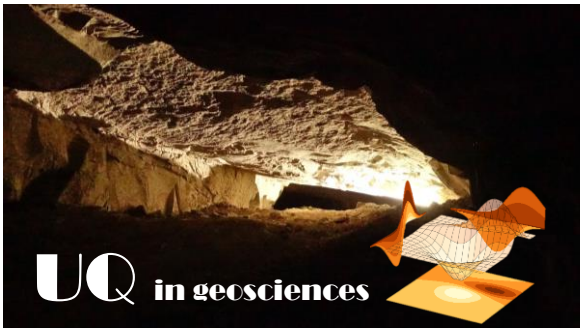

Geoscience for a sustainable Earth

brgm

Sponsors:

Partners:

Symposium on Uncertainty Quantification in Computational Geosciences
Poster presentations

2018, January 15th-17th, BRGM, Orléans, France

UQ applied to natural resources (water and mineral) exploration and exploitation	
1	*A Particle-Tracking Surrogate Model to Improve Stream-Aquifer Invers Modeling, by Yohann Cousquer et al.*
2	*Djebel El Had iron index: Geological approach, geochimics, mineralogical and structural contexts, by Hamida DIAB and Abdelmadjid CHOUABBI*
3	*On the Influence of 3-D heterogeneous geological structures on the seismic response of strategical structures, by Filippo GATTI et al.*
4	*Scenario Reduction in uranium ore deposits mining simulations by In Situ Recovery, by Jean Langanay et al.*
5	*Pre-operational risk study in deep geothermal modeling: insights from a dual medium synthetic model, by Morgan Le Lous et al.*
6	*Recent advances in the environmental characterization of Sfax aquifer, in Southeast of Tunisia: Assessment of groundwater pollution near phosphogypsum disposal, by Samira Melki and Moncef Gueddari*
UQ applied to natural risks related to geological processes	
7	*Hybrid uncertainty propagation: Application to spatialized rockfall hazard characterization, by Guillaume Dupouy et al.*
UQ applied to hydro-meteorological processes	
8	*Bayesian MCMC flood frequency analysis with reconstructed paleofloods, by Guillaume Evin et al.*
9	*Regional possibilistic sea-level projections and their use for coastal adaptation, by Gonéri Le Cozannet et al.*
10	*Extreme dependence models and Environmental contours for safety assessment, by Nicolas Raillard and Marc Prevosto*
11	*Implications of variability and lack of knowledge in future marine flooding assessments as sea-level rises, by Jeremy Rohmer et al.*
12	*Uncertainty Representations of Mean Sea-Level Change: A Telephone Game?, by Thomas van der Pol and Jochen Hinkel*
Methods and theoretical developments for UQ	
13	*Measurement Uncertainty: The First Step in the Overall Uncertainty Quantification, by Stéphane Brochot and Philippe Wavrer*
14	*Descriptive Statistics for Imprecise Data, by Marco de Angelis et al.*
15	*Applicative, methodological and software contributions to uncertainty quantification for environmental modeling, by Matthias De Lozzo et al.*
16	*Bayesian networks with imprecise probabilistic distributions to study overtopping risk in Oscillating Water Columns, by Hector Diego Estrada-Lugo and Edoardo Patelli*

Symposium on Uncertainty Quantification in Computational Geosciences
15 to 17 January 2018, Orleans, FR

A Particle-Tracking Surrogate Model to Improve Stream-Aquifer Invers Modeling

Cousquer Yohann[1], Alexandre Pryet[1], Olivier Atteia[1], Celestine Delbart[2], Remi Valois[3], Alain Dupuy[1]

[1] EA 4592 Géoressources & Environnement, Bordeaux INP and Université Bordeaux Montaigne, ENSEGID, 1 allée Daguin 33607 Pessac, France.
[2] Université François Rabelais de Tours, EA 6293 GéHCO, Parc de Grandmont, 37200 Tours, France.
[3] CEAZA, Raúl Bitrán 1305, Campus Andrés Bello, Universidad de La Serena, La Serena, Región de Coquimbo, Chile.

