PhD Project Proposal 4

Randomized parallel algorithms for data assimilation in numerical weather prediction.

Supervisors: Amos Lawless, Jennifer Scott (UoR), Massimo Bonavita, Nicolas Bousserez (ECMWF).

This is a joint project between the University of Reading and the European Centre for Medium-range Weather Forecasts (ECMWF). Data assimilation (DA) is the process of updating a model forecast with the latest observations to ensure that the forecasting model stays close to reality. It is a crucial component of numerical weather prediction (NWP) systems, ensuring that forecasts are initialised from the best estimate of the current state of the atmosphere. The DA algorithms currently used at ECMWF and elsewhere rely on the numerical minimization of large-scale optimization problems. However, these methods cannot be easily parallelised, which will form a computational bottleneck as NWP systems move to higher resolution. In a current MPECDT project with student Dauzickaite, Lawless and Scott have investigated the use of randomized numerical methods to precondition these optimization problems. At the same time, Bousserez at ECMWF has been investigating new optimization algorithms based on randomized methods. This project will combine this expertise to develop new randomized approaches for solving the DA problem and test them within the ECMWF system.

The first step of the project will be the development of new hybrid optimization methods, combining state-of-the-art solvers such as Lanczos with a randomized singular value decomposition (SVD). The Lanczos algorithm is sequential in nature but offers fast convergence properties due to the optimality of the underlying Krylov subspace to approximate the dominant singular vectors of the matrices. The randomized SVD methods, on the other hand, are highly parallelisable. New methods will be developed that combine the advantages of both approaches in order to find an optimal balance between parallelisation (wall-time efficiency) and accuracy (iterations) in low-rank approximation algorithms.

The second stage of the project will combine the solvers from the first step with the randomized preconditioning methods recently developed by Lawless and Scott. These will be applied to two formulations of the DA problem, strong-constraint (where the model is assumed perfect) and weak-constraint (where model error is accounted for).

The final stage of the project will be to apply randomized numerical methods to sample the uncertainty in the solution to the DA problem. The inverse Hessian of the problem gives an estimate of the error covariance matrix of the solution to the optimization problem. In order to sample this uncertainty ECMWF currently run an ensemble of data assimilations, requiring 50 separate optimizations. However, the randomized algorithm developed in the first part of this project will provide a low-rank approximation of the inverse Hessian. We will investigate the use of this approximation to replace the separate optimizations, thus developing a low-cost method to sample the uncertainty.

All stages of the project will involve theoretical developments using numerical linear algebra and numerical investigation with relevant PDE solvers. This will include a toy problem (e.g. the quasi-geostrophic model) and a realistic NWP framework using ECMWF's Integrated Forecasting System. ECMWF supervisors will join supervisory meetings on a fortnightly basis. The student will have access to the ECMWF computing software and time on the ECMWF supercomputer to perform the numerical experiments. It is expected that the student will also spend up to 3 months on placement at ECMWF.

This project would be suitable for a candidate with a strong mathematical background, and preferably some experience of programming.