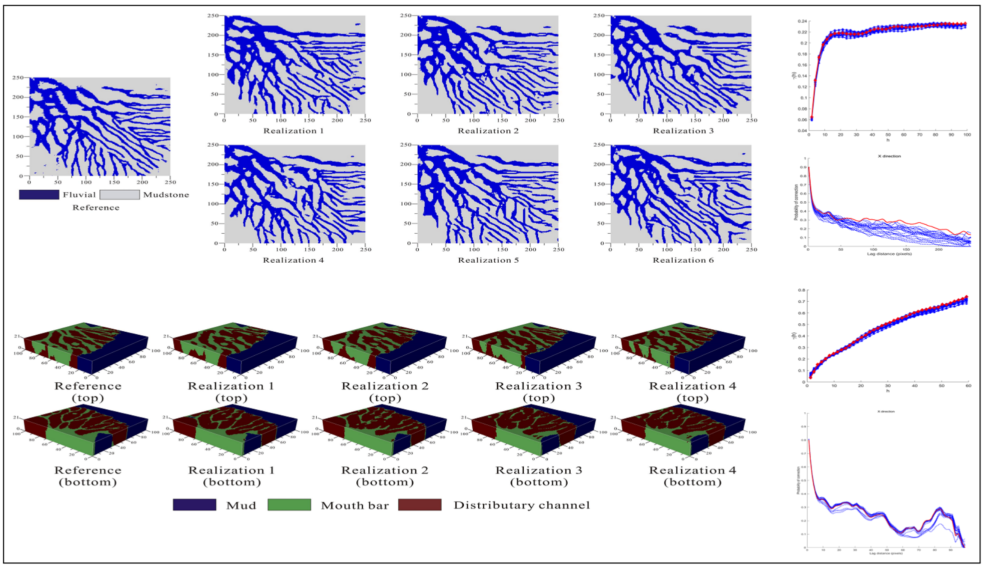


Forward Modelling – Environment of Depositions
Constraints to 3D Facies Modeling
Palustrine
Littoral zone (High Energy)
Sublittoral zone
Littoral zone
Mound
Profundal zone
A
B
113.00 Ma
113.60 Ma
113.80 Ma
Uncertainty Analysis
Environment of Depositions
Base Case
Lower Case

S1404. Stochastic Geological Reservoir Modeling based on Concurrent Multistage U-Net Generative Adversarial Network

Wenyao Fan (School of Computer Science, China University of Geoscience (Wuhan)),
Gang Liu (School of Computer Science, China University of Geoscience (Wuhan)),
Qiyu Chen (School of Computer Science, China University of Geoscience (Wuhan)),
Zhesi Cui (China University of Geoscience (Wuhan))
Room: R5 2023-08-08 10:40

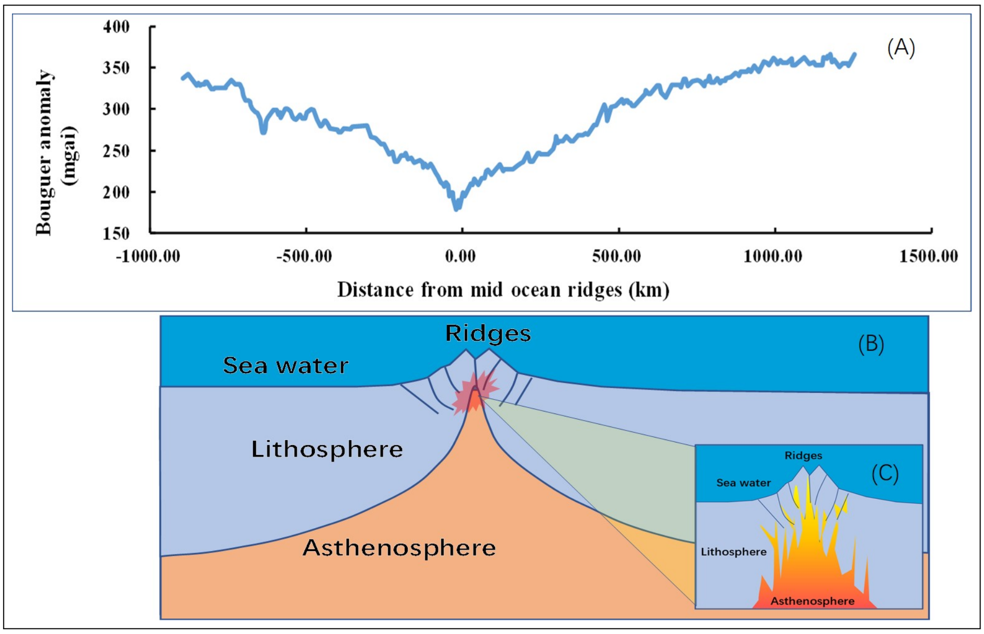
The traditional object-based geological modeling methods always have a low accuracy and uncertainties still exists. The numerical-simulation-based modeling methods are based on the stationary-assumption, although it can quantify the uncertainties, the complex spatial patterns with significant nonstationary properties are hard to reproduce. The Generative-Adversarial-Network-based (GAN-based) methods could overcome these issues, but the mode collapse could be easily occurred when the datasets are limited. Meanwhile, the fixed receptive field is unable to make the reconstruction consider the local features and global features of Training Images (TIs) simultaneously. Therefore, we proposed the Concurrent Multi-Stage U-Net GAN (Con-MSUGAN), which uses a single TI as the input of the fully convolutional pyramid structure and performs the multi-scale representations of spatial patterns, ensuring the feature information of TI could be captured under different scale correspondingly. Meanwhile, the concurrent training patterns were used to ensure the network parameters in different stages were highly related, improving the computing efficiency and realizing the multi-scale stochastic reconstruction. We used the stationary channel TIs and nonstationary delta TIs to verify the Con-MSUGAN reconstruction performances respectively. The results show that the simulation results are similar to the original TIs in terms of spatial variability, connectivity degree and facies type frequency distributions, the Multi-Dimensional Scaling visualization also verify the small differences between the simulation and TIs, indicating that the Con-MSUGAN could the overcome the problem that the complex spatial patterns are hard to reproduce. For the saved network parameters, inputting the data once only, the Con-MSUGAN could obtain different simulation results, thereby realizing the rapid stochastic reconstruction of the geological models.



S1405. Singularity analysis of morphological and geothermal processes over the mid-ocean ridges

Qiuming Cheng (China University of Geosciences/Sun Yat-sen University)
Room: R5 2023-08-08 11:00

Accurate predicting models are essential for mapping anomalous heat flow distribution and seafloor morphology over the mid-ocean ridges where two plates move away from each other, creating a gap or rift, magma rises from the mantle to fill the gap, eventually creating new crust. Theoretical functions derived from classical thermal conductivity models are often used to predict the rate at which heat flow and seafloor depth change with the change of age of the lithosphere or distance from the mid-ridge. These models show the general dependence of the change of heat flux and floor depth on the square root of age. For very young and old lithospheres, however, there are significant deviations between estimates and observations of heat flow and floor depth. Several recent studies have attempted to overcome these long-standing problems by modifying the dynamic model with additional layers of low-density crust or high-temperature mantle. The author adopted a different approach, namely using unconventional lithospheric density (fractal density) to quantify the irregularity and nonlinearity of parameters such as lithospheric porosity on mid-ocean ridges. The new model yields simple solutions in the form of a power-law functions that can significantly improve results for predicting mid-ocean ridge peak heat flow and floor depth. The new model also provides insights into the impact of spreading speed of plates on the distribution of the heat flow and floor morphology.



S1406. Numerical modeling of formation of a porphyry-copper ore shell: implications for the mechanism of metal enrichment

Fan Xiao (School of Earth Sciences and Engineering, Sun Yat-sen University, Zhuhai)

Room: R5 2023-08-08 11:20

Porphyry copper deposits (PCDs) are the most important sources of copper and molybdenum, providing about 75% of the world's copper resources, more than 50% of the metal molybdenum resources, formed by precipitation of metal sulphides from hydrothermal fluids released by magmatic intrusions that cooled at depth within the Earth's crust. The metals are typically concentrated in confined ore shells within vertically extensive vein networks. In this contribution, the finite-element method was used to simulate the fluid migration, heat transmission and mineralization of PCD formation process. Numerical modeling results show that hydrothermal flow is the main factor controlling the mineralization. Dynamic permeability responses to magmatic fluid expulsion can stabilize a metal precipitation above the magma chamber, forming a crescent-shaped metal enriched shell. The metallogenic duration as well as temperature gradient controls the most important economic characteristics of ore shell including shape, aggregation and ore grade.

S1407. Can source area erodibility delay slab break-off signal in a depositional record: insight from stratigraphic forward modelling

Paul Baville (Karlsruhe Institute of Technology),
Nevena Andrić Tomašević (Karlsruhe Institute of Technology (AGW))
Room: R5 2023-08-08 11:40

Geodynamic processes, like slab break-off are suggested to affect overlying sedimentary basin architecture. In this setting, rebound and associated surface uplift following slab-break-off are suggested to lead to a decrease in accommodation and an increase in sediment supply, following source rejuvenation due to uplift. This eventually results in basin shallowing. Although the sediment supply intensity is generally attributed to the tectonic and/or climatic effect, several recent studies emphasize the importance of the source area erodibility as an important factor that can delay tectonic and/or climatic signals.

To test this hypothesis, we modify stratigraphic forward modeling software GPM (Geological Process Modeling software provided and produced by SLB) to account for topography-dependent steady flow. The simulation includes clastic deposition in a deltaic environment. Model inputs include variations in the source area erodibility, sea level variations, tectonic rates, and variable water discharge which accounted for variable precipitation rates. Simulations were conducted in a stepwise increase in complexity to quantify the sensitivity of catchment scale depositional rates to aforementioned variations.

S1408. The control of porphyry magma cooling and crystallization process on mineralization: from the perspective of numerical simulation

Kaiqi Wang (Sun Yat-sen University),
Fan Xiao (School of Earth Sciences and Engineering, Sun Yat-sen University, Zhuhai)
Room: R5 2023-08-08 12:00

Porphyry copper deposits (PCDs) provide about 75% of the world's copper. Copper ore is associated with hydrothermal fluids rising from a porphyry magma chamber, which is mainly consist of metal-rich viscous melt, crystals and magmatic volatile phase (MVP). In the process of porphyry magma cooling and crystallization, the magma progressively evolved from melt dominated to crystal-rich mush, which leads to saturation of the MVP and enrichment of metals in the magmatic hydrothermal phase. During the evolution of porphyry magma chamber, the dynamic process of melt crystallization differentiation is constrained by various factors, including the initial temperature of magma chamber, the thermodynamic properties of melt, the heat exchange efficiency between magma chamber and surrounding rock, and the disturbance of melt supply from deep magma channel. Therefore, a 2D finite element geometric and mesh model of porphyry magma chamber at a certain depth is established to simulate the heterogeneous solidification and crystallization of the melt in the magma chamber. Firstly, based on previous research, the function of magma thermodynamics and fluid dynamics parameters changing with temperature is constructed, and certain initial conditions and boundary conditions are set in the numerical model. Then take the initial temperature of the magma chamber, the pressure disturbance of the deep magma channel, and the heat exchange coefficient between the magma chamber and the surrounding rock as variables, combine these variables and set up several numerical simulation models to explore the cooling and crystallization dynamic process of the porphyry magma and its possible response mechanism to mineralization.

S15 Spatial Association

Yongze Song (Curtin University),
Qiuming Cheng (China University of Geosciences/Sun Yat-sen University)

This session invites researchers to present the latest methods of spatial association and their implementations. Effective characterisation of the spatial association is the basis for spatial factors exploration, spatial prediction, and geographical decision-making. Currently, models of spatial dependence, heterogeneity, singularity, similarity, and other structures of spatial data are the mainstream to describe spatial association. Recent studies demonstrate that geospatial artificial intelligence also provides accurate and effective solutions for spatial association analysis. This session welcomes submissions from both methodology and application studies about the spatial association.

S1501. A comparison of spatial prediction methods

Yongze Song (Curtin University),
Zehua Zhang (School of Design and the Built Environment, Curtin University),
Peng Luo (Chair of Cartography, Department of Aerospace and Geodesy, Technical University),
Qiuming Cheng (China University of Geosciences/Sun Yat-sen University)
Room: R9 2023-08-07 14:40

Spatial prediction, which involves predicting the distribution of Earth and environmental attributes using sample data, is a crucial task in geospatial and Earth data analysis. Current spatial prediction methods can be grouped into three categories: aspatial methods, spatial dependence-based methods, and geostatistical methods. Aspatial methods typically comprise linear regression, nonlinear regression, and machine learning algorithms, while spatial dependence-based methods encompass geographically weighted regression and its extensions, as well as Bayesian hierarchical models. Geostatistical methods, also known as the kriging family of models, are designed to account for the spatial local heterogeneity of data. Given the complexity of geospatial and Earth data, advanced methods are increasingly necessary to effectively capture various data characteristics for more accurate, reliable, and robust spatial prediction. In recent years, a number of advanced geospatial models have been developed to describe diverse data characteristics for spatial prediction, such as the second-dimension spatial association (SDA), generalized heterogeneity model (GHM), and geographically optimal similarity (GOS). In this study, we will evaluate these advanced geospatial models alongside previous methods, focusing on the accuracy, effectiveness, and reliability of spatial prediction. Our findings will help to identify the scope, advantages, and limitations of these methods, enabling broader implementation of these techniques in future spatial prediction efforts.

S1502. a multiscale 3D geological model visualization method based on mesh simplification algorithm

Chen Genshen (China University of Geosciences (Wuhan)),
Gang Liu (Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences),
Qiyu Chen (China University of Geosciences (Wuhan))
Room: R9 2023-08-07 15:00

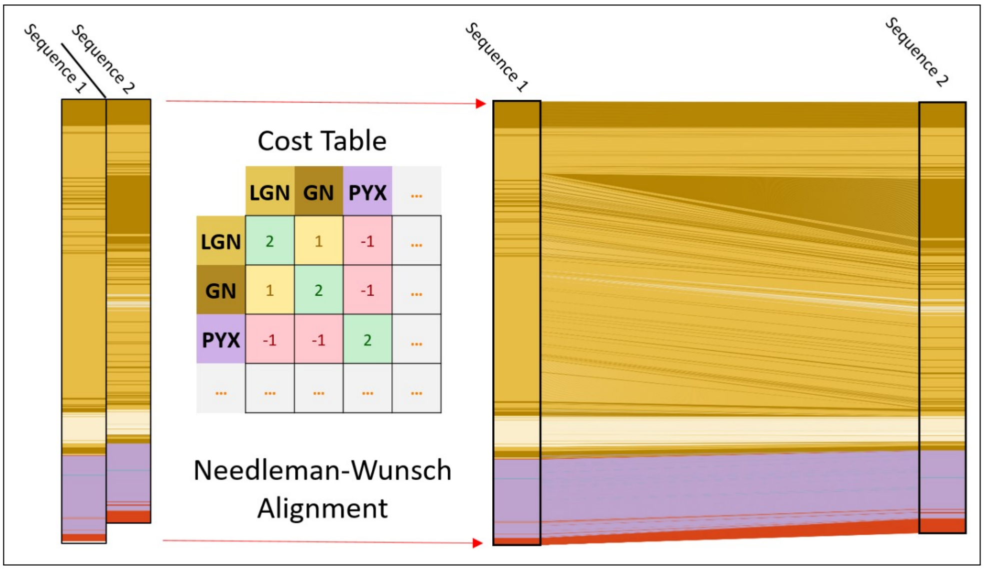
In this paper, a multiscale 3D geological model visualization method based on Mesh simplification is proposed. The main challenge for large-scale 3D geological model visualization is that the memory and computational ability of browsers and computers is not enough for friendly interactive analysis. Firstly, the suitable resolution is pre-defined and the mesh simplification is performed on the models. The key detail information, such as fault, geological boundaries and strata pinching-out, are retained and hierarchical construction of the original model is performed so that it could reduce the number of model's nodes and complexity of the visualization rendering, thereby obtaining the models with different resolutions. Then, in different viewpoint and scene, according to the mapping relationship between viewing distance and viewing area, spatial coordinates and attribute information, the detailed performance with corresponding resolutions is selected, and auto switch of different hierarchy model is realized by defining switching law so that the block details of the model in the current viewing area could be displayed with different resolution. Finally, based on the index of 3D geological models based on octree, the smooth transition of the model with different scale is realized by preloading and caching technologies, thereby improving the efficiency of model switching and data real-time loading. Our method is verified on a large-scale geological model dataset of the Yangtze River Basin. The results show that our method could improve the visualization rendering efficiency of the 3D geological models, accelerating the visual management and open sharing correspondingly.

S1503. Sequence Alignment for Automated Drill Hole Correlation

Tom Buckle (Camborne School of Mines, University of Exeter),
Hannah Hughes (Camborne School of Mines, University of Exeter)
Room: R9 2023-08-07 15:20

Correlation of downhole stratigraphy is routinely performed by geologists in resource exploration, however it can be subjective and time consuming when having to be performed manually on large numbers of drill holes. Automation of this process for continuously sampled data such as wireline logs has been achieved using dynamic time warping (DTW) which aims to find the optimal alignment of two continuous signals by stretching and compressing the signals. DTW was originally used to compare speech patterns in speech recognition tasks, though has since been used in various areas of data science and within the Geosciences to automate the correlation of wireline logs. One constraint of DTW is that it does not account for insertions in the time series sequences it is aligning. The geological equivalent of this would be rapid facies changes or localised features such as xenoliths or intrusions. Furthermore, it requires continuously sampled data that are able to be correlated. Whilst stratigraphy showing a clear wireline response, such as gamma ray in clastic sequences, are well suited to correlation using DTW, other systems such as layered intrusions are not so clear-cut. This is due to subtle changes in lithology or mineralogy, texture and lithology. Likewise, downhole geochemistry does not describe texture and can be noisy, irregularly sampled or not present for an entire sequence.

An alternative to DTW is global sequence alignment via the Needlement Wunsch (NW) algorithm. The NW algorithm was developed for use in bioinformatics for alignment of DNA and protein sequences and is fundamentally similar to DTW, but works on categorical data and can account for the insertion and deletion of categories within a sequence. In this work, we use NW to correlate lithology in a layered intrusion sequence from the Northern Limb of the Bushveld Complex. We demonstrate how use of a custom cost matrix allows for geological domain expertise to be included in the correlation. This initial lithological correlation can be used to quickly construct simple cross sections, or as a constraint for further detailed correlation using DTW.



S17 Uncertainty Modeling

Florian Wellmann (RWTH Aachen University),
Clare Bond

Uncertainties are prevalent in many geoscientific studies: from geostatistics and 3-D geological modelling to coupled process simulations. In this session, we invite contributions covering aspects of uncertainty estimation, modelling and visualisation in geosciences. A special focus will be on (1) methods to estimate uncertainties in data sets, expert elicitation and the analysis of bias (2) approaches to propagate uncertainties through modelling and simulation, including surrogate modelling, probabilistic programming and machine learning approaches, and (3) the quantification and visualisation of uncertainties in temporal and spatial domains. We encourage contributions both addressing theoretical and methodological challenges, as well as applications to geoscientific problems.

S1701. Uncertainty Analyses in Hydrogeology: Review and Applications to NAPL contamination

Rachid Ababou (IMFT: Institut de Mecanique des Fluides de Toulouse),
Juliette Chastanet (GINGER-BURGEAP group, France),
Jean-Marie Côme (GINGER-BURGEAP Group, R&D Department, France),
Manuel Marcoux (IMFT: Institut de Mecanique des Fluides de Toulouse, France),
Michel Quintard (IMFT: Institut de Mecanique des Fluides de Toulouse, France)
Room: R5 2023-08-09 10:00

The different alternatives for managing a contaminated site are usually evaluated in a context of uncertainty, on the basis of cost/benefit balance requiring a predictive modeling step to assess mass depletion and duration of the contamination source (e.g. DNAPL), the dissolved contaminant plume migration, and its impact in the subsurface.

For instance, the ESPER project (Evaluation and Sensitivity of models for Predicting the depletion and Remediation of organic contamination sources in subsurface) was aimed at developing a methodology and a software tool for incorporating uncertainty and sensitivity analyses into such contamination models.

In this communication, we review various approaches to propagate uncertainty through input/output models with uncertain parameters. We present different approaches and results from simple to complex models, and a site-specific study from the ESPER project in particular.

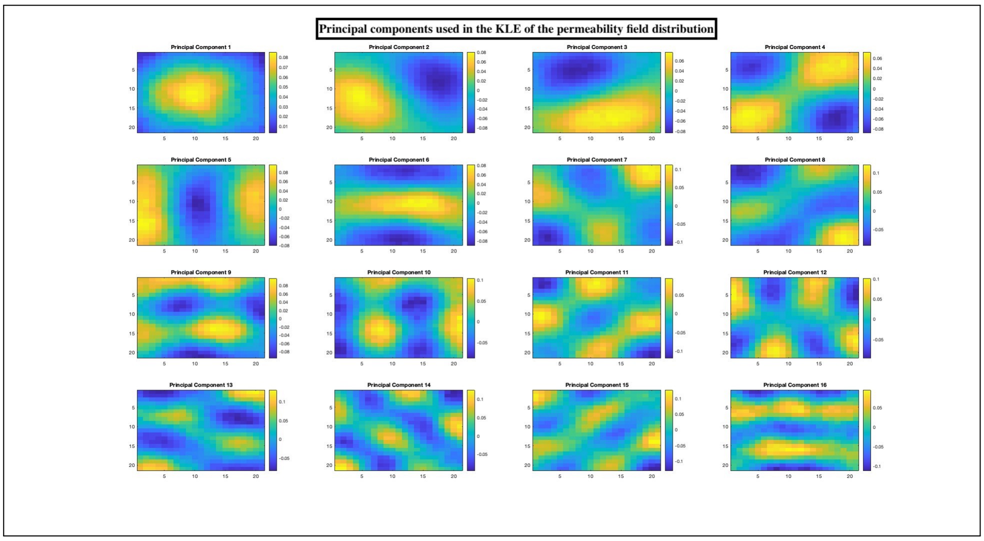
After a first introductive part, the second part of the study covers different methods and concepts for carrying out uncertainty analyses. The third part develops a comprehensive setting for probabilistic uncertainty quantification with random input parameters (including multivariate sets). The fourth part focuses on fuzzy variables, and on combined random/fuzzy approaches (possibility theory).

The rest of the work is devoted to illustrations of several techniques of uncertainty analyses with various simplified models (fifth part), and to full-fledged applications to 3D subsurface contamination with uncertain parameters (sixth part). The latter involves probabilistic Monte-Carlo simulations with or without a meta-model, applied to a new 3D semi-analytical model of solute advection-dispersion due to a decaying DNAPL source, and finally, a complex computer model (MODFLOW-SURFACT 3D code) tested on a real DNAPL contaminated site with uncertain parameters.

S1702. Spectral Methods for Stochastic Surrogate Modeling of Subsurface Flow Dynamics

Tarek Diaa-Eldeen (NTNU),
Morten Hovd (Norwegian University of Science and Technology (NTNU)),
Carl Fredrik Berg (NTNU)
Room: R5 2023-08-09 10:20

Subsurface flow is characterized by high uncertainties that limit the predictability of its dynamical models. This in return hinders implementing an efficient production optimization strategy, which typically uses the model predictions to optimize future production decisions. Therefore, quantifying the model's uncertainty and characterizing its effect on the model predictions is imperative in the decision-making process in petroleum reservoir systems. However, this often involves performing a large number of reservoir model evaluations that are required to propagate an ensemble of realizations that represent the input uncertainty. To be combined with online optimization routines thereafter in the so-called closed-loop reservoir management (CLRM), excessive computational effort is required due to the complexity of the nonlinear and high-dimensional reservoir simulation models. In this study, two spectral methods are combined to obtain a computationally-efficient surrogate model for the stochastic dynamics of reservoir systems in the CLRM. Firstly, the Karhunen–Loève expansion (KLE) is applied to reparameterize the subsurface flow models by representing the uncertain model parameters in terms of low-dimensional independent random variables using the spectral decomposition of the covariance function. Then, a spectral metamodeling approach, namely, the polynomial chaos expansion (PCE) is applied to model the distribution of the system's response representing the stochastic dynamic simulation. A PCE for the reservoir dynamics is computed and employed to propagate the uncertainty without the need for additional expensive model evaluations. This can reduce the computational burden in both the forward and inverse problems of the CLRM. Results show that the PCE surrogate model can accurately quantify the uncertainty and evaluate a large number of model realizations at the cost of evaluating a polynomial, compared to the full model evaluations using Monte Carlo simulations.