The inverse problem of groundwater models is often ill-posed and model parameters are likely to be poorly constrained. Identifiability is improved if diverse data types are used for parameter estimation. However, some models, including detailed solute transport models, are further limited by prohibitive computation times. This often precludes the use of concentration data for parameter estimation, even if those data are available. In the case of surface water-groundwater (SW-GW) models, concentration data can provide SW-GW mixing ratios, which efficiently constrain the estimate of exchange flow, but are rarely used. We propose to reduce computational limits by simulating SW-GW exchange at a sink (well or drain) based on particle tracking under steady state flow conditions. Particle tracking is used to simulate advective transport. A comparison between the particle tracking surrogate model and an advective-dispersive model shows that dispersion can often be neglected when the mixing ratio is computed for a sink, allowing for use of the particle tracking surrogate model. The surrogate model was implemented to solve the inverse problem for a real SW-GW transport problem with heads and concentrations combined in a weighted hybrid objective function. The resulting inversion showed markedly reduced uncertainty in the transmissivity field compared to calibration on head data alone.

Djebel El Had iron index: Geological approach, geochimics, mineralogical and structural contexts

Hamida DIAB[1], Abdelmadjid CHOUABBI[2],

[1] Laboratory of Geodynamics and Natural Resources - LGRN - BADJI Mokhtar University, National Company of Iron Mines SOMIFER, Tebessa, Algeria. E-mail: diabhamida@rocketmail.com
[2] Laboratory of Geodynamics and Natural Resources - LGRN- BADJI Mokhtar University, BP.12, 23000, Annaba, Algeria. E-mail: abdelmadjid.chouabbi@univ-annaba.dz

The region of Tebessa, by the nature of its formations and its geological history, contains great iron and / or polymetallic mining potentialities. A large number of index and deposits are located there.
The powerful marl-limestone formation, of secondary-tertiary age, affected by several tectonic phases accompanied by triasic diapiric halokinesis and by a readjustment tectonic (collapse basins), which characterizes this region, denotes the metallogenic character favorable to the presence of significant mineral concentrations (Iron, polymetals, etc.).
The most important iron deposits, located south of Tebessa, are those of Khanguet el Mouahed and Ain Babouche. Apart from these deposits in operation, many iron ore index are identified and deserve to be developed on the mining plan.
The original ore, in these indexes, is generally composed of Hematite and Siderite, accompanied, in smaller quantity of Copper, Chalcopyrite and in places by small masses of Pyrite. There are also mixtures of magnetite and hematite in the form of clusters and veins.
The mineralization manifests itself in several morphological aspects dictated by the nature of the surrounding formations (veins, clusters, etc ...).
The present study aims to contribute to the development of the Djebel El Had Iron Index, located south of Tebessa. Its objectives are to study and put this index in its geological, structural, and to decide on its scientific interest and its economic perspective.

Key words: Iron Index, Geology, Mineralogy, Geochemistry, Tebessa.

On the Influence of 3-D heterogeneous geological structures on the seismic response of strategical structures

Filippo GATTI[1], Sara TOUHAMI[1], Lúcio DE ABREU CORREA[1], Fernando LOPEZ-CABALLERO[1], Régis COTTEREAU[1], Didier CLOUTEAU[1]

[1]MSSMat, CentraleSupélec 3 Rue Joliot Curie, 91190 Gif-Sur-Yvette

In this study, the influence of the regional 3-D geological structure on the synthetic earthquake ground motion prediction is assessed, quantifying the uncertainty on the geophysical and geotechnical information available. The target is to assess the impact of wave propagation path within the Earth's crust on the seismic response of nuclear reactor buildings located nearby the seismogenic source, supporting the need for a refined source-to-site analysis at regional scale. This issue represent one of the major goal of SINAPS@ project, the first French research project aiming at quantify the uncertainty related to design procedures employed for risk and vulnerability assessment of nuclear power plants.
A suitable reference case-study is represented by the 2007 Niigata seismic sequence (west Japan), that damaged the Kashiwazaki-Kariwa Nuclear Plant. The numerical model of the earthquake scenario of the affected area (≈ 60 km wide) is built-up and small point-wise aftershocks are simulated, changing the relative source-site position. The major 3-D geological interfaces found below the nuclear site are included and the correspondent site response is compared to sub-horizontally layered configuration. Broad-band (0-7 Hz) synthetic wave-forms are obtained for two different aftershocks, reproducing the observed seismic wave focalizion at Unit 1 when the 3-D geology is included, whereas the assumption of layered geology entails a poor fit of the recordings. The different site amplification corresponding to two sources located at opposite sides of the nuclear facility highlight the strong dependency on the wave obliquity. The impact of the 3-D geology is finally quantified in terms of amplification factor at different recording stations located on the structural components of the nuclear facility.

