Project Title	Inference of ecological and environmental models
University (where student will register)	University of Warwick
Theme (Max. 2 selections)	Climate & Environmental Sustainability ☒ Organisms & Ecosystems ☒ Dynamic Earth ☐
Key words	Bayesian inference, Monte Carlo methods, Individual-based models (IBMs), fisheries management, population modelling.
Supervisory team (including institution & email address	**PI: Richard Everitt, University of Warwick, richard.everitt@warwick.ac.uk** **Co-I: Nicola Walker, Centre for Environment, Fisheries and Aquaculture Science (Cefas), nicola.walker@cefas.co.uk**

Project Highlights:

- Development of new Monte Carlo techniques for calibrating ecological models using state-of-the-art statistical techniques
- Use of machine learning and/or data assimilation methods for reducing computational cost of Monte Carlo approaches
- Application of ABC methods to ecological models with the prospect of guiding fisheries and environmental management

Overview (including 1 high quality image or figure): *Maximum 350 words*

Ecological and environmental models are continually evolving. Within fisheries research, individual movement, environmental drivers, and interspecific interactions are key areas of interest stimulating the development of new and complex modelling efforts. Individual-based models (IBMs) are one example of such model in which individual animals interact with one another and the landscape in which they live, with population metrics emerging from the actions of collective individuals. When used in spatially-explicit landscapes IBMs can show how populations are expected to change over time in response to management actions, and have therefore been shown to be effective management tools in many systems. For instance, IBMs are being used to design strategies for the conservation and exploitation of fisheries, and for assessing the effects on populations of major construction projects and novel agricultural chemicals. However, good understanding of fits to data and associated uncertainty are needed before such models can be used to support decision making.

Hence, there is urgent need to improve methods of calibrating complex multiparameter models: existing methods are too slow, and not always accurate. This project aims to improve the best existing method: Approximate Bayesian Computation, ABC. ABC is currently being used for statistical inference in a diverse range of applications in ecology, evolution and more widely, including for example: models of elephants in Amboseli; mackerel in the North East Atlantic; local butterfly populations; but also evolution of pathogens; social network analysis; and statistical physics (see Didelot et al. 2011; Prangle et al. 2016; van der Vaart et al. 2016). In most of these cases the challenges of parameter estimation and model comparison are both of importance, but implementation can prove computationally expensive. This project aims to improve ABC methods and apply them in collaboration with environmental researchers, to help them in fitting models to data. Initial focus will be on IBMs developed for fisheries management by Cefas, the UK governments marine and freshwater science experts, https://www.cefas.co.uk/.

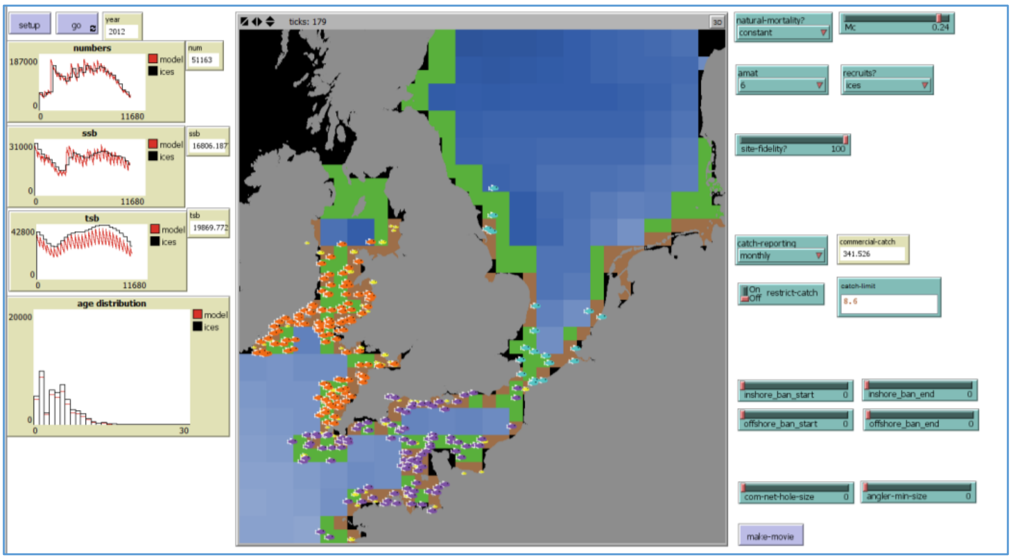

Figure 1: Screenshot of a spatially-explicit IBM for modelling a stock of sea bass.

Alt text: *Screenshot of a spatially-explicit IBM for modelling a stock of sea bass.*

Methodology: *Maximum 150 words*

ABC compares model outputs with data and is particularly useful for statistical inference where the model is only available as a computer simulator such as an IBM. ABC is a relatively new field of research, and is a hot topic in statistics and several applied fields (Beaumont 2010). There are many open problems in this area, some of which lie at the heart of this project, including:

- ABC for high-dimensional parameter spaces. IBMs often have more than 10 parameters that have to be estimated by fitting the model to data: more than in many current applications of ABC.
- ABC for computationally expensive simulators. Some IBMs take several minutes to complete a run. This is a problem because existing ABC methods require thousands of runs to obtain reliable results.