S1703. Quantifying CO2 Leakage Uncertainty for Planning Safe Geological Carbon Storages

Sofia Mantilla Salas (Stanford University),
David Zhen (Stanford University),
Markus Zechner (Stanford University),
Yizheng Wang (Stanford University),
Jonas Kloeckner (Stanford University),
Jef Caers (Stanford University)
Room: R5 2023-08-09 10:40

One of society's main concerns about Geological Carbon Storage (GCS) is the leakage of CO2 that could eventually contaminate freshwater resources or reach the surface. Many stakeholders believe that a prediction of leakage probability below one percent is adequate to ensure the viability of a CO2 storage project. However, the problem is that to be certain in predicting a rare leakage event, an enormous amount of possible models and flow simulations are needed, which is translated into extensive and inefficient computing time.

This study aimed to analyze how the geological structures and the petrophysical properties, particularly the top reservoir surface geometry, and porosity, impact the flow and eventual leakage of CO2. We used the MRST − CO2lab software from SINTEF to perform Monte Carlo flow simulations on synthetic geological models with uncertain top surfaces and porosity derived from a real case. First, we used the data-knowledge-driven trend surface analysis method to construct stochastic geological top surfaces, accounting for the seismic signal and its interpretation uncertainty. Then, the stochastic porosity models were created using geostatistical approaches, accounting for the spatial uncertainty. To estimate the leakage probability with confidence, we used importance sampling to produce a generating function to build more models with top surfaces and porosities with higher leakage probability. Then we computed a correction by weighing the samples.

This methodology drastically reduces the number of models, flow simulations, and computing power to get to the desired confidence interval of the leakage probability. Furthermore, it shows a way to reduce the uncertainty of the different risk factors associated with GCS, thus allowing the planning of storage locations needed for viable GCS projects. Furthermore, this equips society and policymakers with guarantees to rapidly deploy this relevant carbon-neutral technology on a large scale.

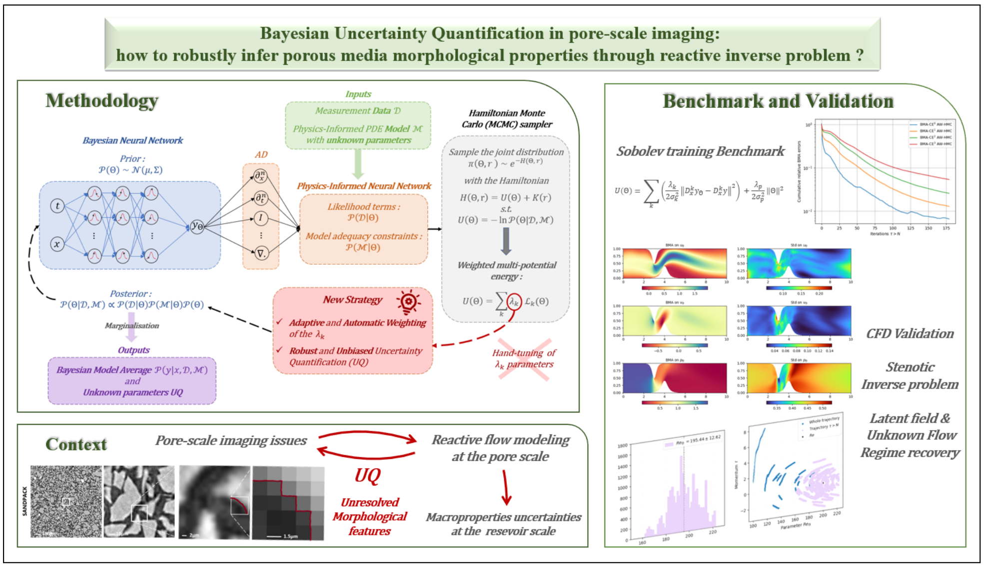
S1704. Bayesian uncertainty quantification in pore-scale imaging: how to robustly infer porous media morphological properties through reactive inverse problems?

Sarah Perez (University of Pau and Adour Area),
Philippe Poncet (Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, LMAP UMR CNRS-UPPA 514)
Room: R5 2023-08-09 11:00

Pore-scale modelling and simulation of reactive flows are usually built upon X-ray microtomography experiments that provide the internal and complex geometries of the porous media. Although these models offer the opportunity to study the evolution of the macroscale properties in the context of CO2 storage, they suffer from measurement uncertainties arising from the pore-scale imaging process. Image artefacts, including noise and unresolved morphological features, are such limitations and require data-driven methods with uncertainty quantification. This aims to improve the reliability of the numerical modelling through better identification of the morphological properties using reactive inverse problems.

However, developing an efficient data-assimilation framework that robustly addresses such multi-objective and multi-scale inverse problems remains challenging. Therefore, we propose a novel adaptive strategy for unbiased uncertainty quantification in Bayesian Physics-Informed Neural Networks, which relies on an automatic weighting of the target distribution parameters and objectives. It benefits from enhanced convergence and stability compared to conventional formulations and reduces sampling bias by avoiding manual tuning of the critical weights. The adjusted weights also bring information on the task uncertainties and improve the reliability of the noise-related and model adequacy estimates.

In this presentation, we focus on the improvements and efficiency of this new approach on various benchmarks, including comparisons with an analytical baseline, and Computational Fluid Dynamics applications. We also apply this auto-weighting in Bayesian Physics-Informed Neural Networks on pore-scale imaging uncertainty quantification. In this direction, we aim to quantify unresolved features in porous media X-ray microtomography and capture morphological uncertainties on the porosity field.

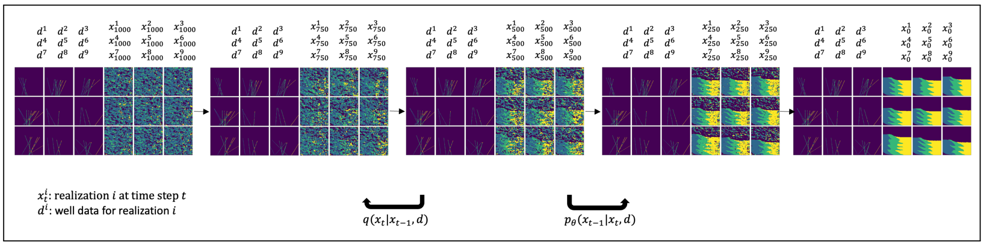

Bayesian Uncertainty Quantification in pore-scale imaging:
how to robustly infer porous media morphological properties through reactive inverse problem ?
Methodology
Benchmark and Validation
Sobolev training Benchmark
CFD Validation
Stenotic Inverse problem
Latent field & Unknown Flow Regime recovery
Context
Pore-scale imaging issues
Reactive flow modeling at the pore scale
UQ
Unresolved Morphological features
Macroproperties uncertainties at the resevoir scale
Bayesian Neural Network
Inputs
Measurement Data
Physics-Informed PDE Model with unknown parameters
Hamiltonian Monte Carlo (MCMC) sampler
Physics-Informed Neural Network
Likelihood terms
Model adequacy constraints
New Strategy
Adaptive and Automatic Weighting
Robust and Unbiased Uncertainty Quantification (UQ)
Weighted multi-potential energy
Hand-tuning of parameters
Marginalisation
Outputs
Bayesian Model Average
Unknown parameters UQ

S1705. Conditional Facies Sampling using Denoising Diffusion Probabilistic Model

Oscar Ovanger (NTNU),
Daesoo Lee (Norwegian University of Science and Technology)
Room: R5 2023-08-09 11:20

We present a novel method for unconditional and conditional facies sampling using Denoising Diffusion Probabilistic Model (DDPM). We propose GeoLatentDiffusion which is a variant of DDPM tailored for facies sampling with improved computational efficiency. This approach is effectively utillizing the latent space of the model formulation. We use training images from an explicit mathematical representation of geobodies based on expert knowledge in the field. In our case this relies on a prograding and aggredating shoreface model with stacked bedsets of sand and shale bodies.

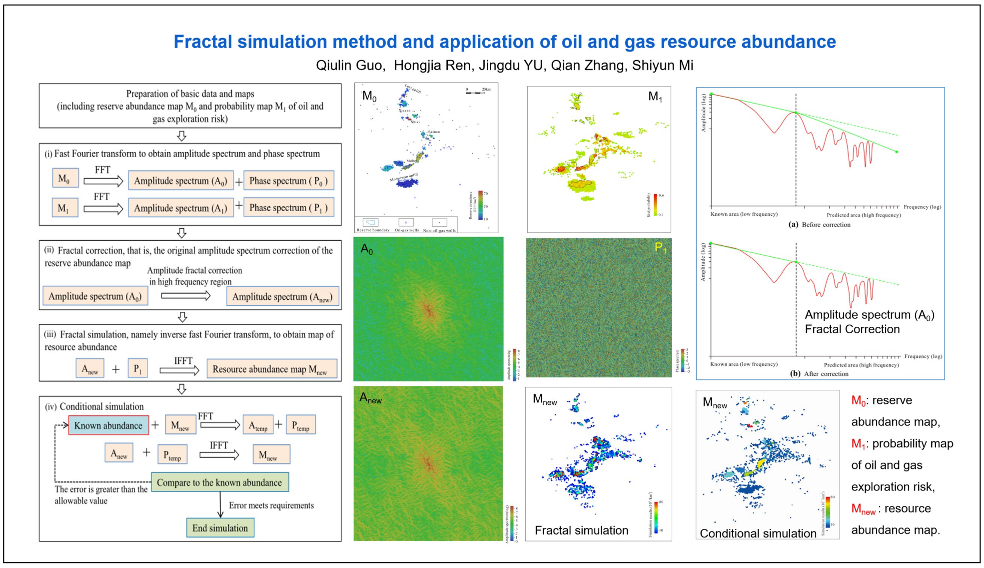
The conventional methods for (unconditional) conditional facies sampling typically require an explicit formulation for the (prior) posterior distribution. Several recent studies have proposed Generative Adversarial Network (GAN)-based methods to eliminate a need for the explicit formulation. However, GANs carry inherent limitations such as non-convergence, mode collapse, generator-discriminator unbalance, and sensitive hyperparameter selection. DDPM is the state-of-the-art generative method in the deep learning literature and has shown superior generative performances. Our experimental results show that GeoLatentDiffusion outperforms both conventional multiple-point statistics methods and competitive GAN-based methods, both in terms of realism and preservation in generated samples.



S1706. Fractal simulation method and application of oil and gas resource abundance

Qiulin Guo (PetroChina),
Jingdu Yu (Petrochina),
Hongjia Ren,
Qian Zhang (PetroChina),
Shiyun Mi (RIPED of PetroChina)
Room: R5 2023-08-09 11:40

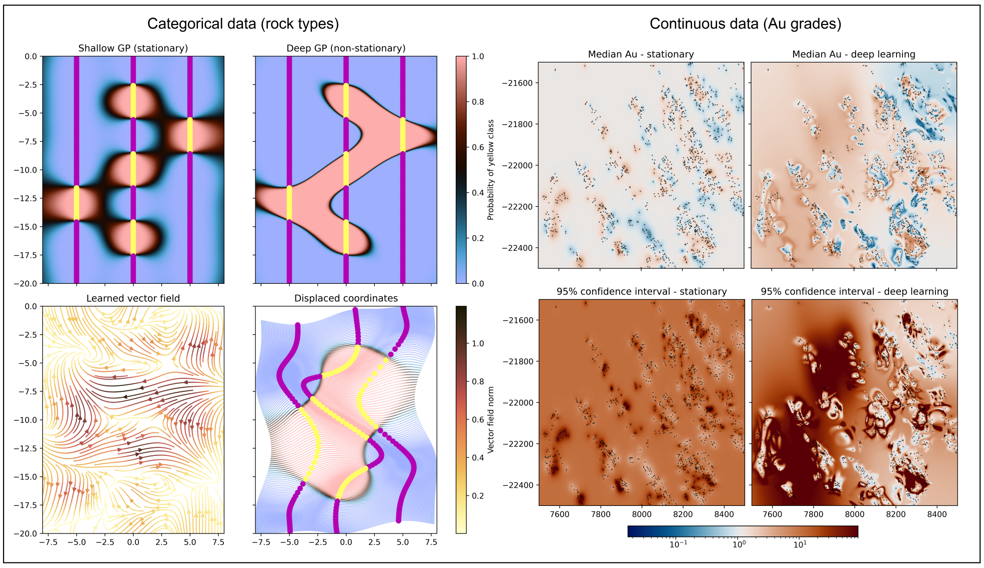
Previous studies believe that the distribution of oil and gas resource abundance has obvious fractal characteristics. This paper proposes a fractal simulation method to predict the abundance of oil and gas resources based on the exploration risk map(M0) and the reserve abundance distribution map (M1). The 4-step key technical process is established: (1) Fourier rapid transformation of M0 and M1 to obtain the corresponding amplitude spectrum and phase spectrum; (2) correct the original amplitude spectrum of M0 to form a new amplitude spectrum; (3) Fourier rapid inverse transformation to form the resource abundance map; (4) condition simulation to form the final resource abundance map. At the same time, the resource abundance prediction model and condition simulation objective function are improved, the problem of risk visualization and spatial distribution prediction of oil and gas resources is solved, and the spatial distribution prediction method of oil and gas resources is formed. The prediction method have been applied to the Jurassic Sangonghe Formation in the hinterland of the Junggar Basin, and good results have been obtained. The results indicated that within the known area, taking the known abundance as the constraint condition, the coincidence rate of the simulated quantities of the original model and the improved model with the actual reserves reached 99.98% after the conditional simulation, indicating that the conditional simulation was effective. In addition, with the improved model, the predicted remaining resources are 0.7899×108 t, which is 65% of the discovered reserves, and the original model predicts that the remaining resources are 3.5033×108t, which is 2.89 times greater than the discovered reserves. Compared with the area in the middle stage of exploration, the results of the improved model are more consistent,and the results of the original model are obviously larger, indicating that the improved model has a good predictive effect for the unknown area. Finally, according to the risk probability and remaining resource distribution, the favorable areas for exploration were optimized. The application results show that these prediction technologies are effective and can provide important reference and decision-making for resource evaluation and target optimization.


Fractal simulation method and application of oil and gas resource abundance
Qiulin Guo, Hongjia Ren, Jingdu YU, Qian Zhang, Shiyun Mi
Preparation of basic data and maps (including reserve abundance map M0 and probability map M1 of oil and gas exploration risk)
(i) Fast Fourier transform to obtain amplitude spectrum and phase spectrum
M0 FFT Amplitude spectrum (A0) + Phase spectrum (P0)
M1 FFT Amplitude spectrum (A1) + Phase spectrum (P1)
(ii) Fractal correction, that is, the original amplitude spectrum correction of the reserve abundance map
Amplitude fractal correction in high frequency region
Amplitude spectrum (A0) Amplitude spectrum (Anew)
(iii) Fractal simulation, namely inverse fast Fourier transform, to obtain map of resource abundance
Anew + P1 IFFT Resource abundance map Mnew
(iv) Conditional simulation
Known abundance + Mnew FFT Atemp + Ptemp
Anew + Ptemp IFFT Mnew
Compare to the known abundance
The error is greater than the allowable value
Error meets requirements
End simulation
M0
M1
A0
P1
Anew
Mnew
Fractal simulation
Mnew
Conditional simulation
(a) Before correction
(b) After correction
Amplitude spectrum (A0) Fractal Correction
Known area (low frequency)
Predicted area (high frequency)
Frequency (log)
Amplitude (log)
M0: reserve abundance map,
M1: probability map of oil and gas exploration risk,
Mnew: resource abundance map.

S1707. Uncertainty propagation in deep Gaussian process networks

Ítalo Gonçalves (Federal University of Pampa)
Room: R5 2023-08-09 14:00

In recent years the well-known kriging model has been adopted by the machine learning community, where it is known by the name of Gaussian process (GP). Variational inference (VI) enables the GP to move beyond simple regression tasks (modeling continuous Gaussian-distributed variables) to non-Gaussian variables, categorical data (implicit modeling), and compositional variables. The variational GP (VGP) consists in specifying a number of inducing points (which may or not coincide with the data points) with corresponding latent variables, function values that are distributed as a multivariate Gaussian. The latent variables' mean and variance can be interpolated from the inducing points to the data points or to a regular grid to generate maps and other products. The VGP is then trained by maximizing the data likelihood with respect to its parameters (range, anisotropy, and others) given a likelihood distribution compatible with the data type, where the latent variables' distribution shifts to provide the best possible fit to the data. The deep Gaussian process model (DGP) builds a deep hierarchy by feeding a latent variable as input to a new VGP layer. As the latent variables are Gaussian, propagation of data through the network would require sampling at each layer, a costly operation. This work employs a non-stationary covariance function that can integrate the uncertainties at each layer analytically. This new network can be customized to allow the machine to learn complex relationships in the data, such as lithological controls on mineral grades, non-sinusoid cyclicities in time series, to compensate for folds and faults, and to use geological proxies to guide the spatial interpolation, all with calibrated uncertainty estimates. The project's website (github.com/italo-goncalves/geoML) contains examples and code.



S1708. Implicit Modelling with GemPy and Spatial Uncertainty Reduction of Potential Nuclear Waste Storage Sites

Carlos Colombo (RWTH - E.On Energy Research Centre CG3),
Florian Wellmann (RWTH Aachen University)
Room: R5 2023-08-09 14:20

Geological implicit modelling methods represent geological surfaces as iso-values of a three-dimensional scalar field (Wellmann and Caumon, 2018). Among many purposes, the methods can be used for the depiction of geological structures for the storage of nuclear waste. In Germany, potential structures for this purpose might be given by crystalline, clay or salt rocks (BGE, 2023). We use GemPy, a special implicit surface representation method (de la Varga et al., 2019) to build a subsurface model from interface contact points with orientations that are obtained from seismic and borehole data. After building the model, we identify and focus our analysis in the geometry of layered formations of crystalline or clay rocks, or salt domes. The spatial uncertainty of the structures is quantified after performing object-distance simulation method (ODSIM) on the model, a method that combines object-based and pixel based modelling for categorical objects (Henrion et al., 2010). Finally, following Wellmann and de la Varga (2016) we combine uncertain prior information with geologically motivated likelihood functions in a Bayesian inference framework to reduce uncertainties in the ensemble of generated models. Our results show a reduction of the spatial uncertainty in the spatial definition of the geometrical structures and provide quantitative information, in the form of probability, for the spatial uncertainty. We believe our method could be a useful contribution for decision-making as the uncertainty is quantified and can be incorporated in the metrics used to validate the selection of a site storage, like thickness or any other spatial requirement.

References:

-BGE (Bundesgesellschaft für Endlagerung); https://www.bge.de/en/repository-search/, accessed: Mar-15-2023, 12:37.

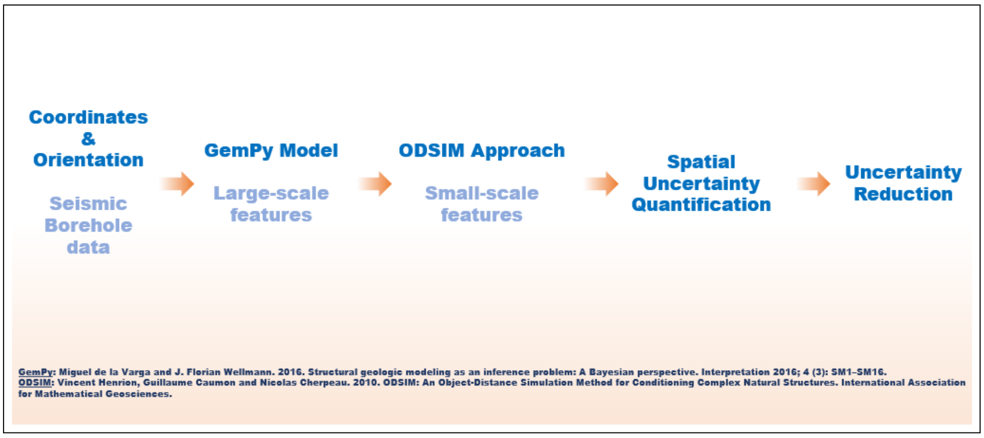
-Wellmann and Caumon. 2018. 3-D Structural geological models: Concepts, methods, and uncertainties. Cedric Schmelzbach. Advances in Geophysics, Elsevier.

-Wellmann and Regenauer-Lieb. 2011. Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics.

-de la Varga, Schaaf and Wellmann. 2019. GemPy 1.0: open-source stochastic geological modeling and inversion. Geoscientific Model Development..

-de la Varga and Wellmann. 2016. Structural geologic modeling as an inference problem: A Bayesian perspective. Interpretation 2016.

-Henrion, Caumon and Cherpeau. 2010. ODSIM: An Object-Distance Simulation Method for Conditioning Complex Natural Structures. IAMGnternational Association for Mathematical Geosciences.


Coordinates
&
Orientation
Seismic
Borehole
data
GemPy Model
Large-scale
features
ODSIM Approach
Small-scale
features
Spatial
Uncertainty
Quantification
Uncertainty
Reduction
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.

S1709. Basin modeling informed Bayesian Networks to update pore pressure prediction with data while drilling

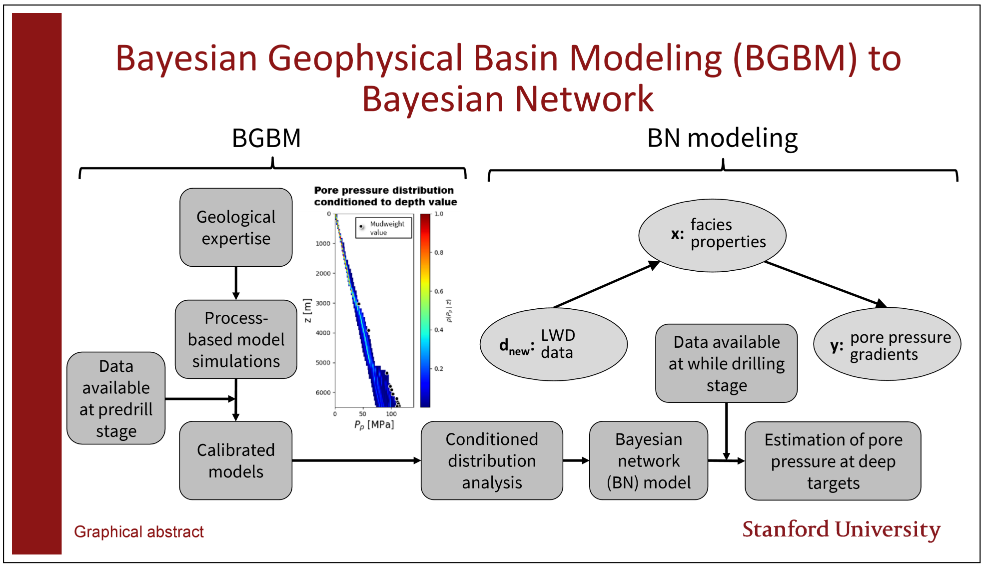
Josue Sa da Fonseca (Stanford University),
Tapan Mukerji (Stanford University)
Room: R5 2023-08-09 14:40

Pore pressure prediction is crucial when drilling wells to reach reservoirs in the subsurface. An accurate pore pressure forecast can help reduce drilling costs and ensure the safety of the drilling operation. In this study, we present a novel approach using Bayesian Network (BN) that is informed by basin modeling to update pore pressure predictions ”on-the-fly” with new data acquired during the drilling procedure.