Scenario Reduction in uranium ore deposits mining simulations by In Situ Recovery

Jean Langanay[1], Thomas Romary[1], Vincent Lagneau[1]

[1]MINES ParisTech, 60, boulevard Saint-Michel, 75272 Paris cedex 06, France

In order to plan and optimize uranium mining by ISR, the recovery of uranium is modeled by reactive transport simulation in a 3D geological model of the deposit. However, processes of reactive transport are strongly influenced by the spatial variability of the geology. The geology of the media is held by the 3D grid and separated in three properties: the oxydo-reduction code (oxidized, reduced and mineralized uranium), the lithotype of rocks (to constraint permeability and porosity) and uranium grades where there is mineralized uranium. The goal is to study the impact of the spatial variability on the simulated uranium recovery by reactive transport simulation. To reach this goal, 100 geostatistical realizations of a production block are used to sample spatial variability uncertainties. Reactive and non-reactive transport simulations are performed with the HYTEC code, developed by the Geosciences Center of Fontainebleau.
This work present a way to represent the simulated uranium recovery variability of the hundred realizations with only a few selected realizations. The method used is scenario reduction. This method is based on a measure of the differences between each realization of the set. We call this difference between realizations the similarity distance. One of the main difficulties is to find a way to compute a relevant similarity distance between each geostatistical realization with the uranium recovery of each realization. The similarirty distance will be computed from descriptors of each realization. These descriptors are computed on geological media properties and on tracer transport simulations. Tracer transport simulations are done in order to try to approximate the uranium dissolution. The scenario reduction method uses the similarity distance and descriptors values to select a subset of realizations in the whole set, through a clustering algorithm. The uranium recovery variability of the subset should be similar to the uranium recovery variability of the whole set of realizations. Results are studied by comparing the uranium recovery of the selected set of realizations and the uranium recovery of the whole set of realization.

Pre-operational risk study in deep geothermal modeling: insights from a dual medium synthetic model

Morgan Le Lous[1], Alexandre Pryet[2], François Larroque[3], Pierre-Clément Damy[4], Alain Dupuy[5]

[1]EA Géoressources & Environnement, 1 allée F. Daguin, 33607 Pessac Cedex, France, morgan.le_lous@ensegid.fr
[2]EA Géoressources & Environnement, 1 allée F. Daguin, 33607 Pessac Cedex, France, alexandre.pryet@ensegid.fr
[3]EA Géoressources & Environnement, 1 allée F. Daguin, 33607 Pessac Cedex, France, francois.larroque@ensegid.fr
[4]Fonroche Géothemie, Technopole Hélioparc, 2 avenue P. Angot, 64053 Pau, France, pc.damy@fonroche.fr
[5]EA Géoressources & Environnement, 1 allée F. Daguin, 33607 Pessac Cedex, France, alain.dupuy@ensegid.fr

Despite their potential, very deep geothermal systems in sedimentary basins remain less exploited than shallow geothermal energy. The uncertain effectiveness and profitability of such deep geothermal systems discourage their use. The potential of heat production and the impact of exploitation of resources must be appreciated with regard to local physical subsurface properties.
The development of geothermal energy generation is closely linked to thermal and hydrogeological knowledge of the subsurface aquifers. Numerical modeling here appears as a tool to delineate development risks induced by limited geological data at great depths.
Computational tools capable of simulating complex geothermal systems coupled to adapted numerical methods of uncertainty design are tailored to evaluate the pre-operational risk associated with the deep geothermal site-specific operation.
IFPEN in-house reservoir numerical model PumaFlow™, coupled with extensive uncertainty and optimization analysis software CougarFlow™, constitutes a reservoir modeling chain capable of investigating effects of input parameters on simulation results.
Investors may further calculate the financial risk, and operators may adjust their exploitation strategy for the entire life-cycle of the reservoir. This integrated work tackles challenges faced in classical stochastic hydrogeological modeling by providing an operational and process-based approach for deep geothermal energy system.

This research is carried out in the FONGEOSEC – AMI project handled by Fonroche Géothermie.

