Workshop

Approximating the response of complex hydrological systems: from theory to real-world applications

Università Ca' Foscari Venezia, 3 giugno 2025

Department of Environmental Sciences, Informatics and Statistics

Venue: Sala conferenze Orio Zanetto, Campus Scientifico via Torino;

Via Torino 155 - 30170 Venezia Mestre

Program:

- **10:00 - 10:40:** Adam Siade, Reduced-dimensional gaussian process machine learning for groundwater allocation planning using swarm theory
- **10:45 - 11:25:** Reygie Macasieb, A probabilistic approach to surrogate-assisted multi-objective optimization of complex groundwater problems

Coffee Break

- **11:40 -12:20:** Federico Piazzon, Surrogate modelling and sensitivity analysis of Kelvin-Voigt viscoelastic flow by polynomial approximation

Lunch Break

- **14:00 - 14:40:** Boumediene Hamzi, Bridging machine learning, dynamical systems, and algorithmic information theory: Insights from sparse kernel flows, Poincaré normal forms and PDE simplification
- **14:45 - 15:25:** Antonia Larese - Damiano Pasetto: RBF-based Surrogates of a high-fidelity simulation model of a debris flow

This workshop is funded by the project Hydro-ROM: "Reduced order models of hydraulic protection systems for extreme water hazards", PRIN222022PXYYK5, CUP: H53D23001350006. Organizers: Damiano Pasetto, Gabriele Santin, Antonia Larese

Reduced-Dimensional Gaussian Process Machine Learning for Groundwater Allocation Planning Using Swarm Theory

Adam Siade

Commonwealth Scientific and Industrial Research Organization (CSIRO), Perth

email: adam.siade@uwa.edu.au

Groundwater management and allocation planning involves a rigorous assessment of the performance of operational decisions such as extraction/injection rates on community and environmental objectives. Maximizing performance through numerical optimization can be essential for high-value resources and is often computationally infeasible due to long simulation model run times combined with nonconvex objectives and constraints. In order to mitigate these drawbacks, surrogate models can be used in place of complex models during the optimization process. There exist a number of machine learning techniques that can be used to develop a data-driven surrogate model. However, the curse of dimensionality, common to groundwater management, limits the use of these techniques due to the necessity for large training data sets. Even though it is now possible to handle large data sets, the generation of these data sets themselves remains computationally prohibitive as they require numerous simulations to produce accurate surrogates. In this study, we integrate a dimensionality reduction method using truncated singular value decomposition to reduce the number of decision variables, thereby reducing the size of the training data set needed. Correspondingly, we demonstrate a simple technique for acquiring an approximate minimax Latin Hypercube design from within the subspace. We also implement a novel technique for adaptive resampling through particle swarm optimization in order to maintain accuracy of the surrogate model throughout the optimization process. The resulting accurate surrogate model for the Perth regional aquifer system of Western Australia runs in a matter of seconds. Adopting this approach can produce timely solutions, making formal optimization tractable for practitioners.

A Probabilistic Approach to Surrogate-assisted Multi-objective Optimization of Complex Groundwater Problems

Reygie Macasieb

The University of Western Australia, Perth

email: reygie.macasieb@research.uwa.edu.au

Groundwater management involves a complex decision-making process, often with the need to balance the trade-off between meeting society's demand for water and environmental protection. Therefore, effective management of groundwater resources often involves some form of multi-objective optimization (MOO). Many existing software tools offer simulation model-enabled optimization, including evolutionary algorithms, for solving MOO problems. However, such analyses involve a huge amount of numerical process-based model runs, which require significant computational effort, depending on the nonlinearity and dimensionality of the problem, to seek the optimal trade-off function known as the Pareto front. Surrogate modeling, through techniques such as Gaussian Process Regression (GPR), is an emerging approach to significantly reduce the number of these model evaluations thereby speeding up the optimization process. Yet, surrogate model predictive uncertainty remains a profound challenge for MOO, as it could mislead surrogate-assisted optimization, which may result in either little computational savings from excessive retraining, or lead to suboptimal and/or infeasible solutions. In this work, we present probabilistic Pareto dominance criteria that considers the uncertainty of GPR emulation during MOO, producing a "cloudy" Pareto front which provides an efficient decision space sampling mechanism for retraining the GPR. We then developed a novel acquisition strategy to manage the solution repository from this cloud and generate an ensemble of infill points for retraining. We demonstrate the capabilities of the algorithm through benchmark test functions and a typical density-dependent coastal groundwater management problem.