The proposed method involves dividing the drilling section into multiple geological compartments (e.g., stratigraphically) and using Bayesian Networks to model the relationships between different geological parameters that influence pore pressure. The conditional distributions of the nodes of the BN model are learned based on the calibrated posterior basin models from a Bayesian Geophysical Basin modeling (BGBM) workflow. This BGBM workflow yields Monte Carlo simulated models of the sub-surface that integrate physical and geological processes during their simulations and are conditioned with predrill offset well and seismic data.

Once the BN is calibrated with the outputs of the basin models, it can be used to make inferences about pore pressure based on new evidence from drilling without running any more computationally demanding basin model simulations. Next, we build a BN with the simulated pore pressure gradients of each geological compartment and train the relationships of the geological parameters inferred from drilling data. Finally, a case study is performed with real data from the Gulf of Mexico.

Overall, the proposed approach demonstrates the potential of using BGBM-informed Bayesian Network modeling of geological compartments to predict pore pressure while drilling. This scheme has the advantage of incorporating calibrated models with geological processes by using BGBM results, which can enhance the accuracy and reliability of the predictions. The method also has the potential to be extended to other drilling-related parameters. It can be integrated into drilling automation systems to provide real-time feedback to the drilling team.



Bayesian Geophysical Basin Modeling (BGBM) to Bayesian Network
BGBM
BN modeling
Geological expertise
Process-based model simulations
Data available at predrill stage
Calibrated models
Pore pressure distribution conditioned to depth value
Mudweight value
z [m]
P_p [MPa]
Conditioned distribution analysis
Bayesian network (BN) model
x: facies properties
d_{new}: LWD data
Data available at while drilling stage
y: pore pressure gradients
Estimation of pore pressure at deep targets
Graphical abstract
Stanford University

S1710. Rigid Transformations in Lower-Dimensional Space for Subsurface Uncertainty Quantification and Interpretation

Ademide. O. Mabadeje (The University of Texas at Austin), Michael J. Pyrcz (University of Texas at Austin)
Room: R5 2023-08-09 15:00

Subsurface datasets commonly have big data criteria, i.e., large volume, significant feature variety, high sampling velocity, and limited veracity. Data volume is enhanced by multiple features derived by imposition of various engineering and geological inputs that invoke the curse of dimensionality. Dimensionality reduction (DR) methods are either linear or nonlinear; however, for subsurface datasets, nonlinear dimensionality reduction (NDR) methods are most applicable due to data complexity. Metric-multidimensional scaling (MDS) is a suitable NDR method that retains the data intrinsic structure and could quantify uncertainty space. Like other NDR methods, MDS is limited by its inability to achieve a stabilized unique solution of the low-dimensional space (LDS) invariant to Euclidean transformations, and has no extension for inclusions of out-of-sample points (OOSP). Hence, it is imperative to transform these datasets into meaningful stable representations of reduced-dimensionality that permit OOSP without model recalculation.

We propose the use of rigid transformations to obtain unique solutions of stabilized Euclidean invariant representation for LDS. First, compute a dissimilarity matrix as the MDS input using a distance metric to obtain the LDS for N-samples and repeat for multiple realizations. Then, select a base case and perform rigid transformation on further realizations to obtain rotation and translation matrices that enforce Euclidean transformation invariance under ensemble expectation. The stabilized-solution identifies anchor positions using convex-hull algorithm in comparison to N+1 case from prior matrices to obtain a stabilized representation of the OOSP. Next, the loss function and normalized-stress are computed via distances between samples in the high-dimensional space and LDS to quantify and visualize distortion. The proposed workflow is tested using different sample sizes for a 2D-registration problem with different distance metrics.

The results show our method is effective for NDR methods to obtain unique, stable representations of LDS invariant to Euclidean transformations. Also, we propose a distortion-based metric that characterizes the uncertainty space for samples in subsurface datasets, which is useful for model updating and inferential analysis for OOSP. Therefore, we recommend the workflow as an invariant transformation mitigation unit in LDS for unique solutions to ensure repeatability in subsurface NDR methods.

S1711. Development of the Transition Risk Analysis Model (TRAM) for Uncertainty Quantification of the Financial Impact of a Carbon-Intensive Company

Soeun Yoon (Department of Climate and Energy Systems Engineering, Ewha Womans University),
Seoyoung Jung (Department of Climate and Energy Systems Engineering, Ewha Womans University),
Minchul Jang (Center for Climate/Environment Change Prediction Research, Ewha Womans Universit),
Jaehak Hwang (Department of Climate and Energy Systems Engineering, Ewha Womans University/Fin),
Baehyun Min (Ewha Womans University)
Room: R5 2023-08-09 15:20

This study presents the Transition Risk Analysis Model (TRAM) as a tool to evaluate the financial impact of carbon-intensive companies exposed to climate change transition risks. The TRAM model estimates a company's annual total cost of carbon emissions based on government policies, financial market trends, and the company's carbon roadmap. This cost comprises the company's investment in emissions reduction, auction costs, and emission liabilities under the Korea Emissions Trading System (K-ETS) from 2020 to 2050. Key variables used for cost estimation are selected from historical data of the Korea Emissions Trading System (K-ETS) and corporate investment roadmaps. The TRAM model utilizes a combination of deterministic and probabilistic approaches for the analysis. Internal data of the industrial cooperator, including estimated annual energy consumption, carbon intensity, carbon emissions, and related costs, are inputs of the TRAM model under a confidentiality agreement. The model builds deterministic emission allocation scenarios based on global emissions scenarios and enhanced domestic emissions scenarios of Nationally Determined Contribution (NDC) based on the 2050 Carbon Neutral Strategy of the Republic of Korea. Probabilistic financial scenarios such as carbon price evolutions are generated using Monte Carlo simulations of geometric Brownian motion. The study tests the feasibility of the proposed model by applying it to a petrochemical company with larger carbon emissions than the allowances and commits to net zero greenhouse gas emissions by 2050. Scenario analysis is conducted to assess the total cost volatility due to changes in policy and carbon market movements. The results indicate that preemptive investments in carbon reduction are beneficial for the sustainable management of a carbon-intensive company, even in a situation where the carbon price increases by around 95 USD and the government protects the domestic industry through free allocation and relaxed emission regulations. The TRAM model can be utilized as an efficient decision support system for managing climate-related business risks toward a low-carbon future.

S18 Marginal Seas Dynamics and Modeling

Jan Harff (University of Szczecin),
Joanna Joanna Dudzińska-Nowak (University of Szczecin),
Wenyan Zhang (Helmholtz-Zentrum Hereon)

Marginal seas form a belt of transition between land and ocean. This chain, crossing all types of climate zones, having formed under various tectonic and geological conditions plays a crucial role in providing people with habitat, food, trade ways, and facilitated socio-economic networking. However, marginal seas are increasingly threatened by rising sea-level, floods, storms, tsunamis, coastal erosion and environmental hazards that endanger livelihoods. These threats have become even more visible in recent times in the face of climate change and anthropogenic impact including industrialization on the natural environment. The DDE Marginal Seas Task Group co-organizing the session is inviting geoscientists, modelers and IT scientists to discuss the opportunities to mitigate the threats by the application of advanced data acquisition such as monitoring concepts, big data analyses, numerical modelling, AI and Machine Learning for sustainable development of the coastal and marine realm of marginal seas. These models allow the reconstruction of the geological past as well as the generation of future scenarios that can be used for effective planning of sustainable development. The dynamics of sea level change, morphodynamic processes, anthropogenic pressure, model design and standardized solutions for sustainable development will play a principle role at the session. Networking of the international scientific teams will be explicitly supported.

S1801. Hydro-eco-morphodynamics in coastal systems

Wenyan Zhang (Helmholtz-Zentrum Hereon),
Peter Arlinghaus (Helmholtz-Centre Hereon)
Room: R8 2023-08-07 10:00

The morphology of the world’s coast is made up of a wide variety of landforms and bedforms manifested in a large spectrum of sizes and shapes depending on the forces acting on the coast (e.g. sea level, tides, waves, wind, river runoff, biota and humans), sediment supply and properties (e.g. grain size, lithology), and underlying geological constraints. The concept of equilibrium serves as a theoretical foundation for studying coastal morphological development in response to climate drivers. In the first part of my talk, I will give a brief overview of historical development of this concept, debates on its validity as well as possible ways to reconcile the debates. Afterwards I will present modeling studies in tide-dominated and wind wave-dominated coastal environments to provide a deepened understanding on this topic. In particular, I will focus on the role of biota in mediating long-term and large-scale coastal morphological development. Our results demonstrate that equilibrium development of coastal systems is largely dependent on the interaction between biotic and abiotic drivers.

S1802. The 4F numerical model – a solution for past reconstructions and future simulations of coastal dynamics

Jerzy J. Frydel (Polish Geological Institute - National Research Institute (PGI-NRI))
Room: R8 2023-08-07 10:20

The impact of climate change, glacial melting and rising global ocean levels, as well as increasing storm impacts on urbanised areas located adjacent to the coast should be considered, globally, as one of the key challenges facing humanity in the 21st century. Throughout the Holocene, the Baltic Sea level was subject to several major shifts caused by glacio-isostasy and climatically driven eustatic sea-level changes resulting in rapid coastline transgression and / or regression. Application of numerical modelling upon coastal zones of marginal seas allows reconstructions of geological processes including: ice sheet retreat / advance, marine regression / transgression and coastline accretion / erosion. Reconstructions and simulations of dynamics of those processes expressed by coastline (or glacial extent) location shift analyses is achievable with the 4F numerical model. The 4F model combines time, space and dynamics within its matrix, for the purpose of multiscale explorations, performed in uniform time-space environment, allowing for both past retrospections and future projections. Whenever possible, utilised mathematical formulas (equations / inequalities), derived from Frydel (2022), utilise variables, so that they can be universally applied upon simulations in different regions and time / space scales (depending on delivered data span). Alongshore quasi-continuous explorations are based on an infinite number of virtual averaged n-profiles, resulting from high-order polynomial regression (or in fact, any other functions) assigned for the glacial phase extents and coastlines (frontiers) themselves (primary inputs). Frontier reproduction is accurate (the coefficient of determination, R 2 0.9), thus assuring reliable secondary inputs for the further application in the process of integral calculus to frontier developmental dynamics modelling are achieved. The relevance of the disputed strategy lies in its direct applicability for both hindcasting dynamics (and extent) of past geologic processes and forecasting purposes, including demands of spatial planning in the coastal zones, based on hypsometry and sea levels projected upon RCP emission scenarios. Moreover, palaeodynamic analyses and simulations using the 4F numerical model, can be combined with past sea level changes and / or palaeogeographic reconstructions, adding a new dimension to palaeoclimatic investigations.

Literature:

Frydel J.J., 2022. https://doi.org/10.1017/qua.2021.64

S1803. Automatic classification of coastline and prediction of change - an exemplary study for the North Sea and Baltic Sea

Peter Arlinghaus (Helmholtz-Centre Hereon),
David Pogorzelski (Helmholtz-Zentrum Hereon, Institute of Coastal Systems - Analysis and Modeling),
Wenyan Zhang (Helmholtz-Zentrum hereon),
Corinna Schrum (Helmholtz-Zentrum Hereon, Institute of Coastal Systems - Analysis and Modeling)
Room: R8 2023-08-07 10:40

Automatic coastline classification based on machine learning is proven to be robust for sandy beaches on regional and global scales. However sandy beaches only make around one third of the world's ice-free shoreline. The rest consists of mudflats, cliffs, different types of vegetation and human constructions. Classification of these feature is more challenging. For instance, mild foreshore slopes resulting in large horizontal tidal excursions and high water content impede shoreline identification and classification in mudflats. Seasonal growth cycles pose difficulties in classification of vegetation.

Here we present a classification method based on a convolutional neural network, which is able to identify the mentioned features including sandy beaches. Based on Landsat and Sentinel (1 and 2) satellite images the coastline evolution in the past 3 decades will be presented for the North and the Baltic Sea, indicating hotspots of erosion and deposition.

The presented results will ultimately serve to predict possible coastline evolution in the North and Baltic Sea until the end of this century. For this purpose coastal climate data (winds, storms, waves, surges, currents) will be used as predictor variable for the identified coastal changes. First ideas on designing an explainable neuronal network to fulfil this task in a subsequent future step will be presented.

S1804. Holocene paleogeographic reconstruction of the Baltic Sea – data collection, synthesis and modeling

Jakub Miluch (University of Szczecin),
Wenyan Zhang (Helmholtz-Zentrum hereon),
Łukasz Maciąg (University of Szczecin, Institute of Marine and Environmental Sciences, Szczecin),
Peter Arlinghaus (Institute of Coastal Research, Helmholtz-Zentrum Hereon, Geesthacht, Germany),
Celine Denker (Univeristy of Hamburg, Faculty of Mathematics, Informatics and Natural Sciences,),
Labiqa Zahid (Univeristy of Hamburg, Faculty of Mathematics, Informatics and Natural Sciences,)
Room: R8 2023-08-07 11:00

Performing a paleogeographic reconstruction requires a specified set of input data: Digital Elevation Model (DEM), eustatic climate-controlled sea level curve, vertical crust movement (either isostatic or tectonic) rate as well as sediment thickness model. Combining such variables leads to generation of paleogeographic models. In the Baltic Sea research case, a set of paleogeographic maps was generated, mirroring the Holocene evolution of the area with 500 years time resolution. Data such as DEM, sea level curve or Glacial Isostatic Adjustment (GIA) model are either accessible online or available in literature, whereas generation of regional sediment thickness model of the Baltic Sea required a synthesis of various local-scale datasets. In overall, eight datasets from national sources or literature were gathered and required either digitalization or processing, such as recalculation from paleo-relief depth to thickness, interpolation from seismic lines or interpolation of point data. Data collection and synthesis allowed to cover majority of the Baltic Sea basin with the thickness grid. The "undersampled" areas required extrapolation using both co-kriging, by correlation of thickness with GIA-corrected bathymetry, as well as Random Forest Machine Learning algorithm. Moreover, a new eustatic regional sea level-curve was developed by correcting a global curve with precisely identified outcrop data corresponding to lake phase of the Baltic, when the basin was disconnected and sea level was independent from the global eustatics. Applied methodology acts as a pattern for the other paleogeographic reconstruction research in marginal seas.

S1805. Long term monitoring of Southern Baltic Sea, Poland, for future use in modeling

Rachel Jankowski (University of Szczecin),
Joanna Dudzińska-Nowak (University of Szczecin)
Room: R8 2023-08-07 11:20

The character of the Baltic Sea, a post glacial marginal sea with a small inlet to the Atlantic Ocean, its lack of tides, the varied composition of cliffs, dunes, spits, and barriers and lagoons, and addition of many anthropogenic engineering structures, makes an interesting model research site. An analysis and comparison of selected sites along the west Pomeranian Polish coast of the Baltic Sea using aerial photographs and sea level data allows for long term morphological changes to be determined. It is well known that the morphology of the southern Baltic coast is governed by glacial isostatic rebound and eustatic sea level rise. Previous research has determined that the main factors influencing erosion are rising water levels and frequency of storm surges as well waves height. Here, we inquire what lithologies and geological structures between the investigated sites are most resistant to erosion, and or support the acceleration of erosion or accretion. Selected parameters, as coastline movement, beach width, volume changes, as well as dune profiles as far as possible are compared at 10-year time steps. To determine historic sea levels with local sea level curves the samples of in situ organic materials, particularly tree trunks and peat, makes possible the calculation of 14c dates, were collected. An estimate of the rate of sea level transgression is calculated and compared to that observed during the present study. From our calculations in the 10-year difference 2012-2022 we observe both dune base line erosion and accretion within the range of 25 meters. For the whole of Western Pomeranian coast, the average trend of dune base movement is close to 0 m, which shows some kind of large-scale balance. Obtained results will be discussed in relation to sea level changes, geological setting and coastal development tendencies for more reliable future projections necessary for coastal protection and safety.

S1806. A Model area for marginal sea research - the Beibu Gulf, South China Sea

Jan Harff (University of Szczecin),
Ryszard K. Borowka (University of Szczecin, Institute of Marine and Environmental Sciences, Szczecin),
Peter Feldens (Leibniz Institute for Baltic Sea Research Warnemünde, Germany),
Hongjun Chen (Guangzhou Marine Geological Survey / China Geological Survey, Guangzhou, PR Chin),
Joanna Dudzińska-Nowak (University of Szczecin),
Andreas Groh (Technical University Dresden, Institute of Planetary Geodesy, Dresden, Germany),
Tao Jiang (China University of Geosciences, Wuhan, P.R. China),
Łukasz Maciąg (University of Szczecin, Institute of Marine Sciences, Szczecin, Poland),
Jakub Miluch (University of Szczecin),
Andrzej Osadczuk (University of Szczecin, Institute of Marine and Environmental Sciences, Szczecin),
Krystyna Osadczuk (University of Szczecin, Institute of Marine and Environmental Sciences, Szczecin),
Ping Xiong (China University of Geosciences, Wuhan, PR China),
Jinpeng Zhang (China Geological Survey),
Wenyan Zhang (Helmholtz-Zentrum hereon)
Room: R8 2023-08-07 11:40

Sustainable development of ocean margins, in particular marginal seas, require management strategies based on a deep understanding of the structure, genesis of sediments in the marine basins and the transition zones between marine and terrestrial realm. Advanced numerical models of complex geo-systems parameterized by multidisciplinary data support the generation of time-space environmental scenarios describing the evolution of marginal seas. As a case study we have investigated the Beibu Gulf at the northwestern continental margin of the South China Sea during the Last Glacial Cycle by a hybrid system approach of geostatistical studies of sediment architecture, sequence stratigraphy, isostatic modeling and morphodynamics of sediments. By geostatistical approaches for interpolation of seismic recordings structural models of the architecture of the sedimentary fill have been achieved. An age model generated using results of OSL- and AMS- dated sediments of an “master core” was used to extend the 3D structural model into a 4D space-time model of the sedimentary basin fill. Mapping of the 4D model to the global sea level curve led to the sequence stratigraphic model, which allows the interpretation of the geological evolution of the Beibu Gulf as a function of the late Pleistocene climate history. “Backstripping” procedures developed in sedimentary basin analysis were used to generate late Pleistocene paleo-geographic scenarios of the Beibu Gulf. Glacio-isostatic modeling of the basin history confirmed the assumption of “ocean syphoning” effects on the Holocene transgression / regression processes expressed by regional empirical relative sea-level curves.

The interdisciplinary numerical model and data-management system for the description of the evolution of a marginal sea can by principally also used for future projections based on climate change scenarios. In so far the project is fostering the development of new special direction of marine research - the investigation of marginal seas as the buffer zone between continents and oceans. The methodology serves as a blue print for applications in other parts of the global continental margins system.

S1807. Modeling Regional Sedimentary Sequence in the North of South China Sea in Late Pleistocene

Jinpeng Zhang (China Geological Survey),
Yufeng Wang (Guangzhou Marine Geological Survey, China Geological Survey, China),
Shiming Dong (Guangzhou Marine Geological Survey / China Geological Survey),
Wenbo Du (Guangzhou Marine Geological Survey, China Geological Survey, China),
Joanna Dudzinska-Nowak (University of Szczecin),
Yuan Cheng (Guangzhou Marine Geological Survey, China Geological Survey, China)
Room: R8 2023-08-07 14:00

Abstract:

In the late Pleistocene the continental shelf of north South China Sea, one of larger marginal seas in west Pacific Ocean, was shaped frequently by sea-level fluctuation, following the glacial cyclicity. Regional marine geoscientific surveys have collected geophysical data including single channel seismic profiles and sediment data from coring campaigns in different spatial scales. These surveys allow to combine geoscientific data sets to model structure and genesis of sedimentary sequences in order to understand the interplay of global change impact on the SCS shelf with regional drivers. On the regional level, the continental shelf of the northern SCS is remarkably influenced by series of rivers discharging to the sea terrestrial matter particularly the Pearl River, Red River. Exemplified by larger river, the Pearl River's paleo-channels and their changing patterns on the continental shelf have been investigated. The results show that the faraway forward channel extent to shelf edge in Last Glacial Maximum, responding to global ice volume expand in climate change. While, during the post-glacial history the sea-level rise played the dominant role driving the shift of paleo-river channels and coastline migration. This data interpretation contributes to a more comprehensive insight into the history of sediment accumulation at the northern continental margin of the SCS and the global and regional drivers.

Key words:

South China Sea, Sedimentary Sequence, Pearl River, River Channel, Late Pleistocene

S1808. Modelling of anthropogenic influenced estuarine evolution by using data-driven modelling approach, case study at the Lingding Bay (Pearl River Estuary)

Hongze Yu (Sun Yat-sen University),
Junjie Deng (Sun Yat-sen University),
Jingyu Hu (Sun Yat-sen University)
Room: R8 2023-08-07 14:20

The Pearl River Estuary (Lingding Bay) is a hotspot in the study of estuarine and coastal geomorphology. Previously, it was generally believed that Lingding Bay would evolve from an estuary to a river based on the linear trend of natural sedimentation and filling. Since 1980s, the disturbance of high-intensity human activities such as land reclamation, channel dredging and sand excavation, have altered the dynamic equilibrium between riverine sediment supply and sediment depositional fluxes in the estuary which has had a profound impact on the evolution trend of estuary evolution and filling. The relative contribution of natural filling processes of different human activities and climate change to estuary filling has not been well quantified. Both human activities and climate change have large stochastic variations, which cannot be accounted for by using stationary forcing. Data-driven models such as Bayesian Network can account for stochastic variations of the forcing and evolution, which have been widely and effectively used in predicting coastline changes, dune evolution, and storm surge surges. This paper intends to use a data-driven model to train and verify the evolution of estuary filling over the past century, and to predict future changes based on different scenarios of sea level rise and riverine sediment supply and so on. The date-driven modeling approach will be compared with the existed process-based morphodynamic model and the historical measurements. The data-driven model considering anthropogenic forcing will help quantify and understand the relative effects of different human activities and natural sediment depositional processes on estuarine evolution.

S1809. Temporal and Spatial Evolution and Numerical Simulation of Sandy Shoreline – A Case Study of Eastern Coast of Laizhou Bay, China

Ke Mu (CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation),
Jun Liu (Yantai Center of Coastal Zone Geological Survey, China Geological Survey, Yantai),
Xuelu Gao (CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation),
Yanfang Li (CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation),
Dong Li (CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation),
Sandra Donnici (Istituto di Geoscienze e Georisorse, Consiglio Nazionale delle Ricerche, Padova,),
Luigi Tosi (Istituto di Geoscienze e Georisorse, Consiglio Nazionale delle Ricerche, Padova,),
Cheng Tang (CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation)
Room: R8 2023-08-07 14:40

Many coastal regions of the world are confronted with changes in their coasts and coastal erosion caused by natural factors such as storm surges and human activities such as land reclamation. Laizhou Bay (LB) in the South Bohai Sea, China, has undergone rapid economic development in recent decades. Its coastal topography varied considerably due to changes in the environment and the intensification of human activities, which provides a good site for studying the impacts of natural factors and human activities on coastal evolution.

Using satellite imagery data (Landsat 5, 7, 8, Sentinel 2) from 1978 to 2022, a neural network classifier was constructed on the GEE platform, with tidal correction, to extract shoreline position in this region. Through digitized historical nautical charts from 1891 and 1940, a 130-year dataset on sandy shorelines was obtained. To analyze the shoreline evolution under the different anthropogenic scenarios, a ShorelineS model was applied using corresponding sediment input and structure parameters, as well as the wind and wave data setup.

The results indicate that decreased terrestrial sediment supply (caused by the upstream dam building), and coastal engineering construction (artificial islands for instance, indirectly alter the surrounding dynamical environment), are two main factors that affect the sandy coastline of east LB. In addition, the transformation of land use patterns and regional sea level rise could also have an impact on the evolution of the coastline, which requires further study in the future.