Recent advances in the environmental characterization of Sfax aquifer, in Southeast of Tunisia: Assessment of groundwater pollution near phosphogypsum disposal

Samira Melki[1 (a)], and Moncef Gueddari[1 (b)]

[1] Laboratory of Geochemistry and Environmental Geology, Department of Geology , Faculty of Sciences of Tunis , University of Tunis El Manar, 2092 Tunis, Tunisia.
[(a)] Corresponding author:
E-mail: melkisamiralefi@hotmail.fr
[(b)] Co-author:
E-mail: moncef_gueddari@gmail.com

The production of phosphoric acid by the Tunisian Chemical Group, in Sfax, Tunisia, led to the degradation of the groundwater quality of the Sfax-Agareb aquifer mainly by the phosphogypsum leachates infiltration. Spatio-temporal monitoring of the quality of groundwater was carried out by performing bimonthly sampling between October 2013 and October 2014. Samples culled in the current study, were subject to physicochemical parameters measurements and analysis of the major elements, orthophosphates, fluorine, trace metals and stable isotopes (18O, 2H). The obtained results show that the phosphogypsum leachates infiltration has a major effect on the downstream part of the aquifer, where the highest values of conductivity, SO4 2- ; Ortho-P and F- and the lowest pH were recorded. In addiction, these results indicated that phosphogypsum leachates contained much higher amount of Cr, Cd, Zn, Cu, Fe and Al compared to the groundwater. Spatio-temporal variation of the conductivity and concentrations of major elements is linked to the phosphogypsum leachates infiltration as well as to a wide range of factors such as the natural conditions of feeding and the water residence time. Contents of 18O, 2H, showed that the water of the Sfax-Agreb aquifer undergo a large scale evaporation process originated from recent rainfall.

Keywords: Groundwater quality, phosphogypsum leachate, trace metals, stable isotopes, Sfax-Agareb aquifer, Tunisia.

Hybrid uncertainty propagation: Application to spatialized rockfall hazard characterization

Guillaume Dupouy[1], Jean-Marc Tacnet[1], Franck Bourrier[1] and Frédéric Berger[1]

[1]IRSTEA, 2 rue de la Papèterie, 38402 Saint-Martin-d'Hères, France, guillaume.dupouy@irstea.fr

Rockfalls can cause severe damage to exposed population and infrastructures in mountainous area. Risk analysis requires characterizing the hazard related to these phenomena, and uncertainty analysis on numerical simulation models is a major tool for such purpose. Though, thematic information about the natural environment (rock volumes, geology…) remains usually scarce: precise and long series of observations are often not available for a specific area of study, and expertise remains essential. Plus, the uncertainties related the local topography, modeled by Digital Elevation Models (DEM), have to be considered, as local slopes influence the trajectory of a rock.
Both thematic and spatial uncertainties have been assessed and propagated through a 3D rockfall propagation simulation model, estimating kinetic energy quantile on a training rockfall simulation case. Two alternative propagation methods have been used, namely the usual probabilistic Monte-Carlo method, and the possibilistic Hybrid approach. This last method uses probability, possibility and belief function theories, used as tools for coding sets of probability distributions, and thus consider the different aspects of information imperfection (randomness, lack of information, inaccuracy of measure…).

Keywords: uncertainty propagation, hybrid, probability, possibility

Bayesian MCMC flood frequency analysis with reconstructed paleofloods

Guillaume Evin[1], Bruno Wilhelm[2], and Jean-Philippe Jenny[3]

[1]Univ. Grenoble Alpes, Irstea, UR ETGR, F-38402 St-Martin-d'Hères, France
[2]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, F-38000 Grenoble, France
[3]Max Planck Institute for Biogeochemistry, 07701 Jena, Germany

The choice of an acceptable and cost-effective solution for the design of hydraulic structures depends upon the estimation of quantiles for different characteristics of floods, usually peak flows. However, series of observed floods have a limited length, and quantile estimates associated to high return periods are subject to large uncertainties. In this study, we propose a novel and complementary approach which aims at combining reconstructed peak flows of the Rhône river with the series of observations. These reconstructions cover the last 350 years and are obtained using measurements of sediment volumes in the Bourget Lake[1] (at Aix-les-Bains, France).
A Bayesian approach is adopted in order to properly treat the non-systematic nature of the reconstructed flow data, as well as the uncertainties related to the reconstruction method. While this methodology has already been applied to historical floods, similar applications to paleofloods are absent and promising. We first estimate extreme quantiles using direct measurements of peak flows (1853-2004). Direct observations are then combined to the sedimentary information (1650-2013). The comparison of the resulting estimates demonstrates the added value of the sedimentary information, and its impact on the associated uncertainties. In particular, 4 major floods which have occurred during the 18th century are very unlikely in comparison to the floods observed during the last 150 years.