Surrogate modelling and sensitivity analysis of Kelvin Voigt viscoelastic flow by polynomial approximation

Nicolò Crescenzo[1], Antonia Larese[1], Federico Piazzon[1], Mario Putti[2]

[1]Department of Mathematics Tullio Levi-Civita, Università degli Studi di Padova; [2]Department of Agronomy, Food, Natural resources, Animals and Environment, Università degli Studi di Padova

email: fpiazzon@math.unipd.it

Kelvin Voigt model is one of the most commonly used mathematical description of linear viscoelastic fenomena. In such a modelization the pair of vector fields *(v, u)* describing the velocity and the displacement of the body is the solution of a linear parabolic system of PDEs.

We tackle the numerical approximation of such a system by Galerkin method (FEM or spectral elements) for the space discretization, and either exact time integration or backward Euler method for time discretization. An interesting feature of our method is the capability of describing heterogeneous or composite materials.

When the mechanical features of the considered media are unknown or possibly affected by error, a study of approximation in the parameters space and sensitivity analysis are mandatory. In such a framework, we heavily exploit the analytic dependence of the solution with respect to the parameters to apply theoretical results of polynomial approximation theory and complex analysis. Finally, the computational issues are treated by the strategies recently developed in the framework of polynomial admissible meshes.

We provide theoretical bounds on the error on solution and on sensitivities. Moreover, numerical results show that, by a reasonable number of solutions of the forward model, our method is capable of constructing a surrogate model having essentially the same precision of the full model.

Bridging Machine Learning, Dynamical Systems, and AlgorithmicInformation Theory: Insights from Sparse Kernel Flows, Poincaré Normal Forms and PDE Simplification

Boumediene Hamzi

Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA, USA, and The Alan Turing Institute, London, UK.

email: Boumediene.hamzi@gmail.com

This presentation delves into the intersection of Machine Learning, Dynamical Systems, and Algorithmic Information Theory (AIT), exploring the connections between these areas. In the first part, we focus on Machine Learning and the problem of learning kernels from data using Sparse Kernel Flows. We draw parallels between Minimum Description Length (MDL) and Regularization in Machine Learning (RML), showcasing that the method of Sparse Kernel Flows offers a natural approach to kernel learning. By considering code lengths and complexities rooted in AIT, we demonstrate that data-adaptive kernel learning can be achieved through the MDL principle, bypassing the need for cross-validation as a statistical method.

Transitioning to the second part of the presentation, we shift our attention to the task of simplifying Partial Differential Equations (PDEs) using kernel methods. Here, we utilize kernel methods to learn the Cole-Hopf transformation, transforming the Burgers equation into the heat equation. We argue that PDE simplification can also be seen as an MDL and a compression problem, aiming to make complex PDEs more tractable for analysis and solution. While these two segments may initially seem distinct, they collectively exemplify the multifaceted nature of research at the intersection of Machine Learning, Dynamical Systems, and AIT, offering preliminary insights into the synergies that arise when these fields converge.

RBF-based surrogates of a high-fidelity simulation model of a debris flow

Eleonora Spricigo[1], Deependra Kumar[1], Antonia Larese[1], Damiano Pasetto[2], Mario Putti[3]

[1]Department of Mathematics Tullio Levi-Civita, Università degli Studi di Padova; [2]Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice; [3]Department of Agronomy, Food, Natural resources, Animals and Environment, Università degli Studi di Padova

email: antonia.larese@unipd.it ; damiano.pasetto@unive.it

In the last decades we have observed a rapid growth of extreme hydrological events, such as floods and rock/debris or mud flows affecting more and more frequently our lives. The detailed physical description of these viscous fluids is fundamental to understand the caused stress on possible flood control structures, such as levees, dams, check dams. However, its simulation through high fidelity physics-based computational models, using for example the Material Point Method (MPM), is extremely computationally demanding, thus limiting the application to real system monitoring.

The development of surrogate models to efficiently replicate the relevant features of the flow is of paramount importance to make a substantial step in the direction of real-time computations, required in any early warning system and to develop mitigation strategies.

Surrogate models have gained significant attention in recent years, especially with the advent of machine learning and the development of neural network-based methods, such as Fourier Neural Operators and Deep Operator Networks, among others.

Here we consider surrogates based on Kernel methods, which demonstrated distinct advantages over widely used neural network-based approaches and provide rigorous error analysis. As fractal functions are pivotal in addressing nonlinear and irregular problems, we propose using the recently developed fractal RBFs as kernel of the surrogate model.

To demonstrate the effectiveness of the proposed approach, we consider a 2D debris flow along a 5m flume as a test scenario, where the outputs of interest are the position of the front and the velocities as functions of the fluid density and the inclination angle of the slope. Our results explore the accuracy and computational efficiency of the fractal RBF surrogate model compared to other kernel-based approaches.