S1810. Holocene sedimentary distribution and morphological characteristics reworked by the East Asian monsoon dynamics in the Mekong River area, South Vietnam shelf

Xiao Wang (China University of Geosciences)
Room: R8 2023-08-07 15:00

The distribution of Holocene sediments in the Mekong River shelf area was investigated using shallow seismic and core evidence. The results reveal that the depocenter of Holocene sediments in the southern Vietnam shelf is predominantly situated within the Mekong estuary, Cape Ca Mau, and incised-valleys. Holocene sediments in the Mekong estuary and Cape Ca Mau are distributed primarily along the coastline, extending a short distance seaward. Holocene sediments in the Mekong estuary and Cape Ca Mau are distributed primarily along the coastline, extending a short distance seaward. Thin sediment deposits are present between the two depocenters. The distribution of Holocene sediments in the incised-valley is strongly influenced by the valley morphology, which is characteristically strip-shaped and can reach a maximum thickness of 20 m. In order to unravel possible driving mechanisms of the sedimentary distribution, 3-Dimensional numerical modeling was applied to investigate oceanographic and morphodynamic on the South Vietnam Shelf. The modeling results reveal that, in addition to sediment supply and sea level changes, the East Asian monsoon exerts a significant influence on sediment redistribution. During the rainy season, a substantial volume of sediments delivered by the Mekong River is deposited in the estuary, with some sediments migrating northeastward under the influence of the East Asian summer monsoon. In the dry season, previously deposited sediments are resuspended by waves and tides, and are transported southwestward under the influence of the East Asian winter monsoon. Upon reaching Cape Ca Mau, these sediments are re-deposited here due to substantially weakened hydrodynamic conditions, resulting in the formation of the sediment depocenter in Cape Ca Mau, which is located away from the Mekong River mouth. Overall, the modeling results indicate that the East Asian monsoon has a significant impact on sediment redistribution in the Mekong River shelf area, in addition to sediment supply and sea level changes.

S1811. Sedimentary characteristics of seamount-related contourite features at Changjun Seamount in the Beikang Bain, southern South China Sea

Zijun Liang (Sun Yat-sen University),
Hui Chen (School of Marine Science, Sun Yat-sen University, Zhuhai 519083, China),
Shan Liu (School of Marine Science, Sun Yat-sen University, Zhuhai 519083, China),
Ming Su (School of Marine Science, Sun Yat-sen University, Zhuhai 519083, China)
Room: R8 2023-08-07 15:20

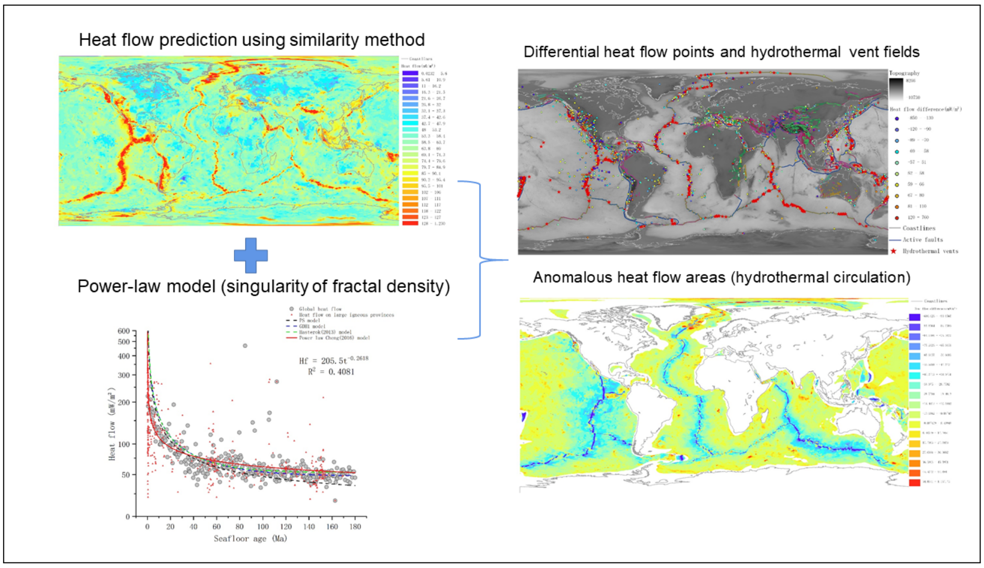
Seamount-related contourite features, generated under the interaction between seafloor topography and ocean bottom currents, are increasingly recognized as a sensitive indicator of flow dynamic changes. The South China Sea (SCS) is the largest semi-enclosed marginal sea in the West Pacific, containing numerous scattered seamounts. However, high-resolution bathymetries or seamount-related depositional features within the Beikang Basin in the southern SCS are rarely documented.

This work depicted Changjun Seamount in the Beikang Basin and identified the relevant contourite features using high-resolution bathymetry data and 2D seismic profiles. There is a small-scale contourite depositional system (CDS) consisting of four moats and four drifts on Changjun Seamount. It is evident from the internal structure and spatial distribution of the drift-moat complexes that SCS Deep Water currents flow eastward. Meanwhile, the synchronous development of multiple moat-drift complexes, developed under multi-scale oceanographic processes (including deep sea circulation currents and tides, deep reaching eddies, etc.) as well as seamount topography forcing, may indicate complex local circulation hydrodynamics. Nevertheless, future numerical simulations and in situ oceanographic observations are required, as always, for further confirmation of this hypothesis.

S1812. Ocean Singularity Analysis and Global Heat Flow Prediction Based on Similarity Method

Yang Zhang (School of Earth Science and Engineering, Sun Yat-Sen University),
Junjie Ji (Sun Yat-sen University),
Qiuming Cheng (China University of Geosciences/Sun Yat-sen University)
Room: R8 2023-08-07 15:22

Ocean heat flow research and global heat flow prediction have been paid much attention, and can be used to study the thermal structures and dynamic processes of the earth. Here we use the singularity theory of fractal density to study the ocean heat flow, and use the latest heat flow and geological and geophysical data to build power-law models. Compared with previous plate models, the power-law model can better describe the trend of ocean heat flow, and overcome the shortcomings of higher heat flow values in young age (<35 Ma) and lower heat flow values in old age (>35 Ma). Then we find that the heat flow of divergent continental margins is similar to the oceanic heat flow, which can be described by power-law model. Moreover, we find that the corrected ocean bathymetry and Bouguer gravity anomaly also conform to the power-law models. Through analysis, it is found that the anomalous subsidence areas of the ocean are related to the large igneous provinces, hot spots and seamounts, and the Bouguer gravity anomaly conforms to the power-law model before 33 Ma seafloor age. Finally, the similarity method is used to predict the global heat flow based on the latest geological and geophysical observables filtered by correlation analysis. The predicted heat flow outlines the general trends of low heat flow in tectonically stable and ancient regions and high heat flow in tectonically active and unstable regions. The regions with large differences between predicted heat flow and actual heat flow are highly related to the hydrothermal vent fields. Combining the predicted heat flow with heat flow models can better describe the anomalous heat flow regions, for example, the regions near the mid-ocean ridges will be affected by hydrothermal circulation.



S1813. Summer flow dynamic characteristics of western boundary currents in the northwestern South China Sea

Hongtao Mai (SYSU),
Hui Chen (School of Marine Science, Sun Yat-sen University, Zhuhai 519083, China),
Dongxiao Wang (School of Marine Sciences, Sun Yat-sen University, Zhuhai 519082, Guangdong, Chi)
Room: R8 2023-08-07 15:24

Summer flow dynamic characteristics of western boundary currents in the northwestern South China Sea

Hongtao Mai1,2, Hui Chen1,2, Dongxiao Wang1,2

1 School of Marine Sciences, Sun Yat-sen University, Zhuhai 519082, Guangdong, China

2 Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519000, Guangdong, China

Deep-water canyons provide sediment transport conduits that can effectively disturb geostrophic currents and promote vertical exchange of sediments and energy across the continental slope. Based on summer cruise observations (CTD, Lowered ADCP, Shipborne ADCP), high resolution bathymetry and HYCOM reanalysis, this research investigates the (seasonal) hydrodynamic characteristics of western boundary currents (WBCs) in the northwestern South China Sea (SCS) within a submarine canyon.

The mainstream of SCS WBCs flows from south to north across the East Vietnam Sea, with their core depth matching that of the SCS Intermediate Water. Branches of the mainstream currents enter the Central Canyon along with the 700-1200m isobath, developing f1u. On the NW side steep slope of the midstream canyon, f1u collides with the oncoming SCS surface currents, causing intense mixing at a depth of 500 m. Mixed dissipation splits f1u into branches. One branch shifts towards NE direction and becomes f1m, while the other turns round, flows towards SW direction along the SE side of the up-midstream canyon as f2.

WBC velocity and direction obviously change due to the strong mixing in the canyon topography, resulting in local circulation patterns. This study has implications for understanding the evolution and topography-current interaction processes of WBCs over continental margins.

S1814. Dynamics of the southern Baltic Sea coastline changes

Joanna Dudzińska-Nowak (University of Szczecin)
Room: R8 2023-08-07 15:26

Morphodynamic coastal processes of the southern Baltic Sea are determined by a complex interplay of the geological setting, eustatic sea-level change, glacio-isostatic adjustment, wave driven sediment dynamics, storm surges and aeolian processes.

The shape of the present Southern Baltic Sea coast is an overlap result of complex of settings and processes affecting lithosphere, hydrosphere, atmosphere as well as biosphere and anthroposphere, which occur in various temporal and spatial scales from millennia to hours and in spatial scale from kilometers to centimeters. The results of the interference is visible as a morphology pattern of beach and dune / cliff as well as a nearshore area.

Use of remote sensing, mainly high resolution aerial photographs and laser scanning data, allow to indicate precisely the characteristic features of the morphology of the terrain, which if repeated, can be treated as indicators of the changes.

Presented research shows the methodology and results of conducted coastal dynamics classification of the southern Baltic Sea coasts, focusing not only on direction of movement and rate of change but also on dynamic of change in different time scales. The main criterion for this classification is the determination of two numerical indicators for each examined section of the coast, synthesizing the image of registered changes in the dune / cliff base line. Changes with medium and high dynamics of both accumulation and erosion processes dominate. The sections with low dynamics of the processes taking place constitute only less than 20% of the studied area. The coast in dynamic equilibrium cover less than 10%. Spatial distribution of individual classes of shore dynamics proves a large diversity and space and time variability of morphodynamic processes along the entire investigated section of the shore, which may substantially bias of the future scenarios.

Further interdisciplinary studies based of measured and modeled data of the long-term variations of waves regime, water level changes, storm surges structure and long-shore sediment transport analyses in relation to coastal zone changes as a consequences of such phenomena's are necessary and strongly recommended in order to reveal a mechanism of the coast development.

S1815. An integrated risk assessment of coastal erosion based on fuzzy set theory along Guangdong coast, southeaster China

Shilong Luo (South China sea bureau, ministry of natural resources)
Room: R8 2023-08-08 14:40

Coastal erosion and accumulation has always existed and contributed to the shaping of the present coastlines. However, coastal erosion now is largely intensified due to human activities, causing severe land loss and aggravated financial burden, which is not acceptable. This paper presents an integrated methodology for risk assessment of coastal erosion, aiming at assessing and mapping coastal erosion risk by integrating hazard, exposure and coping capacity to give risk priority and risk mappings to guide coastal erosion risk management. Delphi method, the analytic hierarchy process and the fuzzy set theory were adopted in the study to establish the methodology. Guangdong province situated at southeast China was taken as a case study; the final result was effectively supported by the data and provided lots of important information for coastal management. The final risk priority recorded the risk disparities from place to place and ranked them with distinct levels to emphasize those places of higher risk levels which needed urgent mitigation measures for risk reduction, and risk mapping gave a visual and clear impression of risk distribution. The risk priority and mapping was only the first step to coastal erosion management; identification of priority areas, implementation of mitigation measures and regular maintenance and monitoring are advanced jobs need to be done.

S19 Reservoir Geostatistics for Geo-Energy Applications

Juliana Leung (University of Alberta),
Sanjay Srinivasan (Pennsylvania State University),
Baehyun Min (Ewha Womans University)

Geostatistics has been widely applied for modelling subsurface flow systems encountered in many geo-energy applications, including but not limited to oil and gas extraction, geological storage of CO2, and geothermal reservoir development. Traditional workflow entails generating multiple reservoir models conditioned to data from diverse scales and sources, subjecting these models to flow or transport modelling, and updating the models with new dynamic data. We are soliciting abstracts that showcase recent advancements in the characterization of complex reservoirs, integration of multifaceted data from diverse sources recognizing the scale and precision of such data, upscaling of flow or transport properties, and coupling machine learning with traditional modelling workflows for uncertainty quantification. Interesting case studies, as well as novel theoretical and computational developments, are welcome.

S1901. Simulating Fracture Networks Using Sequential Gaussian Simulation, LightGBM, and Deep Learning Approaches for Geological Data from Kazakhstan

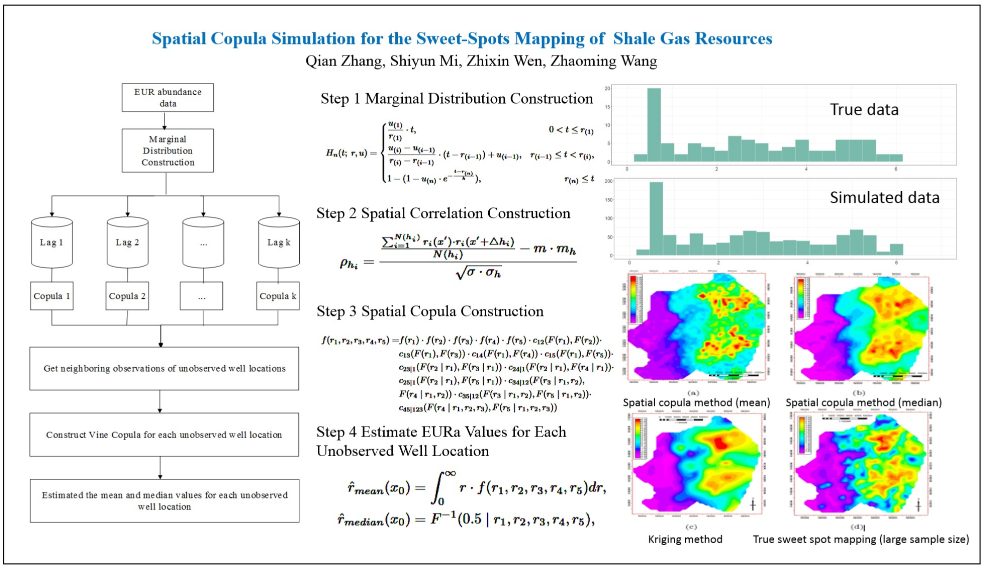
Timur Merembayev (Nazarbayev Univesity),
Yerlan Amanbek (Nazarbayev University)
Room: R3 2023-08-07 10:00

The main goal of fracture modeling is to build a network of subsurface fractures close to reality which is useful for the hydrocarbon production, reservoir evaluation, CO2 sequestration, and groundwater remediation in porous media. Simulating a fracture geometry system using the geotechnical aspects requires significant computational resources. Traditional stochastic algorithms for fracture network characterization are based on statistical inference from fracture inputs. In this talk, we present a numerical model of a fracture network using the Sequential Gaussian simulation (SGS), LightGBM, and deep learning approaches for the realistic data. SGS is a geostatistical tool for designing a fracture network in porous media. LigthGBM is the framework based on gradient boosting algorithms which is good for the tabular data of azimuth and distance. Deep learning approaches are also applied for the detection of fracture network using images. A comparison of methods results will be presented for the geological data of Kazakhstan.

S1902. Spatial Copula Simulation for the Sweet-Spots Mapping of Shale gas Resources: A Case Study

Qian Zhang (PetroChina),
Shiyun Mi (RIPED of PetroChina),
Zhixin Wen (Research Institute of Petroleum Exploration & Development),
Zhaoming Wang (Research Institute of Petroleum Exploration & Development)
Room: R3 2023-08-07 10:20

This study developed a new sweet-spots mapping approach based on copula for assessment of shale gas resources. Variations of spatial correlation structure was ignored in the conventional kriged models. Inaccurate models of the unconventional resources under small sample size can lead to failure of mapping the sweet-spots. In this research, copula-based geostatistical model was used to predict the estimated ultimate recovery abundance (EURa). The model is applied to capture the variable spatial dependence structures. In this research, Fiducial Empirical Distribution (FED) was used to model the distribution of EURa. FED was also applied to acquire simulate random samples from the small sample size data. In most cases, the EURa data is not normally distributed. The fiducial empirical distribution is more flexible than the conventional nonparametric distributions, such as weighted kernel distribution and empirical distribution. The spatial vine-copula model is applied to capture the variable spatial dependence structures. By estimating EURa values at dense grid points and integrate them appropriately, the location of sweet spot can be predicted more precisely because the estimated values have character of EURa spatial dependence, variability and distribution structure. It is showed the new method can effectively express the whole and the local spatial correlation between the sweet-spots, and more accurate sweet-spot mapping than a conventional lognormal kriged model.



S1903. A Recursively Enhanced GRU Model for Real-Time Production Forecasting in Central Montney Shale Gas Reservoirs

Ziming Xu (University of Alberta),
Juliana Leung (University of Alberta)
Room: R3 2023-08-07 10:40

Hydraulic fracturing is essential in shale gas exploration, but it creates a complex fracture system and significant heterogeneity in the reservoir. These challenges can make reservoir characterization and production forecasting difficult using traditional methods that involve complex 3D geological models and numerical flow simulation. However, RNN-based proxy models have recently been widely used to improve production forecasting efficiency. These models learn from historical production data and reservoir characteristics from a database to bypass the need for complex models and constraints. Despite their usefulness, these models can have difficulty capturing the subsurface uncertainties, leading to prediction inaccuracies. In this work, ongoing production data is used to improve the accuracy of the prediction. However, the prevalent RNN-based predicting frameworks often suffer from heavy carry-on errors compromising prediction accuracy.

Our prior research introduced a novel predicting framework called Recursively Updated Forecasting (RUF), which leverages ongoing production data to enhance prediction outcomes without retraining the model. In this study, we integrate RUF with the Gated Recurrent Unit (GRU) and assess its adaptability using a field dataset assembled from 2699 shale gas wells in the Montney Formation in Western Canada. This data encompasses 40 production-related features, including reservoir properties, completion data and wellhead information. For new wells without historical production data, we utilize Gradient Boosting (GB) to initialize the GRU-RUF model. The study evaluates the model's performance in predicting 36-month production using P10, P50, and P90 values of the Root Mean Square Error (RMSE) from 500 test samples. The results show that the prediction accuracy dramatically improves when introducing the production data stream. In addition, we compare RUF with traditional Decline Curve Analysis (DCA) and other RNN-based prediction frameworks. Our proposed model demonstrates superior stability and reliability compared to several alternative prediction frameworks tested in this work. The developed time-series forecasting model can also be adapted for different applications involving multi-faceted inputs.

S1904. Efficient Design of Warm-VAPEX using a Neural Network Coupled with a Multi-objective Optimization Algorithm Under Geological Uncertainty at an Oilsands Reservoir

Seoyoon Kwon (Ewha Womans University),
Juliana Y. Leung (University of Alberta),
Baehyun Min (Ewha Womans University)
Room: R3 2023-08-07 11:00

This study proposes a hybrid workflow using a neural network coupled with a multi-objective optimization algorithm for finding a set of Pareto-optimal operating conditions of the Warm vapor extraction process with dimethyl ether (Warm-VAPEX) at an oilsands reservoir. The proposed workflow is tested for a reservoir model representing the characteristics of Athabasca oilsands, Western Canada. We formulate a set of decision variables (e.g., injection pressure, injection temperature, and solvent fraction) and objective functions (e.g., recovery factor, solvent efficiency, and enthalpy for internal energy). The multi-objective optimization algorithm evaluates the performance of each decision-variable set and searches for the Pareto-optimal front in the objective space. A deep neural network-based proxy is designed to assess the objective functions cost-effectively. The final Pareto-optimal front is derived by repeating the hybrid workflow in both decision and objective spaces. We also analyze the effects of heterogeneities on bitumen production. The proposed workflow can contribute to the efficient design of a solvent-based heavy oil recovery scheme with a balance between accuracy and computational cost.

S1905. Performance Evaluation on CO2 Carbonated Water-Alternating-Gas on Enhanced Oil Recovery and Geological Carbon Storage

Minsoo Ji (Ewha Womans University),
Seoyoon Kwon (Ewha Womans University),
Byungin Choi (Computer Modelling Group),
Baehyun Min (Ewha Womans University)
Room: R3 2023-08-07 11:20

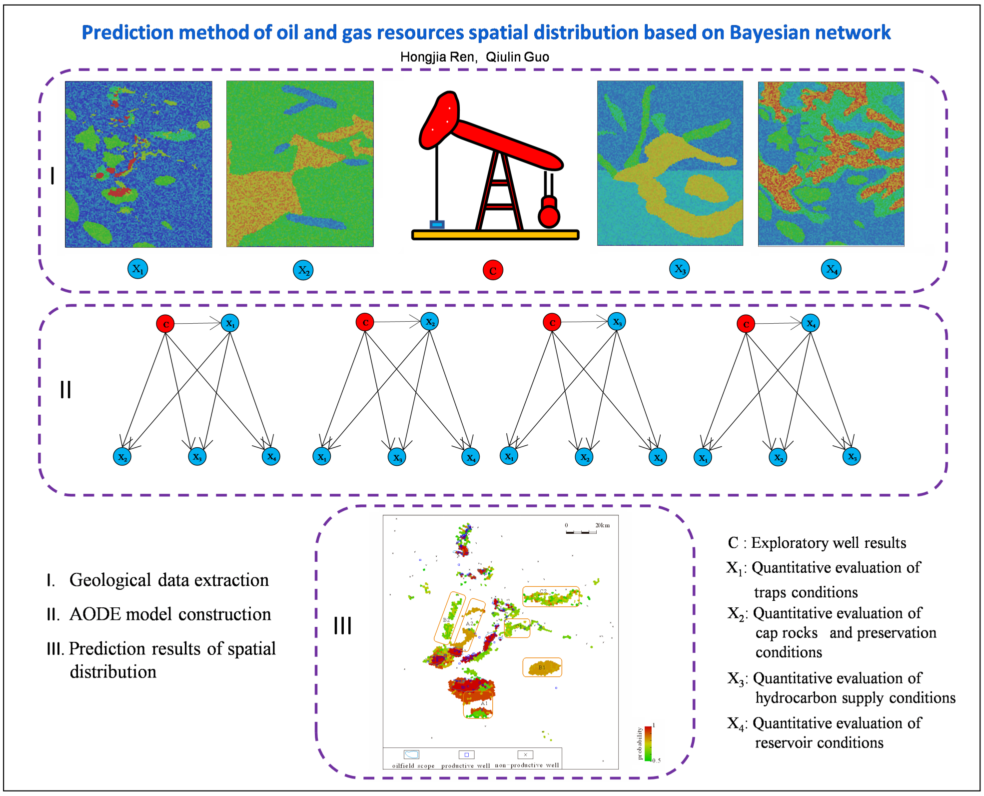
This study proposes a novel CO2-Carbonated Water-Alternating-Gas (CWAG) process that combines carbonated water injection with CO2-WAG to enhance both oil recovery and geological carbon storage. The carbonated water injection process is designed for a multi-component CO2/brine/oil system, and a fluid model is developed based on data acquired from Weyburn, Canada. A case study using a compositional reservoir simulator is performed to evaluate the effectiveness of CO2-CWAG by monitoring the CO2 plume propagation controlled by different trapping mechanisms, such as structural and stratigraphic, residual, and solubility trapping mechanisms. The results show that CO2-CWAG improves oil recovery and CO2 retention compared to conventional CO2-WAG. A sensitivity analysis is conducted to investigate the impact of reservoir properties and operating conditions on the performance of CO2-CWAG. The numerical simulation results demonstrate that the proposed CO2-CWAG process is an advanced enhanced oil recovery scheme that aligns with carbon neutrality. Therefore, this study provides a promising approach for simultaneously enhancing oil recovery and carbon storage.