Keywords: Flood quantiles, long series, sedimentary measurements

[1] Jenny J.-P., Wilhelm B., Arnaud F. , Sabatier P., Giguet-Covex C. , Mélo A. , Fanget, B., Malet E., Ployon E., Perga M.E. (2014) A 4D sedimentological approach to reconstructing the flood frequency and intensity of the Rhône River (Lake Bourget, NW European Alps). Journal of Paleolimnology 51(4): 469-483

Regional possibilistic sea-level projections and their use for coastal adaptation

Gonéri Le Cozannet[1], Gael Durand[2], Jean-Charles Manceau[1], Angélique Melet[3], Benoit Meyssignac[4], Catherine Ritz[2], Jeremy Rohmer[1], David Salas-y-Mélia[5]

[1]BRGM, 3, av. Claude Guillemin, BP 36009, 45060 Orleans Cedex 2, France, g.lecozannet@brgm.fr
[2]Institut des Géosciences de l'Environnement (IGE), 54 rue Molière - Domaine Universitaire - BP 96 - 38402 St Martin d'Hères Cedex, France
[3]Mercator, Parc Technologique du Canal, 8, Rue Hermès, 31520 Ramonville Saint Agne, France
[4]LEGOS - UMR 5566 | CNES **-** CNRS - IRD - UPS Toulouse III. 14, avenue Edouard Belin,Toulouse, France
[5]CNRM – UMR 3589, 42,avenue Gaspard Coriolis 31057 Toulouse Cedex 1 France, France

As climate changes, sea-level rise is aggravating coastal hazards such as flooding, coastal water salinization and shoreline changes. Current sea-level projections are provided in the form of median, likely and probabilistic products. They are essentially useful to plan optimum responses, in cases where the exposure and vulnerability is medium to low. However, the worlds' coastal zones include vital locations such as major cities, marine transportation and industrial centers, so that a failure in their adaptation planning may have foremost macro-economic and social implications. For these areas, there is a need for other types of sea-level projections, conveying uncertain sea-level rise projections appropriately. Here, we provide regional possibilistic sea-level projections compliant with the IPCC assessment and exploring high ends, for each components, based on an extensive review of the literature. We illustrate how these projections can be combined with the knowledge available regarding vertical ground motions to estimate future coastal impacts and support adaptation planning in areas were risk acceptability is low.

Extreme dependence models and Environmental contours for safety assessment

Nicolas Raillard[1], Marc Prevosto[1]
[1]IFREMER, Laboratoire Comportement des Structures en Mer, ZI de la Pointe du Diable, 29820 Plouzané.

In coastal area, both human activities and structures are subject to the impact of extreme events, with possible dramatic consequences. However, high losses may not be caused by a single environmental parameter being extremely high, but by the conjunction of high levels on several variables: for instance, coastal flooding may be caused by a conjunction of high storm surge, high tide and large waves, but not only with one variable being extreme. In such a multivariate context, the environmental contour concept is an effective, risk-based approach in establishing design conditions.
In this work, we will compare two methods for modeling of extremes in three dimensions, via extreme copula approaches and conditional models that have be recently proposed by Heffernan & Tawn [1]. Along with the modeling of extreme values, we will compare two methods for deriving environmental contours: the classical Rosenblatt transformation from the environmental parameter space to a standard normal space, and the newly proposed Huseby method [2] which establishes environmental contours based on direct Monte Carlo sampling from the joint distribution.

References :
[1] Heffernan, J.E., Tawn, J.A., 2004, "A conditional approach for multivariate extreme values (with discussion)", Journal of the Royal Statistical Society: Series B, vol. 66, no. 3, pp. 497-546.
[2] Huseby, A.B., Vanem, E., Natvig, B., 2015, "Alternative environmental contours for structural reliability analysis", Structural Safety, vol. 54, pp. 32–45.