This project will develop new methods to address these issues, driven by the need for accurate ecological models to guide fisheries management.

Training and skills: *Maximum 100 words – excluding CENTA training information*

Students will be awarded CENTA2 Training Credits (CTCs) for participation in CENTA2-provided and ‘free choice’ external training. One CTC equates to 1⁄2 day session and students must accrue 100 CTCs across the three years of their PhD.

The student will spend time at Cefas learning how models are used in managing fisheries. At the University of Warwick, the student will learn to program in several languages such as R or python, and to develop new methods in mathematics and biology. They will be part of the Department of Statistics at Warwick, which contains one of the leading groups in computational statistics and machine learning in the UK. The student will become an expert in this field, aided by participating in the departmental training for new PhD students. This training includes focused study groups, and broader seminar programmes.

Partners and collaboration (including CASE):

Name of L1/L2 Partner (where applicable)	Cefas
Name of CASE partner (where applicable – project proposal **must** be accompanied by a letter of support from the CASE partner)	Cefas

Further information on partners and collaboration (including CASE): *Maximum 100 words*

Cefas is an executive agency of Defra and the UK's most diverse applied marine science centre with over 550 members of staff, including modellers, economists, biologists, chemists, physicists, social scientists, and engineers. Cefas shapes and implements marine policy through internationally renowned science and partnerships that span UK and overseas governments, non-governmental organisations and industry.

The student will collaborate with Prof. Richard Sibly, University of Reading: an expert in behavioral and physiological ecology and IBMs.

The supervisory team have been collaborating >5 years on several projects including 3 NERC SCENARIO CASE studentships. Current projects encompass IBMs of Mackerel and Bass.

Respiratory and Contact Infection Resilience of the Project: *Maximum 100 words:*

Please give concise information about how any respiratory and contact infection pandemic might potentially affect project delivery and how far mitigation efforts can be made available. Give a clear overview of those mitigation efforts, if the project or parts of it might be affected. If a shift of topic would be possible, please indicate in which direction any adjustment could be made. Please note: Project proposals without respiratory and contact infection resilience information will not be advertised.

This project is about the development, implementation, and application of statistical algorithms. As such, most of the work will involve doing either mathematics or coding. None of this work would be directly affected by a pandemic. The effect of a pandemic would be on the interaction between the team members, for planning the direction of the project and fixing problems that arise (e.g. debugging code). If required, in-person meetings could be replaced by video calls, with these taking place more often at times when the student requires more support.

Possible timeline:

Year 1: Gain familiarity with models and methods to be used in the project. Perform initial work on calibrating an IBM for sea bass using existing approaches, paying particular attention to the deficiencies of these approaches. Cefas will support training in use of the packages and models and to understand the ultimate aims for the models that we are developing. In addition, the student will be encouraged to meet and interact with other marine scientists, learn about the science policy interface, and have access to Cefas training courses.

Year 2: **Develop** new methods for calibrating complex multiparameter models. The main focus is on the following three areas:

- **Expensive simulators.** Develop methods that are computationally feasible despite the use of models that take a long time to run.
- **High dimensional ABC.** Investigate the accuracy of existing approaches to high dimensional ABC, developing improvements where necessary, particularly to enable model comparison.
- **Estimating model error.** To develop statistical methods for estimating the error in IBMs; to gain an understanding of this error, and thereby to improve the accuracy with which models are fitted to data.

Year 3: **Apply** methods developed above to IBMs developed at Cefas and the University of Reading. Assess the accuracy of their estimates of posteriors using coverage and show how the uncertainty of predictions can be described. Investigate the use of our new methods in other applications such as data assimilation, as used in weather forecasting. **Deploy** our recommended methods to the ICES (International Council for the Exploration of the Sea) secretariat and other environmental management agencies.

The student will be able to spend up to 3 months at Cefas. The student would take advantage of this when close collaboration with Cefas would be desirable.

Further reading: (*in Harvard Reference Style*)

Journal:

Beaumont (2010) “ABC in Evolution and Ecology.” *Annual Review of Ecology, Evolution, and Systematics* 41: 379–406.
Didelot et al. (2011) “Likelihood-Free Estimation of Model Evidence.” *Bayesian Analysis* 6 (1): 49–76.
Prangle et al. (2016) “A Rare Event Approach to High Dimensional ABC.” *Arxiv*.
van der Vaart et al. (2016) “Predicting How Many Animals Will Be Where: How to Build, Calibrate and Evaluate IBMs.” *Ecological Modelling* 326: 113–23.

Further details:

Contact:

Dr Richard Everitt
Department of Statistics,
University of Warwick,
Coventry, CV4 7AL
richard.everitt@warwick.ac.uk