S1906. Prediction method of oil and gas resources spatial distribution based on Bayesian network

Qiulin Guo (PetroChina),
Hongjia Ren,
Jingdu Yu (Petrochina)
Room: R3 2023-08-07 11:40

Effective prediction of hydrocarbon spatial distribution is of vital importance to optimize hydrocarbon exploration deployment and improve drilling success and economic efficiency. Determining the location of oil and gas existence is a comprehensive and complex uncertainty problem. In this paper, an intelligent prediction method of oil and gas spatial distribution based on Bayesian network classifier is proposed. The proposed averaged one-dependence estimators (AODE) method and workflow are further illustrated by taking the Sangonghe Formation in the hinterland of the Junggar Basin, China, as an example. The prediction results show that: (1) On the data set of 203 exploratory wells, the accuracy of the AODE model can reach 85.2%, which is higher than the state-of-the-art methods; (2) In the discovered reserves area, the predicted results are highly consistent with the exploration results; Outside the reserve area, three types of favorable areas are predicted. Among them, type A, with a high degree of grasp, is the key area for edge expansion and rolling exploration of the next oilfield; Type B, with relatively high uncertainty, is the key area for risk exploration and exploration in the next step; Type C, with complex geological conditions, needs further research and new understanding before formulating drilling deployment and submitting reserves.

The application results show that the AODE model can effectively predict the spatial distribution of oil and gas and visualize the geological exploration risks, so as to optimize the drilling strategy and improve the economic efficiency.



S1907. Conversion of a Depleted Gas Field into a Geological CO2 Storage Site: A Case Study of the Donghae-1 Gas Field, Republic of Korea

Suin Choi (Ewha Womans University),
Seoyoon Kwon (Ewha Womans University),
Minsoo Ji (Ewha Womans University),
Min Kim (Severe Storm Research Center, Ewha Womans University),
Baehyun Min (Ewha Womans University)
Room: R3 2023-08-07 12:00

This study investigates the feasibility of converting the depleted Donghae-1 gas field into the first geological CO2 storage site in South Korea. By using a history-matched reservoir model that produced natural gas for 18 years, we conduct simulation studies to evaluate the CO2 storage capacity, efficiency, and stability. Additionally, we design a new scenario to investigate the efficacy of hydrocarbon production followed by geological carbon storage in a systematic manner by replacing the existing production wells with a horizontal well. We also explore the potential of the underlying saline aquifer as CO2 Storage site by running CO2 injection simulations to determine the optimal injection rate and pressure, as well as injection well trajectory, for maximizing CO2 storage. We track CO2 plume propagation based on diffusion caused by pressure gradient and buoyancy of the injected CO2. The CO2 storage efficiency of each scenario is assessed in terms of three trapping mechanisms (i.e., structural, stratigraphic, residual, and solubility trapping mechanisms). This study's results can serve as preliminary data for the planned Donghae-1 carbon capture and storage (CCS) project after 2025.

S1908. Data cross-validation technique to improve the accuracy of geological model in terms of reservoir properties prediction

Elena Kharyba (NTC NIS - Naftagas LLC Novi Sad),
Stefan Stankovic (NTC NIS - Naftagas LLC Novi Sad),
Leonid Stulov (NTC NIS - Naftagas LLC Novi Sad),
Vladimir Miroshkin (LLC "GAZPROMNEFT SCIENCE& TECHNOLOGY CENTRE"),
Rustam Gaynanshin (LLC "GAZPROMNEFT SCIENCE& TECHNOLOGY CENTRE")
Room: R3 2023-08-07 12:02

The reliability of reservoir predictions depends on accurately describing the reservoir geology in a model. In the case of a complex object with deep-water and shelf sedimentation environments, it's essential to consider the diversity of facies and patterns of changes in properties by facies when distributing petrophysical properties in the model.

To address this, the present study used cross-validation analysis to select the optimal algorithm and parameters for 3D porosity modeling based on log data and a 2D effective thickness trend map. The study was conducted on one sector of a large oil field with thousands of production wells that still carries great geological uncertainty.

The approach was applied to a deep-water part of the clinoform complex. The study found that the map of effective thicknesses can be used as a trend for porosity modeling, but it's necessary to control and compare the modeling results with well information to select the optimal modeling algorithm.

The study used a cross-validation technique, excluding seven wells located evenly over the area in different parts of facies zones. Porosity modeling was performed for seven options, and for each option, an average map of porosity was made, synthetic porosity logs and the average porosity for the seven wells were calculated. The difference parameter between the average synthetic and real data porosity for all seven wells was calculated in absolute and relative values. The average value and standard deviation were analyzed based on the delta values of the seven wells participating in the analysis.

The results showed that the Gaussian Sequential Simulation algorithm using the Co-kriging option with two secondary variables was the most optimal. The secondary variables included the horizontal trend of the effective thickness and the vertical function of reservoir porosity for each facies zone. The collocated co-kriging coefficient was 0.9, and the form of secondary variables transformation was normal score transform with a variance reduction factor of 0.6.

This study is crucial for understanding the geological structure and reservoir properties and making decisions for further drilling plans. Additionally, the study provides guidance for addressing uncertainties and choosing modeling approaches.

S22 Flow simulations in geo-energy applications and gas storage

Francesca Watson (SINTEF Digital),
Elyes Ahmed (SINTEF Digital)

Numerical flow simulations are crucial for all kinds of geo-energy and geological gas storage applications. This is a broad area with aspects relating to spatial discretisation of the flow equations over complicated domains, numerical solution of potentially large systems of equations, models with high levels of heterogeneity and limited data availability, to name but a few. In this session we invite speakers to present recent advances on topics regarding flow simulations in the subsurface. Contributions are welcome in areas such as effective solution methods, data-driven modelling, flow simulations for various applications e.g. geothermal, fractured reservoirs, CO2 storage, as well as anything else relevant to the field.

S2201. Use of hybrid mesh for flow simulations: application on a North Sea CO2 storage study

Margaux Raguenel (TotalEnergies),
Antoine Mazuyer (TotalEnergies),
Wan-Chiu Li (Tessael),
Cédric Borgese (Tessael)
Room: R9 2023-08-10 10:00

Carbon Capture and storage is a critical technology for mitigating climate change. To follow the recommendations of climate agencies, the number of projects has to increase and the decision-making processes need to be accelerated.

To address this challenge, a precise 3D modeling of the subsurface geology and of the physical behavior of the injected CO2 is crucial. An approach using multi-physics simulations (coupled flow and geomechanical effects) has been recently of interest for these projects. These fully-coupled simulations require a unique mesh, as a support suitable for the different physical processes, especially to model flow and geomechanics around faults. Furthermore, this mesh should respect some geometrical constraints, such as cell conformity and respect of the fault geometry.

To overcome the limitations of structured meshes (non conformity and inaccurate description of geological faults) and of tetrahedral meshes (need for complex numerical schemes to account for geometrical deformation), we propose to use a new generation of hybrid meshes, which consist of different predefined types of cells (hexahedra, pyramids and tetrahedra). These cells are built in an optimal and adaptive framework to respect the above mentioned constraints, and they are compatible with classical discretization schemes (Finite Element Methods for mechanics and Finite Volume Method for flow). So, hybrid meshes appear to be of particular interest in terms of flexibility and performances.

In this work, we propose first to present the construction of such a mesh on a model using the constraints of a North Sea CO2 storage study, where the aquifer targeted for CO2 injection is highly faulted. In particular, we show that this type of mesh is able to represent the structural complexity of the underlying geological target, with a focus on the management of the fault pattern. Then we show that this mesh is successfully used to perform flow simulations, and comparisons are done between several discretization schemes and several simulation tools, which confirm the consistency of flow simulation results, opening the path to fully coupled simulations.

S2202. Rapid Prototyping of Subsurface CO2 Sequestration with Flow Diagnostics

Dmytro Petrovskyy (Delft University of Technology),
Sebastian Geiger (Delft University of Technology),
Carl Jacquemyn (Imperial College London),
Matthew D Jackson (Imperial College London),
Gary J Hampson (Imperial College London),
Julio D Machado Silva (University of Calgary),
Sicilia Judice (University of Calgary),
Fazilatur Rahman (University of Calgary),
Mario Costa Sousa (University of Calgary)
Room: R9 2023-08-10 10:20

Traditional reservoir modelling workflows are often too restrictive for substantial modifications to the underlying geological hypotheses, and therefore restrict the breadth of interpretations considered in subsurface evaluation (Bentley and Smith, 2008). Emerging geoengineering challenges associated with the energy transition, such as carbon dioxide sequestration, hydrogen storage, and aquifer thermal energy storage, require extensive study of geological heterogeneity and its impact on storage and flow properties from sparse data (e.g. Alshakri et al., 2023; Jackson et al., 2022).

We propose an alternative and complementary approach, referred to as rapid reservoir modelling (RRM), that aims to address the drawbacks of conventional modelling methods. RRM differentiates itself by using sketch-based interface and modelling (SBIM) technology that constructs 3D surfaces from simple 2D sketches and geologically consistent surface-interaction operators that allow for fast and intuitive subsurface modelling. (Jacquemyn et al., 2021b). Here, we further advance RRM capabilities by connecting on-the-fly flow diagnostics (FD) computations to streamline and enhance the prediction of subsurface flow (Petrovskyy et al., 2023). The application of FD methods is particularly efficient when ranking and comparing model ensembles (Watson et al., 2022). FD brings the prototyping approach of SBIM into the dynamic space of injection/production and well patterns.

Our FD implementation enables the specification of custom open-boundary conditions that are usually closed in previous FD studies, rapid and on-demand grid resampling for higher accuracy of calculations, and an interactive graphical interface to navigate and specify the well placements and perforation directly in 3D. We demonstrate the FD dynamic modelling approach using a model of deep water slope channels (Jacquemyn et al., 2021). Specifically, we study the impact of well placement on the viscous-to-gravity force balance during the active CO2 injection phase, to illustrate how different offtake well placements influence the viscous flow regime. FD calculations acquire fast and informative flow prediction in a matter of seconds. Such calculations facilitate rapid prototyping, screening, and ranking of reservoir models. The FD computation may serve as the guide and first-order approximation to inform expensive full-physical modelling and simulation.

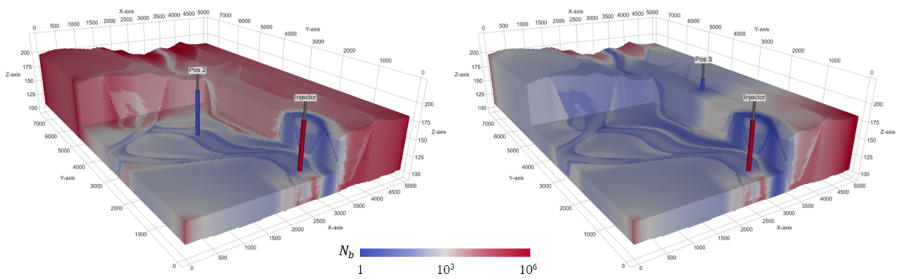


Figure 1 – Flow diagnostics (FD) calculations of gravity number N_b in deep water channelised reservoir model with vertical CO_2 injection (in red) and water offtake (in blue) wells. Lower values of N_b correspond to regions where flow is primarily driven by viscous forces. Geological heterogeneity significantly influences the dominance of viscous forces for different locations of the offtake well (left vs. right image).

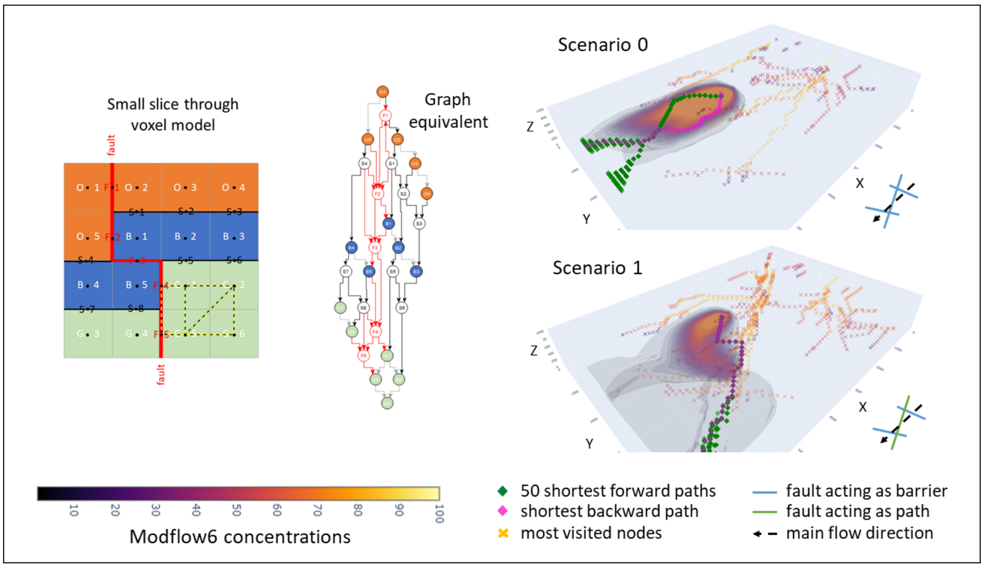
S2203. How good can graph distances approximate flow and transport simulations in faulted aquifers?

Guillaume Pirot (The University of Western Australia),
Kerry Bardot (The University of Western Australia),
Mark W Jessell (The University of Western Australia, School of Earth Sciences, Crawley, WA 6009),
Jim McCallum (The University of Western Australia)
Room: R9 2023-08-10 10:40

In groundwater application, data and knowledge are often very limited. Thus, the characterization of subsurface properties remains highly uncertain and proper predictive uncertainty quantification requires to consider and propagate, conceptual, parametric and data uncertainties. However, groundwater flow and transport simulations rapidly become computationally expensive as we increase further the level of spatio-temporal discretization. It can then be a strong limiting factor for proper uncertainty quantification. One way to alleviate these constraints is to use faster but approximate solutions. In some particular settings, such as in recent work by Rizzo and de Barros (2017), Chen et al. (2020) and Tang et al. (2021), the use of graph distances for approximations of flow and transport seems promising.

Here we propose to assess the potential of such methods in the context of faulted heterogeneous aquifers. Varying the number of faults, and their connectivity (acting as a barrier or a preferential pathway) and geometric characteristics, we run a sensitivity analysis on a contaminant transport experiment. Each hydrogeological model is represented dually as a meshed voxet and as a graph, to allow for both full-physics numerical simulations and graph-distance computations.

We acknowledge the support from the ARC-funded Loop: Enabling Stochastic 3D Geological Modelling consortia (LP170100985) and the Mineral Exploration Cooperative Research Centre. This is MinEx CRC Document 2023/**.



S2204. Fast Modeling of Fault Leakage during CO2 Storage

Hariharan Ramachandran (Institute of GeoEnergy Engineering, Heriot-Watt University, Edinburgh, UK),
Florian Doster (Heriot-Watt University),
Sebastian Geiger (Delft University of Technology)
Room: R9 2023-08-10 11:00

Storing carbon dioxide (CO2) in geological formations is a crucial method for mitigating climate change. However, CO2 leakage during injection and storage poses a significant risk, and faults represent a particular concern as they can act as major structural traps or connect pathways to shallower geological layers. Therefore, it is vital to understand the behavior of faults and related structures such as micro-cracks, joints, fracture networks, deformation bands, and fault cores to assess the risk of CO2 leakage and ensure safe and effective storage.

Accurately simulating fault-related properties at different scales is crucial to predict the consequences of CO2 injection and storage. However, this task can be challenging, particularly in the early stages of a storage project when knowledge of the storage reservoir is limited, and the cost of obtaining high-quality well logs, cores, and seismic data is high. To address this issue, this study proposes a workflow for ultra-fast screening of fault leakage risk during injection and storage at the concept selection stage. The workflow employs a vertically integrated reservoir model coupled with an upscaled fault leakage function.

Simulation results of various CO2 injection scenarios in a storage reservoir with potential for fault leakage demonstrate that the workflow can produce reliable saturation profiles with substantially reduced computation time compared to fine-scale models. Matching CO2 saturation profiles obtained from fine-scale and vertically integrated models adds confidence to the proposed workflow. The fast workflow presented in this study provides a useful tool for identifying the uncertainties associated with key fault parameters, reservoir architecture, and other constitutive relations that affect the behavior of the storage reservoir and potential fault leakage outcomes. By incorporating this workflow at the concept selection stage, stakeholders can quickly assess the risk of CO2 leakage and evaluate the feasibility of the storage site. Overall, the proposed workflow provides a cost-effective and efficient method for screening fault leakage risk during CO2 injection and storage, helping to ensure safe and effective carbon storage.

S2205. Sparsified coarse-scale operators for multiscale methods

Omar Chaabi (Khalifa University),
Mohammed Al Kobaisi (Khalifa University)
Room: R9 2023-08-10 11:20

Generating a solution for the pressure in reservoir models with fine-scale heterogeneities, modeled at a meter scale, can be compute-intensive. Multiscale methods aim at reducing the runtime for the solution of Poisson-type equations by solving coarsened flow problems that account for fine-scale variations. To obtain a coarse-scale operator one needs to project the fine-scale operator using prolongation and restriction transfer operators. The choice of transfer operators determines the sparsity pattern and the quality of the coarse-scale operator. Multiscale restriction-smoothed basis (MsRSB) method and the classical multiscale finite volume (MSFV) method provide coarse-scale operators with MPFA-like stencils even when TPFA fine-scale operator is used at the fine level. Typically, the use of a denser MPFA operator is justified by its ability to capture tensorial permeability flow and non-K-orthogonal grids unlike TPFA or multiscale coarse-scale operators.

In this work, we show an algorithm that reduces the sparsity pattern of the coarse-scale operator for multiscale methods irrespective of the choice of prolongation and restriction operators. Once the choice is made on the transfer operators, the algorithm sparsifies the coarse-scale operator to the desired sparsity pattern. We test the algorithm on the state-of-the-art MsRSB method to solve the elliptic pressure equation for the full 3D 10th SPE comparative solution project. The SPE10 problem has uniform structured grids with diagonal permeability tensor and this motivated our choice of reducing the coarse operator sparsity to a TPFA-like stencil structure. Results show the potential of this approach as we were able to get excellent solutions with tolerable errors compared to the reference solution and with less runtime using the sparsified coarse-scale operator compared to the original MsRSB coarse-scale operator.

S2206. Learning the Solution Operator of Flux Functions for the Parametric Transport Equation in Porous Media Using Physics-Informed DeepONet

Waleed Diab (Khalifa University),
Mohammed Al Kobaisi (Khalifa University)
Room: R9 2023-08-10 13:40

Traditional finite difference and finite volume methods for simulating transport in porous media can be computationally expensive. Recent advances in machine learning for scientific computing may help speed up the simulation of the transport equation in porous media.

DeepONets has recently emerged as a powerful tool for accelerating the solution of partial differential equations (PDEs). This is achieved by learning implicit operators (mapping between function spaces) of deterministic PDEs. In this work, we learn the mapping between the space of non-convex flux functions of the Buckley-Leverett PDE and the space of solutions (saturations). We leverage Physics-Informed DeepONets (PI-DeepONets) to achieve this mapping without any labeled data. As opposed to Physics-Informed Neural Network (PINN) approach that solves the PDE deterministically, that is it must be re-trained for every new mobility ratio, PI-DeepONets solves the parametric PDE and hence, the trained model can be prompted with any mobility ratio to momentarily retrieve the correct saturation solution without retraining.

The proposed PI-DeepONets model is trained on 1000 non-convex flux functions with mobility ratios sampled from a uniform random distribution from 1 to 10. The model is then tested on 100 new flux functions from the same distribution. Our results show that the proposed model achieves a mean and standard deviation of relative L2 prediction error of 3.20×10 (-2)±5.80×10 (-3) when compared to high-fidelity numerical solutions. This demonstrates PI-DeepONets generalizability beyond the training data, making it a promising tool for accelerating the solution of PDEs in porous media.

S2207. Robust Optimization of Brine Extraction Well Placement Using Deep Learning and Fast-Marching Method

Hoonyoung Jeong (Seoul National University),
Hyunjee Yun (Seoul National University)
Room: R7 2023-08-09 15:00

In geological carbon storage projects, it is essential to manage reservoir pressures for long-term safety. The pressure build-up caused by CO2 injection may lead to serious safety issues; caprock damage, induced seismicity, and leakage. One of the practical solutions for mitigating reservoir pressure buildup is brine extraction which relieves the pressure buildup by extracting the brine. However, the brine extraction well should be optimized because the performance of brine extraction is affected by the well placement in heterogeneous reservoirs.

Optimization of brine extraction well placement is computationally expensive because many simulation runs are required to find optimal solutions. Robust optimization is to find optimal solutions for multiple reservoir models representing reservoir uncertainty. Thus, robust optimization of brine extraction well placement requires the higher computational cost.

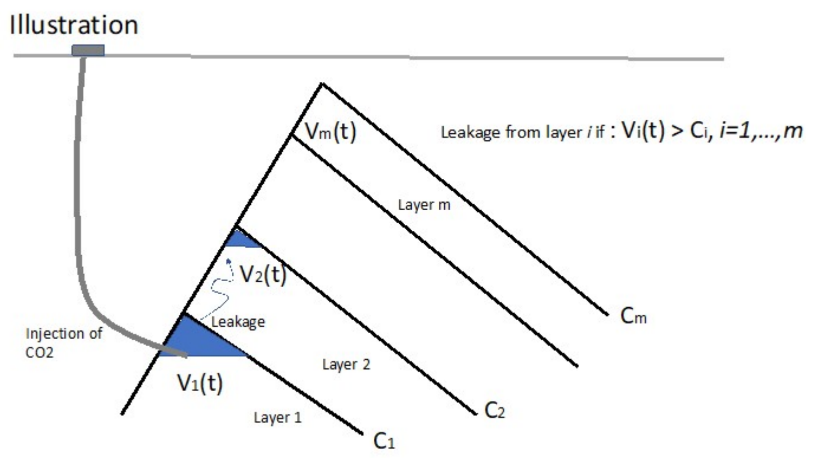
In this study, we propose a surrogate model that quickly evaluates the net present value (NPV) for brine extraction well locations using the fast-marching method and convolutional neural network. Our surrogate model was applied to a CO2 injection site in the Pohang basin, Korea. Compared to conventional optimization methods, our surrogate model saved 90% of the computational cost in finding the global optima of the brine extraction well location.

S2208. Value of seismic monitoring of CO2 storage in a multi-layer stratigraphic system

Geetartha Dutta (NTNU),
Jo Eidsvik (NTNU)
Room: R9 2023-08-10 14:20

A framework for value of information analysis of time-lapse seismic data in the context of CO2 storage monitoring is presented. It is considered that CO2 is injected in the bottommost layer (layer 1) in a stratigraphic model of m layers with uncertain petrophysical and topographic properties. The injected CO2 volume over time is known, but with uncertainty in the layer properties, the volume and pressure that builds up in layer 1 is uncertain and hence modelled stochastically. A linear model is assumed for the variation of volumetric pressure in layer 1 with time, the slope of which is considered random. Inspired by invasion percolation theory, when the volumetric pressure of layer 1 exceeds the capillary threshold of the cap rock for this layer, the residual volumes leak to the layer above. The capillary threshold is also random. The time of leakage from the first layer can then be calculated by finding the minimum time when the volumetric pressure exceeds the capillary threshold. Similarly for the other layers, the pressure builds up from the residual volumes leaking from the deeper layers if and when the volumetric pressure reaches the capillary threshold for the corresponding layer.

By including the volumes in different layers explicitly, the linkage to monitoring is facilitated, assuming that the differences in time-lapse seismic are indicative of the volumes that build up in various layers over time. Considering that time-lapse seismic data gives information about the CO2 volumes in various layers, a value of information analysis is then conducted to determine when to collect such data to maximize the value. The value of information analysis is based on statistical machine learning approaches working on Monte Carlo samples from the random slopes, volumes and capillary thresholds as well as associated synthetic seismic data.