Implications of variability and lack of knowledge in future marine flooding assessments as sea-level rises

Jeremy Rohmer[1], Goneri Le Cozannet[1], Jean-Charles Manceau[1]

[1]BRGM, 3, av. Claude Guillemin, BP 36009, 45060 Orleans Cedex 2, France, j.rohmer@brgm.fr

Despite progresses in climate change science, projections of seal level rise (SLR) remain highly uncertain. Recent studies have highlighted the limitations of probabilities as a tool to account for the intrinsically imprecise nature of SLR, i.e. the lack of knowledge surrounding it (aka epistemic uncertainty). The added value of using alternative representation tools have recently been discussed in the literature by focusing on extra-probabilistic theories (like the possibility theory), which avoids selecting a unique probability laws when the level of uncertainty is too high and enables to bound all the possible ones consistent with the available data. By adopting these new mathematical tools, we aim at evaluating how the epistemic uncertainties related to SLR accumulates to other sources of uncertainties (epistemic or aleatoric) for local flooding impact assessment, namely: the choice in Representative Concentration Pathway (RCP) scenario, the high-end scenario, the regional bias, the contributions of extremes and wave effects. The case study corresponds to a local low-lying coastal urban area exposed to storm surge and waves in the north-western Mediterranean coast. We focus on the time evolution of the probability of flooding Pf defined as the probability of exceeding a critical threshold corresponding to the height of coastal defenses or low walls. The joint sensitivity analysis of the probabilistic, possibilistic and scenario-like sources of uncertainty enables to highlight the key contributions of RCP and global SLR both on the best estimate of Pf and on the imprecision affecting it. To a lesser extent, extreme water levels can also play a key role.

Uncertainty Representations of Mean Sea-Level Change: A Telephone Game?

Thomas van der Pol[1], Jochen Hinkel[2]

[1] Global Climate Forum, 6, Neue Promenade, 10178 Berlin, Germany, thomas.van.der.pol@globalclimateforum.org
[2] Global Climate Forum, 6, Neue Promenade, 10178 Berlin, Germany, hinkel@globalclimateforum.org

In the chain from producing sea-level rise (SLR) information to using this information for coastal decision-making, a variety of uncertainty representation methods (URMs) and conversions between these methods are applied. In every conversion, additional assumptions are made. This has been a source of confusion and misinterpretation in SLR literature. To address this limitation, this paper first categorises URMs along two dimensions: (i) scenarios or a prediction, and (ii) a deterministic, interval, imprecise probabilistic, partial probabilistic, imprecise partial probabilistic or probabilistic representation. Next, the URM categories are applied to analyse how total mean SLR uncertainty is represented and converted within the following three strands of literature: SLR projections, impact assessments and decision analyses. We find that SLR projections mostly use partial probabilistic or probabilistic scenarios. We show that the likely SLR ranges in the IPCC report fall within the URM category of imprecise partial probabilistic scenarios. Impact assessments of SLR and decision analyses mostly use deterministic scenarios for representing mean SLR uncertainty. We demonstrate that the transfer of SLR information across this interface involves multiple conversions and a sequence of additional assumptions. The potential for a more consistent treatment of mean SLR uncertainty across studies is discussed. The latter may help to prevent the reduction of the quality and loss of SLR information in order to better prepare coastal zones for uncertain changes in relative mean and extreme sea levels.

Measurement Uncertainty: The First Step in the Overall Uncertainty Quantification

Stéphane Brochot[1], Philippe Wavrer[1]

[1] CASPEO, 3, av. Claude Guillemin, BP 36009, 45060 Orleans Cedex 2, France,
email s.brochot@caspeo.net, p.wavrer@caspeo.net

Most of the models used in geosciences are based on measurements, or are validated or calibrated using measurements. Consequently, the quality of a model is dependent on the quality of those measurements. There are then two main sources of uncertainty for a model: the uncertainty due to the fact a model is just mind's eye and present a gap with reality, and the uncertainty linked to the measurement used to confront the model with reality. Keeping in mind a measurement is also based on a model with its inherent uncertainty (what Pierre Gy was naming the scientific error), any uncertainty quantification of a model starts with the deconvolution of these different uncertainty sources.

Measurements are mainly performed on samples which are supposed to represent the lot subject of the study. Therefore, the measurement error is not limited to the analysis error, and a sampling error must be accounted too. In geosciences, the lots of material are characterized by their heterogeneity in constitution as well as in distribution. In some cases, this heterogeneity is so large that the sampling error can be dramatically high.

