Aims

The aims of this course are to develop techniques of data assimilation in mathematical biology, ecology and medicine. This will be achieved by a mixture of lectures on basic methodology, tutorial exercises, computer labs, case studies and a large group-based modelling, implementation and simulation project. We will introduce a number of data assimilation approaches that are widely used in applications to mathematical biology, ecology and medicine, including basic regression analysis, variational approaches, Kalman filtering, extended and ensemble Kalman filtering and the Bayesian inference approach. The course will teach practical implementation of these data assimilation techniques in the context of computer simulations, which will be illustrated by prototype applications from predictive modelling in biology, ecology and medicine. These methodologies will form the basis for a series of modelling case studies as well as the group-based project component of the course.

Syllabus

Background/Review:: Data sets, statistics, re-analysis; polynomial interpolation (Legendre, Chebychev, etc) and errors; Fourier Transform and Fast Fourier Transform (theory and practical implementation). (3 lectures)

Variational Approach:: Cost functions; regression analysis review (least squares, linear and nonlinear models); Optimal Interpolation, 3DVar and 4DVar approaches, applications to Numerical Weather Prediction. (4 lectures)

Kalman Filtering:: Basic ideas (filtering/smoothing) of Kalman Filter; Extended Kalman Filter (nonlinear); Ensemble Kalman Filter (sequential MC methods), numerical implementation, Case study on the ensemble Kalman filter applied to the Lorenz system. (4 lectures)

Bayesian Inference Approach:: Basic ideas, Bayes Theorem, prior and posterior distributions (selection and interpretation); implementation: acceptance-rejection sampling, MCMC approach to target distributions (Metropolis-Hastings), examples from diffusion problems, wave equation, fluid mechanics, geophysics and molecular dynamics. Discussion of general Inverse Problems, Uncertainty Quantification and Extreme Events. (4 lectures)

Modelling, Data Assimilation and Simulation Project:: Group-based work on an extended project on Ecology, including the modelling and subsequent direct numerical implementation of one or more of the data assimilation approaches above. The project includes a background literature search, development of the underlying model, assessment of the data and appropriate data assimilation techniques that can be applied, hands-on simulation of one or more of these techniques, a group-based presentation and a written report. (15 lectures)
Teaching and Assessment

Contact Hours: 3 lectures and 1 tutorial per week
Assessment: 50% by class tests or other continuous assessment
50% by end of course 2-hour exam
Resit Type: exam
By the end of the course, students should be able to:

- **Understanding, Knowledge and Cognitive Skills**
  - Understand the role of data inclusion in models and the improvement in model prediction attainable.
  - Collate background knowledge on polynomial interpolation and associated errors, re-analysis, least squares for linear and nonlinear models.
  - Implement the Fast Fourier Transform to solve partial differential equations (1+1 dimensions) using spectral methods.
  - Formulate and implement the Optimal Interpolation variational approach to Numerical Weather Prediction, in particular the 3DVar and 4DVar approaches.
  - Understand the basic ideas underlying the Kalman filter and how to implement it.
  - Understand the basic ideas underlying the extended Kalman filter, and its implementation in practice.
  - Understand the role of ensemble Kalman filtering, its use in practice and its practical implementation.
  - Formulate Bayes theorem, the selection and interpretation of the prior and posterior distributions and their consequences in practice.
  - Apply various implementations of the Bayesian approach, including acceptance-rejection methods, MCMC approach to target distributions including the Metropolis-Hastings algorithm.
  - Formulate the Bayesian approach in example problems, including: diffusion problems, wave equation, fluid mechanics, geophysics and molecular dynamics.
  - Understand the derivative problems of Uncertainty Quantification and Extreme Events.
  - Implement one or more of the above data assimilation approaches as part of a group-based project on Ecology (eg. predict future vegetation distribution).
  - Understand the relative capabilities and limitations of all of the above data assimilation techniques.

- **Scholarship, Enquiry and Research (Research-Informed Learning)**
  - Understand how to apply a range of modelling, data assimilation and simulation methodologies in real research problems.
  - Critically interpret the modelling, data assimilation and simulation components of research papers in mathematical biology, ecology and medicine papers.
  - Use Matlab for applications to research problems.
  - Apply and implement working modelling, data assimilation and simulation methodologies they have learned in the context of a specific project.

- **Industrial, Commercial and Professional Practice**
  This course teaches a variety of transferable skills which have applications within industry, commerce and research institutes:
  - Application of theory on computational methodologies in the practical context of computer simulations.
  - Ability to represent real-world processes in various mathematical models.
  - Group-based work.
  - Written and oral presentation skills.
• Autonomy, Accountability and Working with Others
  In the second half of the course ("Modelling, Data Assimilation and Simulation Project") students will work on a particular problem (an extended project) involving modelling, data assimilation and simulation in ecology. In this activity they will learn how to work effectively as part of a research team. When subsequently writing their report students will also learn how to develop team-initiated ideas autonomously. When subsequently giving their oral presentation students will also learn presentation skills in the context of a team presentation and how to develop and present a coherent project narrative together.

• Communication, Numeracy & ICT
  This course has a very high numeracy and ICT component since a key aspect is the writing, implementation and testing of computer code to simulate data-infused models for phenomena in ecology. The final part of the course ("Modelling, Data Assimilation and Simulation Project") will teach important skills in written and verbal communication.