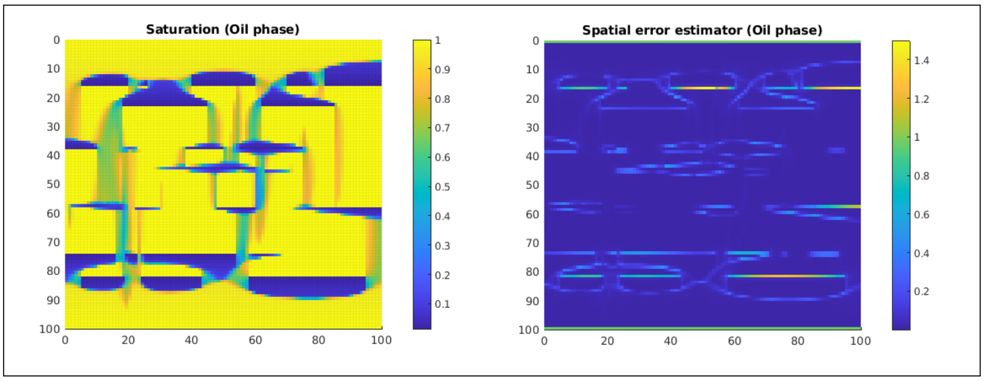


CO_2 is injected in the bottom layer. When the pressure in this layer exceeds the capillary threshold of the cap rock, the residual volumes start to leak into the layer above.

S2211. Adaptive Timestepping, Linearization, and A Posteriori Error Control for Multiphase Flow of Immiscible Fluids in Porous Media with Wells

Elyes Ahmed (SINTEF Digital),
Øystein Klemetsdal,
Xavier Raynaud,
Olav Møyner,
Halvor M. Nilsen
Room: R9 2023-08-10 14:00

This work focuses on the development of adaptive timesteps, stopping criteria, and error control strategies for reservoir simulations with fully implicit (FIM) solvers. Using a rigorous error control framework, an adaptive time selector combined with nonlinear stopping criteria is used to control nonlinear iterations as well as to balance accuracy and robustness for challenging nonlinear simulations.



S23 Automated Lithology

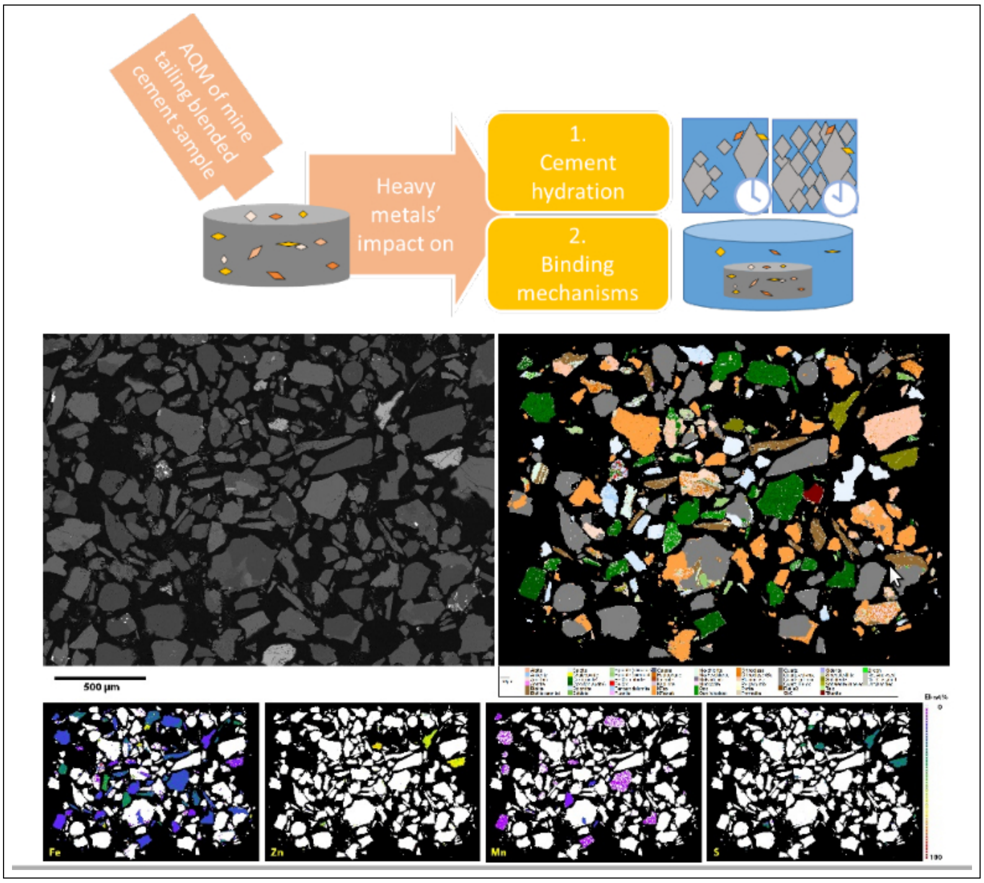
Kurt Aasly (NTNU),
Eric Pirard (University of Liège - GeMMe)

Automated characterization of minerals and rocks is of major importance for the future development of numerical geology. Exemplified with the mining industry, where this is known in core-scanning and automated mineralogy. The vast amount of data created from various analytical methods, creates sets of «big data» that, if properly analyzed, may be significant to improve the understanding of mineral deposits, their exploration, and their downstream processing. Recent technologies, based on both 2D and 3D imaging systems, in combination with the large number of data from geochemical analyses provides sets of big data that should be efficiently analyzed together using AI and ML methods. This session welcomes submissions that offer contributions on the approach for analyses of such big datasets and case studies.

S2301. Quantification of heavy metals in mine tailing blended cementitious samples

Anne Mette T. Bagger (Technical University of Denmark),
Stefanie Lode (Norwegian University of Science and Technology),
Wolfgang Kunther (Technical University of Denmark),
Pernille E. Jensen (Technical University of Denmark)
Room: R7 2023-08-10 10:00

Utilizing mine tailings as a partial replacement for cement offers the possibility to reduce waste deposition in the mining industry and CO2 emission associated with cement production. Mine tailings is the fine-grained waste fraction from ore processing which may still contain elevated amounts of metals including heavy metals. The impacts of heavy metals on cement hydration and binding mechanisms are yet unclear, and reportedly range from retarding to accelerating cement hydration. In addition, heavy metals may pose an environmental risk if they leach. To understand the metal-cement interaction, we propose applying automated quantitative mineralogy analysis (AQM) known in geology, which provides detailed information on modal mineralogy, mineral and element associations (incl., metals) as well as microtextures and grain properties. This methodology thus represents a new advanced microanalytical tool in the cement and concrete industry to investigate and quantify the heavy metals impact on cement hydration and their binding mechanisms.



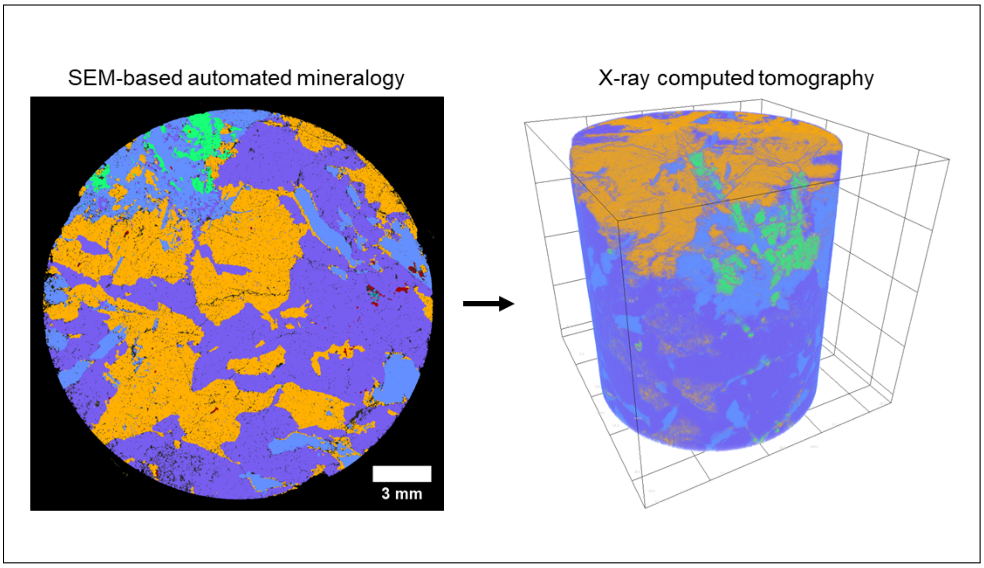
S2302. Combination of automated mineralogy and X-ray computed tomography for the 3D characterization of ore samples.

Florian Buyse (Ghent University),
Stijn Dewaele (Laboratory for Mineralogy and Petrology, Department of Geology, Ghent University),
Matthieu Boone (Radiation Physics Research Group, Department of Physics and Astronomy, Ghent Uni),
Veerle Cnudde (Pore-Scale Processes in Geomaterials Research group (PProGRess), Department of G)
Room: R7 2023-08-10 10:20

An accurate three-dimensional (3D) microscopic characterization of mineralized samples and associated data analysis are of major importance for the study of ore deposits. X-ray computed tomography (µCT) is a technique that allows for the non-destructive 3D imaging of geological samples. The principle of µCT is based on the reconstruction of the distribution of the X-ray attenuation coefficient. This parameter is based on two material properties (i.e., density and effective atomic number) which are impossible to disentangle in conventional µCT. Most applications with µCT in mineral characterization are, therefore, currently limited to the differentiation between the major phases: pores, low-density gangue minerals and high-density ore minerals. However, using µCT in combination with scanning electron microscopy-based automated mineralogy allowed us to differentiate between different minerals in 3D and quantify their structural and stationary textures (Buyse et al., 2023). Different mineral phases have been distinguished in a pegmatite sample and their orientation and spatial correlation were calculated by using statistical descriptors. In addition, Li-mica has been identified with µCT due to the chemical attribution of lithium on the resulting X-ray attenuation coefficient. The combination of µCT with automated mineralogy systems offers a helpful tool in the 3D visualization and quantification of the mineralogy of ore samples at the microscopic scale. The application of these techniques to whole core sections will open a new depth of information in genetic studies of ore deposits.

References

Buyse, F., Dewaele, S., Boone, M.N. & Cnudde, V., 2023. Combining Automated Mineralogy with X-ray Computed Tomography for Internal Characterization of Ore Samples at the Microscopic Scale. Natural Resources Research, 18 p.


SEM-based automated mineralogy
3 mm
X-ray computed tomography

S2303. Automated mineralogy using LIBS and deep learning

Simon Nachtergaele (University of Liège - GeMMe),
Christian Burlet (RBINS-GSB Royal Institute for Natural Sciences - Geological Survey of Belgium),
Renata Barros (RBINS-GSB Royal Institute for Natural Sciences - Geological Survey of Belgium),
Sophie Verheyden (RBINS-GSB Royal Institute for Natural Sciences - Geological Survey of Belgium),
Jean-Marc Baele (UMONS - Géologie fondamentale et appliquée),
Severine Papier (UMONS - Géologie fondamentale et appliquée),
Anca Croitor (KU Leuven - Biomedical MRI Unit, Institute for Artificial Intelligence),
Hassan Bouzahzah (University of Liège - GeMMe),
Eric Pirard (University of Liège - GeMMe)
Room: R7 2023-08-10 10:40

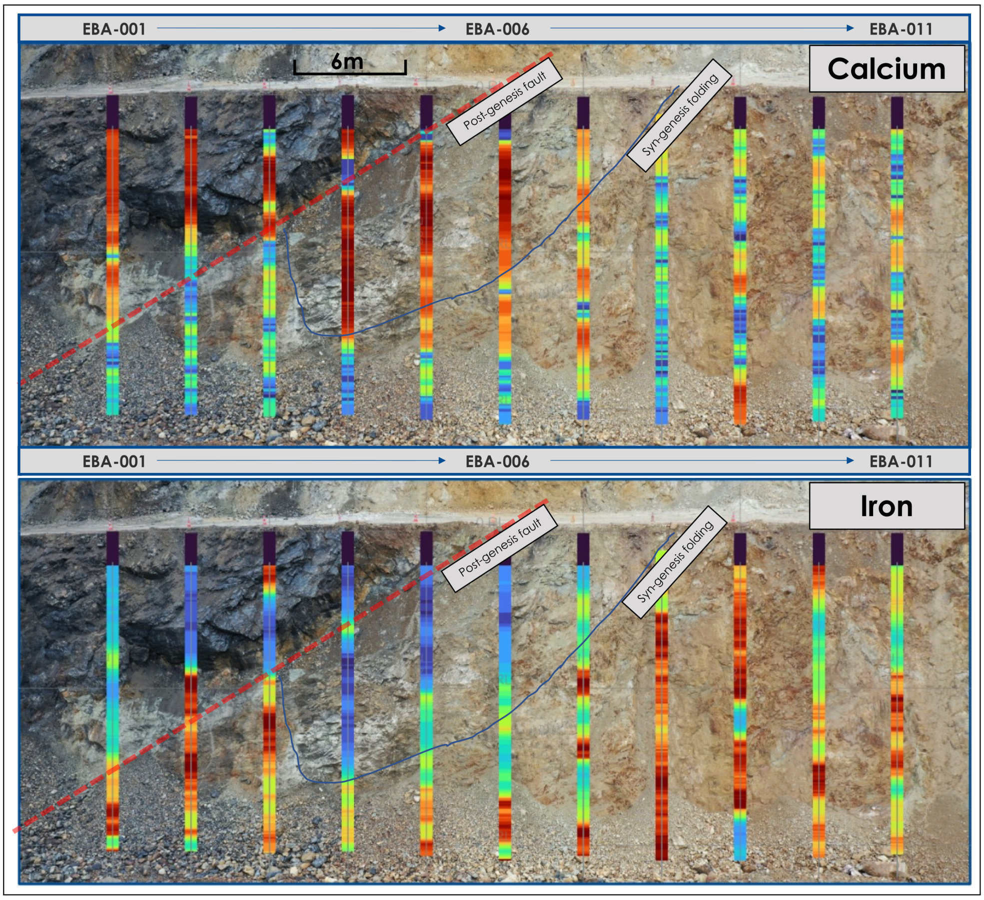
Laser Induced Breakdown Spectroscopy (LIBS) analysis was done on mineralized rock cores from the Pb-Zn ores of Plombières (Belgium) in the scope of the LIBS-SCReeN project (https://sites.google.com/view/libs-screen/). Mineral classification using modern data science techniques such as neural networks showed promising results. The neural networks were already pre-trained on LIBS data from an annotated mineral LIBS database. The classification results and accuracy will be evaluated during this presentation. Also, the LIBS instrumentation of the Belgian LIBS cluster will be explained in detail.

This research highlights the highly inter-disciplinarity character of LIBS automated mineralogy. In this subject, geology is combined with data science techniques which are necessary to process the "big data" originating from spectroscopy techniques. Rapid data handling using neural networks (AI) has become essential to increase the productivity of a potential automated mineral core logging system. Therefore, different neural network architectures will be compared and the most optimal data processing pipeline will be proposed.

S2304. Improved mineralogy from multivariate analysis of data from the Orelog spectrometric logging tool measured at the Erzberg iron ore deposits (Austria)

Manuel Queisser (UIT GmbH),
Matthew Tudor (UIT),
Thomas Heinig (UIT),
Alexander Domula (UIT),
Jens Schubert (UIT),
Horst Märten (UIT),
Alfred Stadtschnitzer (VA Erzberg)
Room: R7 2023-08-10 11:00

Erzberg (Styria, Austria) hosts the largest known siderite deposit in the world. Open pit mining is applied to extract iron ore. A novel, compact (76 mm diameter, 3.1 m length) geophysical logging tool of type Orelog (UIT Dresden, Germany) with a pulsed, miniaturized Deuterium-Tritium (DT) neutron generator was used to measure elemental abundances in the Erzberg ore formations. Orelog uses time-resolved prompt gamma-ray spectrometry (distinguishing between spectra from inelastic fast-neutron scattering, thermal neutron capture and other spectral components) as well as time-resolved neutron detection at various detector locations (source, near, far) and in several energy groups (fast, thermal, optionally epithermal). The tool measures the elemental composition and geophysical characteristics in real-time. Eleven 23 m deep DC boreholes, drilled along a bench face of the Erzberg mine, were logged to determine the mass fractions of 7 main elements (Fe, Mn, Ca, Mg, Si, K, S). Multivariate analysis of the data, including principal component analysis, depth resolved cross plots and cumulative probability plots was used to map and quantify three main sections of the bench face, each with a distinctive mineralogy, including an intermediate section. The top 10 m of this intermediate section are dominated by calcite and clay minerals, followed by a transition zone with ankerite/siderite and a bottom section dominated by siderite. In summary, the tool data provided a more detailed picture of the mineralogy of the bench face otherwise hidden to the eye and moreover a quasi-real-time grade control in mining, minimizing waste material and, thus, increasing sustainability. Its versatility makes Orelog and the data it provides an excellent tool for explorational drilling, providing swift quantitative insights. Orelog's geochemical data is furthermore eminent in constraining ore-deposit models, including multi-parameter models, such as geophysical inverse models combined with geochemical/mineralogical models, e.g., reactive transport models of mineral precipitation in carbon sequestration, where knowledge of the mass fractions of mineralizing cations (including Na^+, Ca^{2+}, Mg^{2+}) in the rock matrix is crucial.


EBA-001
EBA-006
EBA-011
6m
Calcium
Post-genesis fault
Syn-genesis folding
EBA-001
EBA-006
EBA-011
Iron
Post-genesis fault
Syn-genesis folding

S2305. Lithofacies Classification for In-situ Leaching Uranium Mining Deposits using Machine Learning Algorithms

Timur Merembayev (Nazarbayev Univesity),
Yerlan Amanbek (Nazarbayev University)
Room: R7 2023-08-10 11:20

Machine learning algorithm is an effective tool in reducing the uncertainty of physical properties from the subsurface. The lithofacies classification is essential for several applications in geoscience issues including the uranium production using in-situ leaching extraction. The automatic classification has been faced with several specific geological challenges for accurate detection of borders of lithofacies due to a small thickness of litho type in some reservoirs. In this talk, we present the model of lithofacies classification using machine learning algorithms from the various well log data of the uranium mining deposit in the Kazakhstan reservoir. We will compare results in the detection of lithology boundaries using machine learning algorithms for the well log data. The evaluation of the model score is conducted by metrics such as f1 and penalty matrix. In addition, a validation of the result is performed on core images by applying the convolutional neural network method for classification lithofacies.

S24 Real Options in Geosciences

Verena Hagspiel (NTNU),
Milan Stanko (NTNU),
Luis Martinez Tipe (MCEISA)

Energy and mining sectors play a critical role in the modern economy. These are capital intensive sectors tied to products that are commonly traded on financial exchange. Another characteristic is that projects in these sectors are prone to large private risk related to resources as well as available technology. Therefore, advanced valuation techniques and forecasting can significantly improve economic aspects associated with the development and operation of projects. Real options analysis allows to quantify the value of managerial flexibility inherent in potential projects. Managerial flexibility stems from the decision maker's ability to react on uncertainty as it unfolds over time. Project development related to mining or energy can in many cases be split into a sequential problem with multiple decisions over time. Using the knowledge gained between stages, the management can optimize future decisions. This allows to some extend mitigate downside risk and exploit the upside potential of uncertainty. Real options valuation provides us with a tool to quantify the value of such flexibility and identify the most profitable decision strategies.

In this session we aim to bring together experts as well as newcomers to the topic that are interested to learn about practical applications of real option valuation to decision problems in energy and mining sectors.

S2401. Valuing an open pit mine as a portfolio of European call options: assessing the effext of metal price and metal grade uncertainties in an open pit gold mine

Luis Martinez Tipe (MCEISA)
Room: R8 2023-08-09 10:00

When starting a mine operation, owners and stakeholders of the project need to know the value of the mining project and the cash flow that will generate over its operating life. These are the fundamental bases for making final irreversible decisions about going ahead with the project investment; the main consequence of not including uncertainty in mine project evaluation is that it misleads decision-makers and investors to a static description of the economic and technical risk of the mine project.

As such, the valuation of the mine is seen as the valuation of a multi-period portfolio of European call options; i.e., at each production period, the decision to either mine and process a block from the mine, or to mine and send it to the waste dump, are seen as a European option.

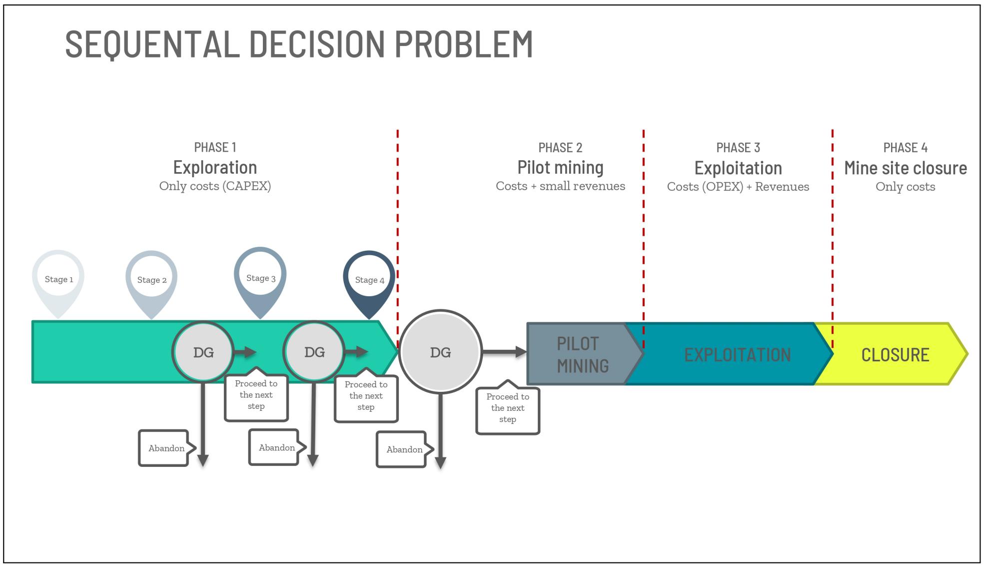
In this work, geostatistical simulations in conjunction with metal price uncertainty are used in a practical framework to show that the value of an open pit mine depends on the level of geological uncertainty and price volatility. Furthermore, this work also focused in confirming the theoretical results obtained by Henry, et.al. (2005), who stated that: i) metal grade uncertainty may introduce asymmetries in the block value greater than metal price uncertainty; and ii) that the benefits of optimising an open pit mine under geological uncertainty is an order of magnitude below the benefits for resolving the geological uncertainty. The contribution of this work rescinds in that the mine optimisation process is done including time discounting and slope angles constraints.

To achieve these objectives, 100 conditional simulations in conjunction with simulated prices are used to value an open pit gold mine project. Here, gold prices are modelled as a Geometric Brownian Motion process. To capture the intrinsic optimised mine value over time (visualised here as a portfolio of European options), the Direct Block Schedule technique, implemented in the Mining Math software, is used to generate an optimum long term production schedule that includes the strategic mine plan and design.