The representativeness of samples has been largely studied by the Theory Of Sampling (TOS, based on the Pierre Gy's theory) and well associated with the theory of geostatistics. TOS offers a set of mathematical tools allowing quantification of the main components of the sampling error. But these tools necessitate to get a good description of the material heterogeneity. This description is called "heterogeneity model". The quality of quantification of the sampling error depends on the pertinence of the heterogeneity model.

For its activities in process modelling and simulation, and in material balance calculation, Caspeo has developed an expertise in sampling and measurement uncertainty quantification. This includes the development of a methodology to build the heterogeneity model and then to choose the most appropriate mathematical tool to quantify the uncertainty. The heterogeneity model gives also a lot of information about material behaviour during operations of sample taking and preparation, allowing to recommend high care during these operations.

When applied to exploration and mining exploitation, the measurement uncertainty has an economic impact, generating financial risk. When applied to environmental studies, it takes part in the environmental risk assessment. In the field of predictive models, largely based on statistics over measured data, measurement uncertainty must be introduced to calculate the prediction uncertainty.

Descriptive Statistics for Imprecise Data

Marco de Angelis[1], Scott Ferson[1], Luke Green[2]

[1]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, L69 7ZF, UK, marco.de-angelis@liverpool.ac.uk
[2]Vivaldi Analytics, Stony Brook, New York, US, luke@dataclimate.co

As a discipline, the theory of imprecise probabilities may be pricing itself out of the market in the sense that its complexity, computational burden, and requisite mathematical sophistication required for nontrivial applications are prohibitive in many subject domains. For the discipline to grow, it is essential to foster broad interest and use across science and engineering. This will involve recruiting a class of users who may not develop methods but who will apply them in their routine work. Their applications give evidence of the utility of the imprecise probabilities approach and its underlying philosophy. This implies that someone who sees imprecision in a data set but who lacks special training in uncertainty quantification or imprecise probabilities should be able to apply convenient algorithms for basic statistics.
When the data set has imprecision, computing statistics can be challenging. For example, for data in the form of intervals, using naïve interval analysis yields results with inflated uncertainty because of repetitions of variables in the formulas. Moreover, finding optimal bounds on many basic statistics are NP-hard problems that grow in difficulty with the size of the data set. It is practically impossible to solve these problems for large data sets with a simple sampling strategy, such as Monte Carlo, in which the formula for the variance is treated like a black box evaluated for many possible configurations of the data points within their respective intervals.
Over the last century, statistics has focused on developing methods for analyses in which data sample size is limiting. But not all uncertainty in data has to do with small sample sizes. Although most statistical analyses today ignore the uncertainty reported by laboratories and empiricists as interval measurement uncertainty, this is clearly not always because this uncertainty is negligibly small. We believe it may instead be due to the lack of friendly software to handle it. We announce and describe a software library that is intended to provide convenient access to basic statistics for interval and censored data. The library of algorithms is being used to develop on-line and stand-alone software for analyzing data sets containing imprecision as well as sampling uncertainty. The algorithms in the library require users to make fewer dubious assumptions about the data set than currently popular methods for handling data censoring, missingness, and lack of independence. The library currently supports methods to compute over two dozen measures of location, dispersion and distribution shape, including arithmetic, geometric and harmonic means and median, variance, confidence intervals, histogram, and several inferential methods for linear and logistic regressions, t-tests, F-tests, and outlier detection. We show the accuracy of the proposed rigorous approaches via numerical comparisons between them and other bounding techniques like global optimisation, and with other traditional statistical methods for handling censored data.

Applicative, methodological and software contributions to uncertainty quantification for environmental modeling

Matthias De Lozzo[1], Sophie Ricci[1], Mélanie C. Rochoux[1], Pamphile T. Roy[2], Paul Mycek[2], Vanessya Laborie[4], Anne-Laure Tiberi-Wadier[3], N. Frebourg[5], C. Lamotte[5], G. Rea[1]

[1] CECI, Université de Toulouse, CNRS, CERFACS, 42 avenue Gaspard Coriolis, 31057 Toulouse, France
[2] CERFACS, 42 avenue Gaspard Coriolis, 31057 Toulouse, France
[3] CEREMA, 155 rue Pierre Bouguer, BP 5, 29280 Plouzané Cedex, France
[4] CEREMA/LHSV, 134 rue de Beauvais, CS 60039, 60280 Margny Lès Compiègne
[5] ENM, 42 avenue Gaspard Coriolis, 31057 Toulouse, France