S2402. Practical application of real options valuation: a case for deep-sea mining activities on the Norwegian continental shelf

Farida Mustafina (NTNU),
Verena Hagspiel (NTNU)
Room: R8 2023-08-09 10:20

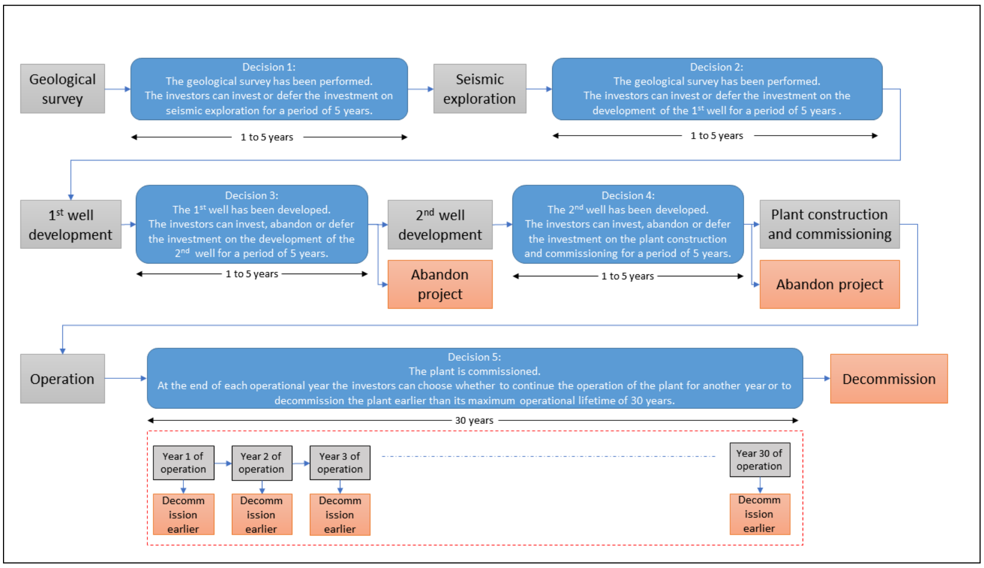
One of the essential tasks of the current generation is creating a sustainable global economy for future generations. Given the increasing demand for minerals and decreasing ore grades in land-based deposits, deep-sea mineral deposits are considered a potential source of critical minerals. The question is whether deep-sea mining can contribute to the new economy in a socially acceptable, economically viable, and ecologically sustainable manner. Answering this question is the main task of the "The Deep Dilemmas: Deep Sea Mining for the new Deep Transition? (TripleDeep)" project at the Norwegian University of Science and Technology (NTNU). As a part of this project, we aim to consider the economic issues of deep-sea mining activities with a particular interest in decision-making under uncertainty. Exploration of deep-sea mineral deposits will require substantial investment in technology and infrastructure. From the corporate perspective, without proven technology and preexisting infrastructures, investors may prefer to wait due to the vast costs, large uncertainty about future revenues, and future regulatory framework. It is crucial for the regulatory side to understand how uncertainties and risks affect the corporate side. For that purpose, we will apply a real options approach, which allows us to analyze the effect of risk and uncertainties on investment decision making. Here we will focus on some of the main uncertainties regarding deposit characteristics and regulatory decisions arising during the exploration phase of projects. We will consider possible types of regulations and incentive mechanisms and study their effect on investment behavior with a specific focus on the Norwegian continental shelf. We aim to provide important insight into the decision making regarding the future of deep-sea mining activities.



S2403. The value of flexibility in geothermal energy investments: an environmental impact driven real options analysis

Spyridon Gkousis (University of Antwerp),
Kris Welkenhuysen (Geological Survey of Belgium),
Tine Compernolle (Geological Survey of Belgium; University of Antwerp)
Room: R8 2023-08-09 10:40

Geothermal energy is considered a renewable energy source with a low carbon footprint that can provide a reliable alternative to fossil-fired energy. In this context, the interest for exploiting medium enthalpy deep geothermal resources for heating purposes in Belgium has been increasing. Deep geothermal energy development occurs in separate steps (e.g. surface explorations, first drillings) and after each step the developers can pick over a range of options for how to proceed with the development (continue, abandon, defer). Traditional economic and environmental evaluation methods cannot capture the additional economic and environmental benefit this flexibility brings to the project development. In this context, this study combines the real options theory with techno-economic and life cycle analysis to investigate the economic feasibility and environmental desirability of deep geothermal heating in Northern Belgium. A traditional real options solution, that calculates the decisions with the aim to maximize the economic value acquired, is compared to an impact driven real options solution that aims to minimize the environmental impact of the plant. Results show that depending on the underlying geological and market conditions the results of the two solutions can vary marginally or more. Regardless, the economic and the environmental performance of the plant are improved, compared to when the developer's flexibility is omitted, in both real options solutions.



S2404. A Monte Carlo and decision tree approach to Real Options Analysis for CO2 capture and storage

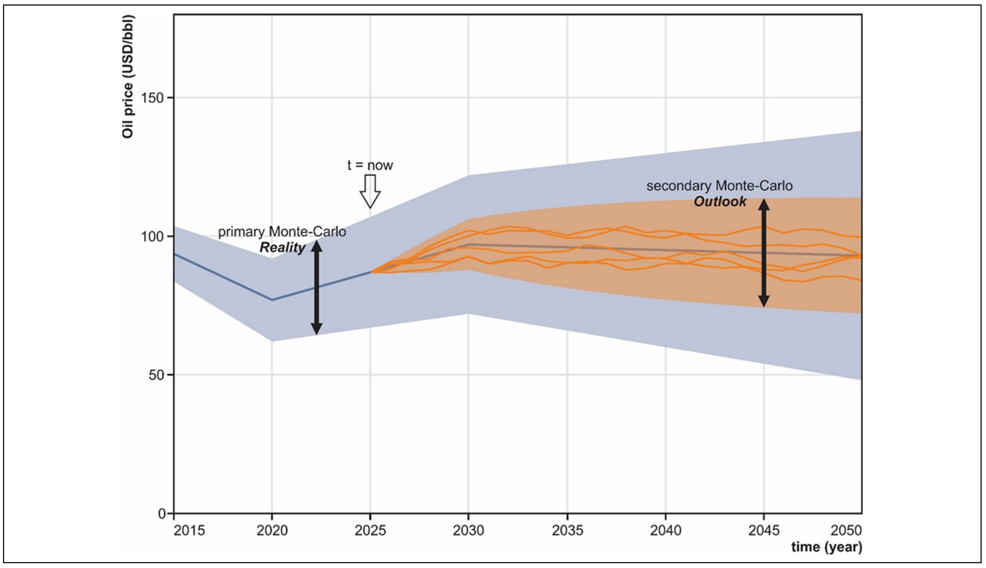
Kris Welkenhuysen,
Alejandra Tovar (Geological Survey of Belgium - Royal Belgian Institute of Natural Sciences, Jenn),
Tine Compernolle (Department of Engineering Management, University of Antwerp, Prinsstraat 13, Ant),
Kris Piessens (Geological Survey of Belgium, Royal Belgian Institute of Natural Sciences, Jenne)
Room: R8 2023-08-09 11:00

The development of large-scale infrastructure for the capture, transport and storage of CO2 (CCS) as a climate change mitigation measure requires large investments and careful planning and coordination between the stakeholders involved. The CCS chain of technologies is gradually being developed and only sparsely being deployed worldwide, while climate ambitions require a fast ramp-up of implementation. Many factors introduce uncertainty for decision makers, such as energy and CO2 market prices, the development of capture technologies, and especially uncertainty on geological parameters which will dictate how suitable a subsurface volume is to be developed as a CO2 storage reservoir. This makes realistic forecasts on the deployment of CCS technology and its true potential challenging.

The concepts of Real Options Analysis (ROA) allow for introducing flexibility in decision making, while considering uncertainty. An analytical ROA solution to simulate CCS investment decisions was found to be restrictive in the number of uncertain parameters that have a significant impact, and in the complexity of successive project decisions. Instead, a numerical approach is developed based around a decision tree which defines the timings and potential investments or changes in such a project. Two levels of Monte Carlo (MC) calculations are applied. In the primary MC, the stochastic parameters are attributed a value according to a probability density function, for each iteration. These values serve as "reality" in the simulation process. A secondary MC provides a level of uncertain outlooks or limited foresight on which investment decisions are based. A single primary MC results in a succession of project decisions that are a best assessment considering uncertain outlooks. A full simulation delivers development pathways with probabilities attached.

A particular benefit of this method is that geological parameters and their resulting reservoir behavior are treated at the same level as the techno-economic ones. This also means that uncertainties at reservoir level are carried throughout the simulation process and are represented in the results.

Our approach creates imperfect decisions to better approach realistic decision making. Because of the numerical nature of the calculation, the number of uncertain parameters and flexibility options is only limited by calculation time.


Oil price (USD/bbl)
150
100
50
0
t = now
primary Monte-Carlo
Reality
secondary Monte-Carlo
Outlook
2015
2020
2025
2030
2035
2040
2045
2050
time (year)

S2405. Investment planning under uncertainty: a multi-objective real options analysis of offshore platform electrification

Olga Noshchenko (NTNU),
Verena Hagspiel (NTNU)
Room: R8 2023-08-09 11:20

The petroleum industry faces a growing environmental pressure that raises a set of complex questions about the ways of fulfilling environmental requirements in an economically viable manner. There is an increasing demand to improve the platform operation schemes that are responsible for the largest share of CO2 emissions produced by powering gas turbines. In addition to environmental challenges, due to ongoing global market fluctuations offshore field projects are also challenged by market uncertainty. Hence, more advanced economic valuation methods such as real options and decision analysis are required to correctly treat uncertainty and assess the value of investment opportunities and inherent managerial flexibility.

To demonstrate application of the advanced economic valuation methods, a synthetic case study considering a construction of offshore platform is examined. In the research, we propose to optimize planning of investments in "green field" offshore petroleum project considering market uncertainties and environmentally friendly electrification alternatives. Prices of oil, gas, and electricity as well as EU CO2 allowances are selected as the most relevant factors of uncertainty that affect the project's value and investment decision. Connection to power from shore grid and offshore wind farms represent two electrification alternatives. To perform the optimization, we develop a methodological framework for decision-making under market uncertainty based on real option modelling approach. We apply Least-Squares Monte Carlo (LSM) algorithm to precisely estimate the value of waiting and assess the impact of uncertainty factors. The modified two-factor short-term/long-term (STLT) stochastic model using future price was applied to forecast oil and gas prices. The development of EU CO2 allowances rate is described by GBM process, while STLT price model proposed by Schwarz and Smith is applied to evaluate electricity prices.

The combination of the modified price models and LSM dynamic programming allows to receive more precise estimation of the optimal timing to exercise the option to invest in the project and select electrification alternative. The results of the optimization provide price values from which it is economically viable to exercise the investment options in environmentally friendly solutions.

S2406. Decision-support methods for early-phase design of offshore oil and gas fields using model-based optimization and uncertainty quantification

Milan Stanko (NTNU)
Room: R8 2023-08-09 11:40

In early phases of offshore oil and gas field development, it is often necessary to take decisions about the main features of the field such as production strategy, drilling and production schedules, processing capacities, among others with limited and uncertain information. This presentation will provide a summary of the work performed by the research group of the presenter on using proxy-modeling, numerical optimization and probabilistic analysis to provide decision support to field planners during early stages of offshore field development.

Some issues that will be presented are: development of proxy models to capture the field production performance, efficient non-linear and piece-wise linear optimization formulations, multi-objective optimization including environmental factors such as carbon emissions and footprint, effect of uncertainties in the subsurface, cost and price, sampling versus stochastic optimization, inclusion of subsea processing, abandonment time, recovery mechanism, price shock testing. The methods have been tested in literature cases, synthetic cases and real cases.

The methods presented could significantly improve early phase offshore oil and gas field development workflows as they run in short time and allow to take into account several uncertainties. Moreover, there are some practical learnings that can be incorporated into field design guidelines.

S25 Multiple Point Statistics

Jacob Skauvold (Norsk Regnesentral),
Håkon Tjelmeland (NTNU)

Multiple Point Statistics (MPS) has developed tremendously during the last 20 years - the main idea being to move away from the traditional (Gaussian) two-point dependency models. Based on extensive training data, MPS can provide realistic discrete (or continuous) realizations of geomodels for facies or other attributes. Over the years more complex MPS models have been developed, and one can say that this has sparked interests in using recent machine-learning approaches such as Generative Adversarial Networks which are currently quite popular in the field.

S2501. 2D Stochastic Structural Geomodeling with Deep Generative Adversarial Networks

Charlie Garayt (Geovariances / Mines Paris PSL),
Nicolas DESASSIS (Mines Paris - PSL),
Samy BLUSSEAU (Mines Paris - PSL),
Jean LANGANAY (Geovariances),
Pierre-Marie GIBERT (Geovariances),
Thomas ROMARY (Mines Paris - PSL)
Room: R5 2023-08-10 10:00

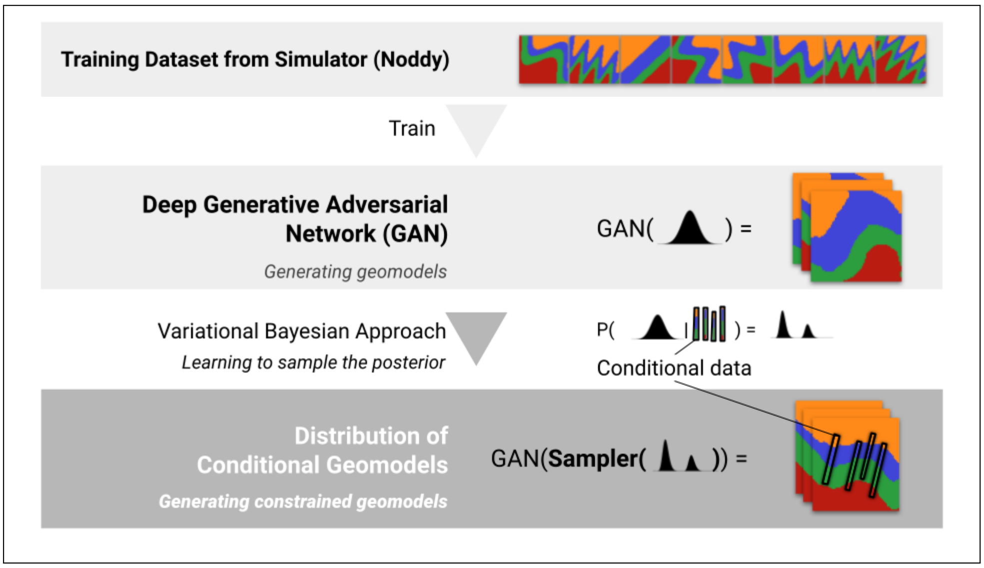
This work aims to simulate 2D structural geological models, or geomodels, that respect given knowledge and data. By definition, geomodeling is an ill-posed problem due to the limited quantity and quality of available data. Current geomodeling methods struggle to both characterize uncertainties and produce realistic geomodels.

To achieve this goal, a deep generative adversarial network (GAN) has been implemented. GANs, which are usually used in image generation, need a large training dataset. For instance the ImageNet dataset contains more than 1.5 million images. In geomodeling, such a large real dataset does not exist. Fortunately, geological structures are a consequence of physical and chemical processes, so creating a synthetic dataset is feasible from the simulation of these processes.

The training dataset is created from Noddy, which can be viewed as an object-based simulator. The use of advanced GANs like Least-Square GAN (LSGAN) and Wasserstein GAN allows training a deep neural network called the Generator. The Generator defines an implicit distribution of geological models. This is a function that transforms a random vector into a geomodel similar to the training dataset.

However, the Generator produces unconstrained geomodels. A variational Bayesian approach is used in order to train a Sampler, which enables the generation of geomodels that fulfill constraints, or conditioned geomodels. Thanks to the versatility of the variational Bayes approach, constraints can be of different types and quality, for instance rock type, rock orientation or geophysical data. The goal of the Sampler is to find the posterior distribution where the Generator produces the desired conditioned geomodels. Using a variational Bayesian approach makes it possible to take into account different types of data, with their own quantity and quality without re-training the Generator.

Finally, the combination of the Sampler and the GAN allows the generation of conditioned geomodels. In addition, this approach enables handling uncertainties and to perform computation, since the resulting generator produces an implicit distribution of conditioned geomodels.



S2503. Quantifying verbatim copy in MPS simulations using location Maps

Mathieu Gravey (Austrian Academy of Sciences)
Room: R5 2023-08-10 10:20

Multiple point statistics (MPS) simulations are a powerful tool for modeling subsurface reservoirs and predicting their properties. However, one of the challenges of MPS simulations is the potential for verbatim copy, where a complete section of the training image is copied into the simulation. This can artificially reduce variability and provide overconfident results. In order to address this issue, a new approach has been proposed to quantify verbatim copy in MPS simulations, specifically in Direct Sampling (DS) and QuickSampling (QS). The proposed approach utilizes the index/location map generated as a side product of DS and QS simulations to inform you of the origin of each simulated pixel. By studying the relationship between the original position of pixels and positions of simulated pixels, it is possible to extract some information about verbatim copy. The advantage of this approach is that it doesn't rely on the value of pixels and therefore extremely robust against false positive. This approach can help to improve the accuracy and reliability of studies relying on MPS simulations.

S2504. Reconstructing 3D Overthrust Model by Multiple-point Statistics and Fully Connected Networks Constrained with Geological Knowledge

Hou Weisheng (Sun Yat-sen University),
Yanhua Li (Sun Yat-sen Universit),
Shuwan Ye (School of Earth Sciences and Engineering, Sun Yat-sen University),
Songhua Yang (School of Earth Sciences and Engineering, Sun Yat-sen University)
Room: R5 2023-08-10 10:40

An accurate 3D geological model is of great significance for many applications, such as understanding geological processes and revealing geodynamic mechanisms. However, reconstructing complex fault systems is still a bottleneck problem in 3D geological modeling. In multiple-point statistics (MPS) simulation, the relations in known data (termed training images, TI) are characterized by patterns instead of an implicit or explicit function. The patterns extracted from TI are reorganized with a stochastic process, to reproduce a new result. The MPS has been successfully used in many applications. However, the constraints of geological knowledge from geologists are difficult to add to the MPS simulation process.

In this study, a novel MPS method combined with a fully connected network (FCN) is proposed, which geological knowledge constraints including stratigraphic sequence, fault activity period, and stratum thickness. The joint objective function of the presented method is composed of the loss function of the FCN and the conventional MPS objective function constrained by the geological semantics of geological objects. The global features of each geological object are generated from the trained FCN of which the parameter is the subsurface elevation. A 3D model at the coarsest scale obtained with the trained FCN is used as the initial model for the following iterative MPS process. The stratigraphic sequence and geological object relationship are checked before the initial model is output. The thickness of each geological object is used as a soft constraint during the simulation process. A concrete example of reconstructing the overthrust model is given, in which 10 orthogonal cross-sections are extracted from the real model as the modeling data source. The realizations illustrated that the geometry and spatial distributions of strata intersected by three groups of faults, and the stratigraphic sequence is kept well. The comparison of virtual boreholes and the real model shows that the accuracy of the geological object reaches 75%. The detail of the complex faults and stratigraphic system can be obtained accurately by the presented method.

This study was supported by the National Natural Science Foundation of China (NSFC) Program (41972302, 41772345).

S26 Compositional Data Analysis

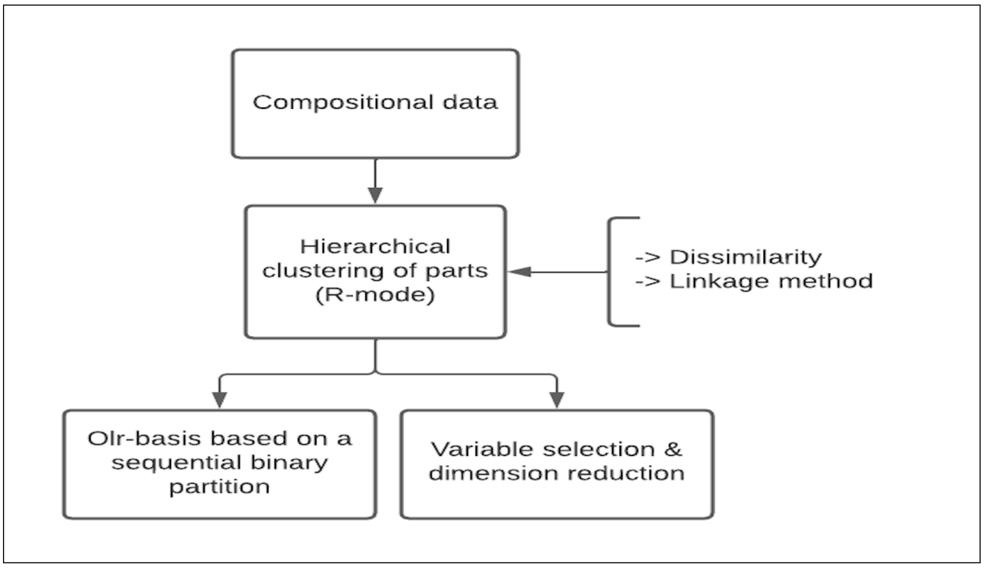
Juan Jose Egozcue (Technical University of Catalonia),
Vera Pawlowsky-Glahn (University of Girona)

Compositional Data Analysis (CoDA) can be applied advantageously in many fields. However, CoDA still needs methodological improvement. This session is focused on such applications in geosciences and includes methodological contributions.

S2601. Exploring Hierarchical Clustering of Compositional Variables

Josep Antoni Martin-Fernandez (University of Girona),
Valentino Di Donato (UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II),
Vera Pawlowsky-Glahn (University of Girona),
Juan Jose Egozcue (Technical University of Catalonia)
Room: R7 2023-08-10 13:40

Cluster of variables (i.e., R-mode clustering) can be useful for dimension reduction and variable selection. In addition, in compositional data analysis, hierarchical clustering (HC) methods can be used for creating an orthonormal log-ratio basis easily interpretable when the number of variables is not large. R-mode agglomerative HC requires the calculation of dissimilarities (or similarities) between variables. For measuring the difference (or association) between variables the most popular options are based on distances, angles, and indices of correlation. In addition, the clustering method is defined by the way in which two clusters (which may be single a single part) are combined to form a larger cluster. The classical options are single-linkage (or nearest-neighbour), complete-linkage (or farthest-neighbour), a compromise between these two (average-linkage methods), and the Ward method. In this work, we explore both elements, types of dissimilarities and linkage options, when the compositional variables are expressed using log-ratio coordinates. The identification of groups of species characterized by the similar response is of considerable importance in ecological or palaeoecological studies. A dataset of planktonic foraminiferal assemblages retrieved from a Mediterranean Sea core expressed in terms of log-ratio coordinates is shown as an illustration.



S2602. Geographically weighted regression for compositional data (GWR-CoDA): An application to a geochemical data

Takahiro Yoshida (The University of Tokyo),
Daisuke Murakami (The Institute of Statistical Mathematics),
Hajime Seya (Kobe University),
Narumasa Tsutsumida (Saitama University),
Tomoki Nakaya (Tohoku University)
Room: R7 2023-08-10 14:00

Geographically weighted regression (GWR) is a widely used spatial data analysis technique across various fields. Additionally, the extension for non-Gaussian distributed data has been progressing. However, studies on the extension for compositional data are limited. Spatial regression model developments for compositional data are crucial topics in compositional data analysis (CoDA) literature. Geostatistical compositional models, such as the compositional kriging model, are popular approaches because CoDA has been historically developed in geosciences in which a continuous spatial process can be assumed. Other study approaches employ conditional autoregressive models or simultaneous autoregressive (spatial econometric) models. Although spatial autocorrelated errors are considered in these models, models for compositional data with spatial heterogeneity or spatially varying relationships remain limited. The objective of this study is to build a GWR model for compositional data to consider both spatial heterogeneity and the constant-sum constraint. We accommodated the GWR model and log-ratio techniques of CoDA and then formulated the GWR model in the simplex space. We applied this model to analyze an actual data set from a geochemical mapping project. The interpretational usefulness of the results of spatially varying compositional semi-elasticities was empirically verified.

S2603. First-order compositional differential equations in geochemistry: a compositional linear model for world oil characterization

Eusebi Jarauta-Bragulat (UPC),
Carme Hervada-Sala (Universitat Politecnica de Catalunya (UPC)),
Jose Gibergans-Baguena (Universitat Politecnica de Catalunya (UPC))
Room: R7 2023-08-10 14:20

First-order ordinary differential equations are applied in process of characterization models in geochemistry. In particular, in models applied in oil, the growth curves that fit the data correspond to ordinary differential equations. Usually these models are applied to annual oil production data and in some models, hypotheses about oil reserves are mentioned.