Uncertainty quantification (UQ) has become a mandatory step in numerical sciences in a context of massively parallel computation with high fidelity modelling software. UQ provides a framework for sensitivity analysis to quantify the uncertainty in the model outputs that are due to various model uncertainties. UQ leads to classifying sources of uncertainties and identifying which ones should be reduced with data assimilation (DA). We focus on non-intrusive ensemble methods in order to deal with both black-box and open-source computer codes, using APIs when available. Our applications relate to computational fluid dynamics (CFD), especially environmental risk related topics such as flood forecasting, wildland fire spread, micro-scale dispersion of air pollutant, but also industrial CFD with combustion chamber ignition.

Sensitivity analysis allows to identify the major sources of uncertainty for the quantity of interest, and thus to identify the control vector for DA. Our focus is on Sobol' indices which measure shares of output variance imputable to these sources. We propose new methods in the case where inputs and outputs are functional, notably in terms of visualization. For instance, we improve recent graphical techniques for representing the statistics and Sobol' indices associated with 1D, 2D and 3D variables and open the way to audio features. We also propose new methods to better track position errors and go beyond Euclidean measures to represent discrepancies in terms of front shape and topology and analyze their sensitivity to uncertain input parameters.

We currently develop strategies to reduce the cost of stochastic estimations (expectation, variance, threshold excess, probability density functions, covariance matrices, sensitivity indices, ...). A first one consists of building a surrogate model from a limited number of model evaluations, whose statistics and surface response can be obtained at a limited cost. Using a polynomial chaos expansion in place of the full direct model is considered as a promising strategy to reduce the ensemble size and alleviate the computational cost of the forecast step in ensemble-based DA. We also consider the multifidelity EvoFusion technique, which consists of combining different levels of accuracy when building the surrogate model. Furthermore, we pursue our efforts to develop cost-effective Gaussian process and polynomial chaos methods able to address large uncertain input parameters that are indexed by time or space through Karhunen-Loève transformation. Lastly, we currently develop multilevel Monte-Carlo strategies, which combine solutions of the direct model discretized on coarse-to-fine resolution meshes.

CERFACS is actively developing the open source BATMAN Python package which is dedicated to statistical analysis of numerical simulators based on non-intrusive ensemble simulation (available on GitLab). It provides a convenient, modular and efficient framework for design of experiments, resampling, surrogate modelling, uncertainty quantification and sensitivity analysis relying on existing libraries (OpenTurns, Scikit Learn). It can be used either by writing a Python script or filling a JSON file.

Bayesian networks with imprecise probabilistic distributions to study overtopping risk in Oscillating Water Columns

Hector Diego Estrada-Lugo[1], Edoardo Patelli[2]

[1]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, L69 7ZF, United Kingdom. h.d.estrada-lugo@liverpool.ac.uk
[2]Institute for Risk and Uncertainty, University of Liverpool, Peach Street, L69 7ZF, United Kingdom. epateilli@liverpool.ac.uk

Bayesian networks enhanced with structural reliability methods are a useful and reliable tool to study complex engineering systems that involve a number of variables that influence the occurrence of wanted or unwanted events. Despite the great number of advantages that this method has, traditional Bayesian networks are restricted to the use of discrete or Gaussian distributions that bring some disadvantages, e.g. discretization of data and impoverishing of information. However, the enhanced Bayesian networks inherit all the advantages of the traditional approach and allows working with continuous distributions without losing of information. As a result of this approach, uncertainty of the variables involved in the network is taken into account since the behaviour of the mentioned parameters is described through the use of probabilistic functions instead of only discrete probability values. Nevertheless the number of advantages of this method, some studies are affected by the lack of data and the high uncertainty of variables involved, especially those concerning to climate events that affect technological facilities, that poses a great contribution to imprecise predictions or analyses with low-veracity results. In the present work, the implementation of the so-called Credal networks permits the representation of the lack of information through the use of probability intervals, or p-boxes. Such a network is a representation of a set of Bayesian networks that owns the same graphical disposition but are associated to different probability measures. The methodology was implemented in the general-purpose software OpenCossan due to the high flexibility and potential to improvement.

As a case study, data of an Oscillating Water Column experimental model was considered to study the risk of overtopping event for different configurations of harbour walls.