World oil can be characterized compositionally as a system of three parts: cumulative production, known reserves, and a third unknown component, such that the sum of the three parts is equal to the total amount of oil on the planet. Total world oil is, by definition, constant and bounded, and quantitative hypotheses can be made about it. This approach allows the application of Compositional Data Analysis concepts and tools.

This paper shows the analysis of this compositional system and as a result, a linear model to characterize the world oil system, from the ILR coordinates of the three parts considered in the system. Thus, the compositional derivative of the system is obtained, which turns out to be constant. In addition, the equation of the compositional line that describes this system is obtained, which is also shown in the three-part simplex.

Keywords: world oil system, simplicial derivative, compositional linear model, isometric logratio coordinates.

References

Buccianti, A. and E. Grunsky (2014). Compositional Data Analysis in geochemistry: are we sure to see what really occurs during natural processes? Journal of Geochemical Exploration, Volume 141, June 2014, pages 1-5.

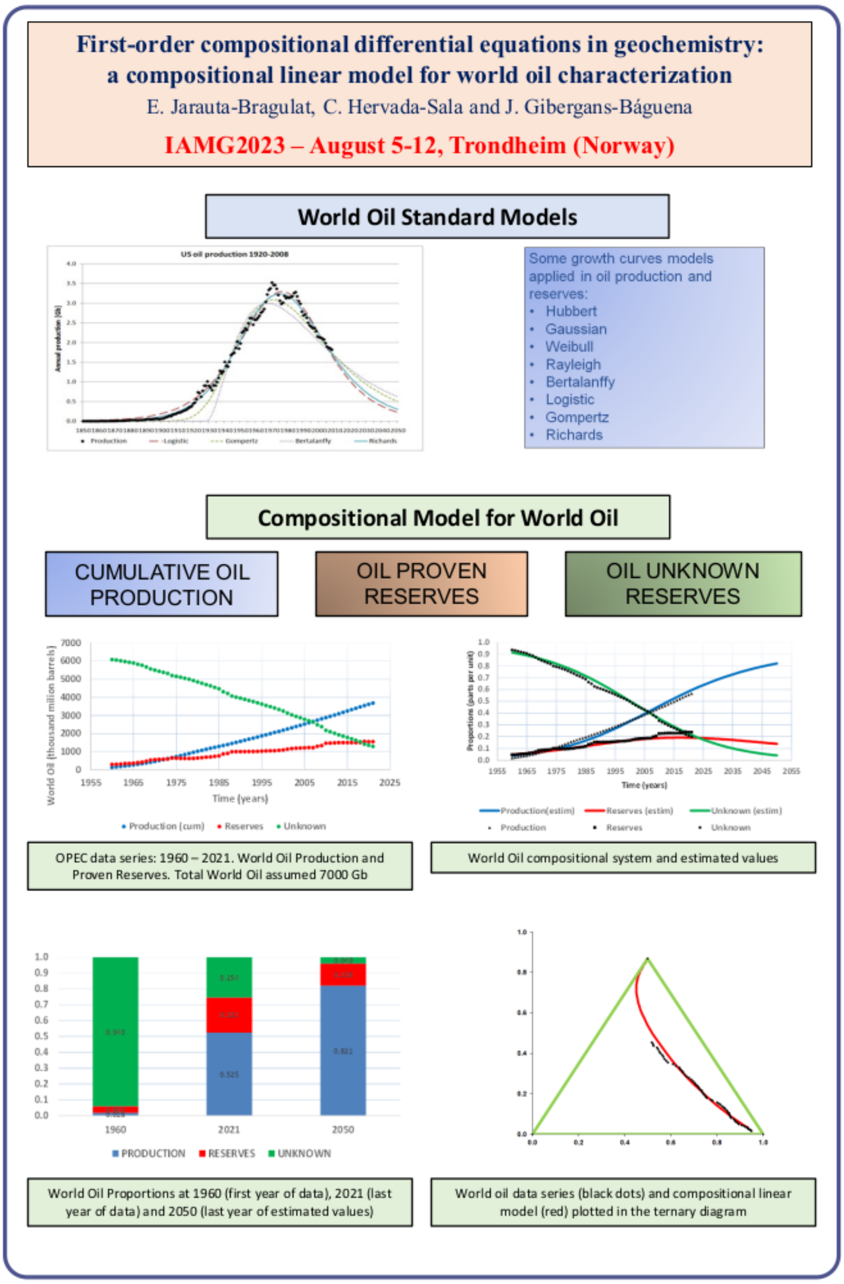
Egozcue J.J., E. Jarauta-Bragulat (2010). An approach to growth curves analysis from a simplicial point of view. Proceedings of IAMG2010, Budapest.

Egozcue J.J., E. Jarauta-Bragulat (2014). Differential models for evolutionary compositions. Math. Geosciences. Volume 46: 381-410.

Höök, M., Junchen, L., Noriaki, O. & Snowden, S. (2011). Descriptive and predictive growth curves in energy system analysis. Natural Resources Research, Vol. 20, Issue 2: 103-116.

Pawlowsky-Glahn, V., J.J. Egozcue and R. Tolosana-Delgado (2015). Modeling and Analysis of Compositional Data. Wiley.

Wang, J. and L. Feng. (2016). Curve-fitting models for fossil fuel production forecasting: Key influence factors. Journal of Natural Gas Science and Engineering, Vol. 32: 138-149.


First-order compositional differential equations in geochemistry:
a compositional linear model for world oil characterization
E. Jarauta-Bragulat, C. Hervada-Sala and J. Gibergans-Báguena
IAMG2023 – August 5-12, Trondheim (Norway)
World Oil Standard Models
US oil production 1920-2008
Some growth curves models applied in oil production and reserves:
• Hubbert
• Gaussian
• Weibull
• Rayleigh
• Bertalanffy
• Logistic
• Gompertz
• Richards
Compositional Model for World Oil
CUMULATIVE OIL PRODUCTION
OIL PROVEN RESERVES
OIL UNKNOWN RESERVES
World Oil (thousand million barrels)
Time (years)
Production (cum)
Reserves
Unknown
OPEC data series: 1960 – 2021. World Oil Production and Proven Reserves. Total World Oil assumed 7000 Gb
Time (years)
Production(estim)
Reserves (estim)
Unknown (estim)
Production
Reserves
Unknown
World Oil compositional system and estimated values
PRODUCTION
RESERVES
UNKNOWN
World Oil Proportions at 1960 (first year of data), 2021 (last year of data) and 2050 (last year of estimated values)
World oil data series (black dots) and compositional linear model (red) plotted in the ternary diagram

S2604. Principal component analysis for distributional data with application to particle size distributions

Ivana Pavlů (Palacký University Olomouc),
K. Gerald van den Boogaart (HZDR, HIF),
Raimon Tolosana-Delgado (Helmholtz Institute Freiberg),
Jitka Machalová (Department of Mathematical Analysis and Application of Mathematics, Palacký Univ),
Karel Hron (Palacky University, Olomouc)
Room: R7 2023-08-10 14:40

Particle or grain size distributions often play an important role in understanding processes in the geosciences. Functional data analysis allows applying multivariate methods like principal components and discriminant analysis directly to such distributions. These are however often observed in the form of samples, and thus with a sampling error, i.e. each data point is a distribution, but one where the sampling error is present. This additional sampling error changes the properties of the multivariate variance and thus the value, number and direction of the principle components. The result of the principal component analysis becomes an artefact of the sampling error and can negatively affect the following data analysis.

Our contribution presents how to compute this sampling error and how to confront it in the context of principal component analysis. We demonstrate the effect of the sampling error and the effectiveness of the correction with a simulated dataset. We show how the interpretability and reproducibility of the principal components improve and become independent of the selection of the basis. We also demonstrate how the correction improves interpretability of the results on a grain size distribution dataset from river sediments.

S27 AI and Geostatistics

Jo Eidsvik (NTNU)

There is currently tremendous progress in the field of artificial intelligence (AI) providing optimized decisions, plans or learning systems via sophisticated computer programs. In this session authors study how the development in AI links to geostatistical models and methods. Some presentations combine AI and geostatistics for various geoscience applications, while other outline AI methods that can potentially overcome challenges with more traditional geostatistical approaches.

S2701. AI and Geostatistics: Similarities, Differences and Opportunities

Dionysios Christopoulos (Technical University of Crete)
Room: R5 2023-08-10 11:00

This presentation gives a conceptual, personal perspective—in the spirit of a biased random walk—on the relationship between AI and geostatistics. We first illustrate the close connection between geostatistics and the AI method of Gaussian process regression (GPR), thus showing that kriging is undead and thriving in the GPR framework. For quantitative data that follow skewed probability distributions or contain heteroskedastic noise, we discuss how Gaussian anamorphosis and the equivalent AI procedure known as data warping can benefit from new advances such as modified exponential and logarithmic transforms with flexible tails and non-parametric warping. AI involves a broader array of methods than geostatistics. While the latter is mostly based on function approximation models, AI also uses algorithms that employ general logical rules (e.g., methods based on decision trees). Such methods can be quite accurate in classification and regression tasks and allow incorporating various types of datasets at the expense of reduced interpretability. We will briefly compare results for time series forecasting using GPR, tree-based methods and LSTM neural networks, and discuss advantages and disadvantages.

The main engine of geostatistical methods, as well as of GPR, is the covariance kernel which incorporates information regarding spatial and temporal correlations. There are two main problems with covariance-based methods: firstly, large, dense covariance matrices need to be inverted for prediction and simulation of big datasets, a task which is computationally costly or even prohibitive; secondly, it is mathematically challenging (or impossible) to test whether specific kernels provide a permissible covariance when non-Euclidean distance measures or multivariate datasets are used. Stochastic local interaction (SLI) models can alleviate these problems. We discuss the benefits of SLI models and how they bridge the geostatistical and machine learning frameworks. Finally, we comment on the relation between deep neural networks (DNNs) and Gaussian processes. Surprisingly, even though the two approaches seem radically different, at certain limits they become equivalent. However, whether the above limits are in the "interesting" regime of DNNs and whether the covariance kernels of the equivalent Gaussian processes can be obtained in explicit form, are open questions for further investigation.

References

Hristopulos, D. T., Pavlides, A., Agou, V. D., & Gkafa, P. (2021). Stochastic Local Interaction Model: An Alternative to Kriging for Massive Datasets. Math. Geosci., 53(8), 1907-1949.

Hristopulos, D. T. (2020). Random fields for spatial data modeling. Springer Netherlands.

De Iaco, S., Hristopulos, D.T. & Lin, G (2022). Special Issue: Geostatistics and Machine Learning. Math. Geosci. 54, 459–465 (2022).

S2702. Some considerations on the usage of compositional data in artificial intelligence

Raimon Tolosana Delgado (Helmholtz Institute Freiberg)
Room: R5 2023-08-10 11:20

Spatial dependence between samples is not the only particularity that data science, machine learning, deep learning or artificial intelligence should deal with in the geosciences: quite pervasive also is the existence of data with multivariate unconventional statistical scales. As an example, this contribution discusses the effects of the compositional metric on predictive algorithms, i.e., the way to compare compositional data, or data informing of the relative importance of some parts forming a whole.

Most machine learning methods (lasso or ridge regression, partition trees, random forests, artificial neural networks of any depth, support-vector machines (SVM) and other kernel methods, etc.) are mostly used to extract a response prediction rule out of a training data set with known responses. If compositional data play any role, either as explanatory or as response variable, there is a chance that its specific scale affects the performance of these methods.

The compositional scale is induced by the fact that the information conveyed by compositional data is only relative (of one part to another) and for the need to work with closed and non-closed sub-compositions (formed by a subset of the parts either summing to 100% or not). Dealing with these issues is particularly eased by taking an invertible set of (log)ratios of the parts, and several alternative logratio transformations have been proposed in the literature. However, for each of the methods above there is either an objectively preferred transformation, or else they provide all identical results. In this contribution we will review the properties of the compositional scale, of the transformations to logratio scores available, and how do they interact with the methods mentioned before.

S2703. Deep autoregressive diffusion models for facies simulation

Lukas Mosser (Aker BP)
Room: R5 2023-08-10 13:40

We present an exploration of the application of deep learning in geological property modelling, specifically focusing on order-agnostic diffusion models (OADG) previously introduced by authors at Google.

This machine learning-based approach offers an effective representation of categorical properties, a crucial aspect of reservoir characterization studies in exploration and carbon capture & sequestration applications.

The OADG approach combines the strengths of diffusion models with autoregressive deep generative models, providing a granular and explicit probabilistic representation of geologic properties, while accommodating a random sequence of sampling paths for generating new realizations.

A particular advantage of OADG models over other generative models like GANs is their ability to naturally condition realizations to observed data at the grid block level and a probabilistic formulation allowing for uncertainty quantification. We introduce the technique based on connections to traditional autoregressive geostatistical models such as SIS and show examples on both simple and complex synthetic datasets.

S2704. Applications of Machine Learning in Geophysical Frontiers

Ajeet Ram Pathak (NTNU),
Anne C. Elster (NTNU),
Ole Jakob Mengshoel (NTNU)
Room: R5 2023-08-10 14:00

We are witnessing a remarkable rise in the use of data-driven Artificial Intelligence (AI) techniques, specifically in machine learning and deep learning, to address geophysical applications. This trend reflects the potential of AI-based methods in tackling the challenges posed by high dimensionality in complex networks of spatio-temporal geophysical systems. Notably, deep learning has emerged as a powerful tool alongside machine learning, with applications spanning climate modeling, earthquake prediction, seismic imaging, hydrology, temperate and precipitation forecasting, and more.

In this paper, we present a comprehensive analysis that explores the utilization of AI-based techniques to overcome diverse challenges encountered in geophysical applications. These challenges include addressing high dimensionality, quantifying uncertainty, solving higher order differential equations, and more. Our research significantly contributes by assessing approaches that combine AI and geophysical applications, paving the way for novel advancements in the field.

S2705. A Spatial-statistical model to analyse historical rutting data

Natoya O.A.S. Jourdain (NTNU),
Ingelin Steinlsland (NTNU),
Mamoona Birkhez Shami (NTNU),
Alex Klein-Paste (NTNU),
Dagfin Gryteselv (Statens Vegvesen),
Doreen Siebert (Statens Vegvesen)
Room: R5 2023-08-10 14:20

Pavement rutting is a significant concern for road authorities and transportation agencies, as it directly affects the quality, in terms of friction reduction due to water percolation, and road safety, such as impaired steerability of road users. Accurate and reliable prediction of pavement rutting can aid in proactive maintenance planning and resource allocation. We propose, fit and evaluate a statistical model for pavement rutting in a hierarchical Bayesian framework of latent Gaussian models with spatial components. Our model provides insight into pavement behaviour and deterioration on a fine-scale e.g., 20-metre stretches of the road, and has the potential to identify areas that exhibit unexpectedly high rutting rates (i.e., 'hot spots'), and predict rutting development. The model is fitted to eleven years (2010-2020) of data, including asphalt type, road width, traffic in tensity (average annual daily traffic) and rut depth measured in the previous year as explanatory variables for the European route E14, from Stjørdal to Storlien on the Swedish border. Rutting prediction is made for year 2021. We find that asphalt type is the driving factor for rutting, and there are spatial dependencies and road stretches with more than 1 millimeter rutting than expected annually. This means that the expected lifetime is halved at these stretches. We provide maps with expected rutting, and some locations have been identified for accelerated rutting, with reduction in pavement life expectancy of at least 10 years.

Keywords: Flexible pavement; predictive maintenance; rutting; spatial-statistical model; uncertainty

S2706. Using AUV to measure high gradient zones in a river plume

André Julius Hovd Olaisen (NTNU)

Room: R5 2023-08-10 14:40

This project uses AUVs to sample and explore mixing zones between fresh river water and saline ocean water. The goal is for the AUV to run transacts where the directional gradient of the salinity is the highest. The best transact is calculated by combining a physical ocean model and gaussian processes to predict where we expect to find an improvement in the highest gradient observed. The AUV has to continually assimilate the new measurements to make new predictions where the next best transact will be.

Contributions by topic

M01 Two Point Geostatistics

S0101, S0301, S0309, S0310, S0311, S0312, S0322, S1011, S1101

M02 Multi Point Geostatistics

S0321, S0803, S1005, S1201, S2503, S2504

M03 Model-Based Geostatistics

S1013, S1016, S1407, S1806, S1902

M04 Point Processes

S0114, S0323, S0605, S1002

M05 Time Series Analysis

S0111, S0304, S0504, S1007, S1008, S1010, S1011, S1015, S1016, S1017, S1503

M06 Space-Time Processes

S1001, S1002, S1013, S1014, S1207, S1304, S1401, S1806, S1809, S1813, S1814

M07 Spatial Statistics

S0101, S0104, S0201, S0321, S0601, S0602, S1001, S1003, S1004, S1009, S1012, S1017, S1101, S1105, S1501, S1902, S2503, S2602

M08 Fractal and Multi-Fractal Modelling

S1706, S1812

M09 Image Analysis

S0104, S0110, S0506, S1201, S1204, S1803, S1805, S1814, S2302

M10 Stochastic Geometry and Stereology/Mathematical Morphology

S0323, S1807, S2302

M11 Other Spatial and Space-Time Statistical Methods

S1001, S1004, S1011, S1013, S1014, S2602

M12 Electromagnetic Data: Spectral, Hyperspectral, SAR, ...

S0105

M13 Compositional Data Analysis (including distributions)

S0605, S0701, S0702, S0703, S0704, S0705, S0706, S0707, S0708, S0709, S1908, S2304, S2601, S2602, S2603, S2604

M14 Spherical Mathematics, Probability and Statistics

S2601

M15 Multivariate Statistics

S0302, S0308, S0309, S0311, S0601, S0603, S0703, S0706, S0807, S1006, S1305, S1312, S1710, S2304, S2601

M16 Bayesian Statistics

S0114, S0320, S1007, S1307, S1309, S1311, S1709, S1906, S2501

M17 Machine Learning

S0101, S0102, S0103, S0104, S0110, S0111, S0201, S0302, S0305, S0306, S0309, S0320, S0502, S0503, S0504, S0506, S0508, S0603, S0703, S0801, S0802, S0803, S0804, S0805, S0806, S0807, S0809, S0902, S0903, S0904, S1012, S1014, S1016, S1017, S1102, S1104, S1105, S1202, S1203, S1204, S1301, S1304, S1310, S1313, S1704, S1705, S1707, S1710, S1803, S1804, S1808, S1901, S1903, S1904, S1906, S2203, S2208, S2303, S2305, S2501, S2504, S2704

M18 Meshing

S2201

M19 Optimisation and Operations Researc

S0304, S0306, S1308, S1701, S1904, S1908, S2405, S2406

M21 Numerical Modelling and Numerical Simulation

S0105, S0107, S0110, S0501, S0505, S0901, S0902, S1301, S1304, S1306, S1308, S1403, S1404, S1406, S1407, S1408, S1701, S1703, S1801, S1802, S1806, S1808, S1810, S1813, S1904, S1905, S1907, S2201, S2202, S2203, S2204, S2404, S2406, S2501

M22 Data assimilation and data integration

S0805, S0807, S1007, S1103, S1201, S1202, S1312, S1704, S1709, S1805

M23 Inverse Problem solving

S1205, S1302, S1306, S1307, S1308, S1309, S1405, S1708

M24 Uncertainty Quantification

S0102, S0103, S0106, S0114, S0303, S0304, S0311, S0312, S0313, S0320, S0323, S1301, S1307, S1309, S1407, S1701, S1702, S1703, S1704, S1708, S1709, S1710, S1711, S1808, S2208, S2403, S2404, S2405, S2406, S2503, S2705

M25 3D Geodata and modeling of 3D objects

S0322, S0902, S1002, S1203, S1204, S1404, S1503, S1804, S1908, S2202, S2302, S2504

M26 Geographic and Geoscience Information Systems

S0201, S0801, S0802, S0805, S0808, S1103, S1804, S1805, S1812

M27 space/space-time geo-databases

S1103, S1502

M29 Other Computer Sciences Methods

S0808, S1101, S1202, S1706, S1708

M30 Other Mathematical Methods

S0301, S1010, S1503, S1812, S1906, S2202, S2204

M31 Other Methods

S0113, S0301, S1203, S1312, S1811, S1813, S2304, S2401, S2402, S2403

A01 Geomorphology and Quaternary geology

S1801, S1802, S1803, S1804, S1805, S1806, S1807, S1810, S1811

A02 Volcanology and Geothermal Research

S0605, S1812

A03 Stratigraphy and Sedimentology

S0110, S0901, S0902, S1015, S1401, S1403, S1407, S1503, S1705, S1804, S1806, S2202, S2501, S2604

A04 Structural Geology and Tectonics

S0802, S0804, S1101, S1102, S1201, S1204, S1206, S1308, S1309, S1404, S1407, S2202, S2501, S2504

A05 Petrology, Mineralogy and Crystallography

S0605, S0807, S1017, S1103, S1906, S2302

A06 Characterization of Oil and Gas Resources

S0101, S0103, S0104, S0110, S0111, S0114, S0808, S1004, S1007, S1010, S1014, S1205, S1311, S1312, S1403, S1701, S1702, S1706, S1708, S1709, S1710, S1901, S1902, S1904, S1905, S1906, S1907, S1908, S2201, S2406, S2504

A07 Mineral Resources

S0301, S0305, S0308, S0310, S0311, S0312, S0320, S0321, S0322, S0323, S0602, S0703, S0708, S0801, S0803, S0805, S0806, S0807, S0809, S1101, S1103, S1104, S1201, S1202, S1203, S1206, S1406, S1408, S1503, S1704, S1706, S1707, S1708, S2203, S2302, S2303, S2304, S2402, S2406, S2504

A08 Mining

S0201, S0301, S0302, S0304, S0306, S0309, S0311, S0312, S0313, S0320, S0321, S0323, S0325, S0809, S1007, S1503, S2304, S2305, S2401, S2501

A09 Geometallurgy

S0303, S0304, S0320, S0323

A10 Natural Hazard and Geological Risk Assessment

S0106, S0501, S0502, S0504, S0505, S0508, S1013, S1207, S1704, S1708, S1709, S1805, S1814, S2204

A11 Soil and Rock Mechanics and Geotechnics

S0501, S0504, S0506, S1007

A12 Palaeonthology, Paleoecology, Geobiology

S2601

A13 Geochemistry

S0602, S0603, S0605, S0701, S0704, S0706, S0707, S0709, S0803, S1103, S2304, S2601, S2602, S2603, S2604

A16 Geodynamics, Geophysics and Geodesy

S0105, S0201, S1011, S1013, S1202, S1203, S1306, S1405, S1808, S1812, S1814, S2704

A17 Remote Sensing

S0201, S0805, S1005, S1014, S1201, S1803, S1805, S1814

A18 Water: Sea, Surface and Subsurface

S0107, S0110, S0903, S1002, S1008, S1009, S1011, S1016, S1017, S1101, S1301, S1302, S1304, S1306, S1309, S1804, S1808, S1813, S2203

A19 Climatology, Meteorology, Cryology, Atmosphere and Ocean Sciences

S1001, S1005, S1006, S1011, S1013, S1017, S1806, S1808, S1812, S1813

A20 Soil Science

S1701

A21 Biosphere, Ecology, Biogeochemistry and Health

S0702, S0705, S0904, S1704

A22 Geography

S1012, S1501, S1809, S2602

A23 Other Topics in Geosciences

S0102, S0104, S0105, S0113, S0503, S0601, S1002, S1003, S1010, S1014, S1203, S1205, S1307, S1310, S1312, S1313, S1502, S1701, S1703, S1709, S1710, S1711, S1903, S1906, S2201, S2202, S2204, S2208, S2403, S2404, S2405, S2503, S2601

A24 Other Topics Outside Geosciences

S0325, S0807, S0904, S1105, S1305, S1307, S2404

Author Index