Machine learning driven balance relationships for next generation data assimilation systems

Lead Supervisor: Ross Bannister, Univ of Reading, Dept of Meteorology and National Centre for Earth Observation
Email: r.n.bannister@reading.ac.uk

Co-supervisors: Christopher Thomas, Met Office; Varun Ojha, Univ of Reading, Dept of Computer Science; Hong Wei, Univ of Reading, Dept of Computer Science

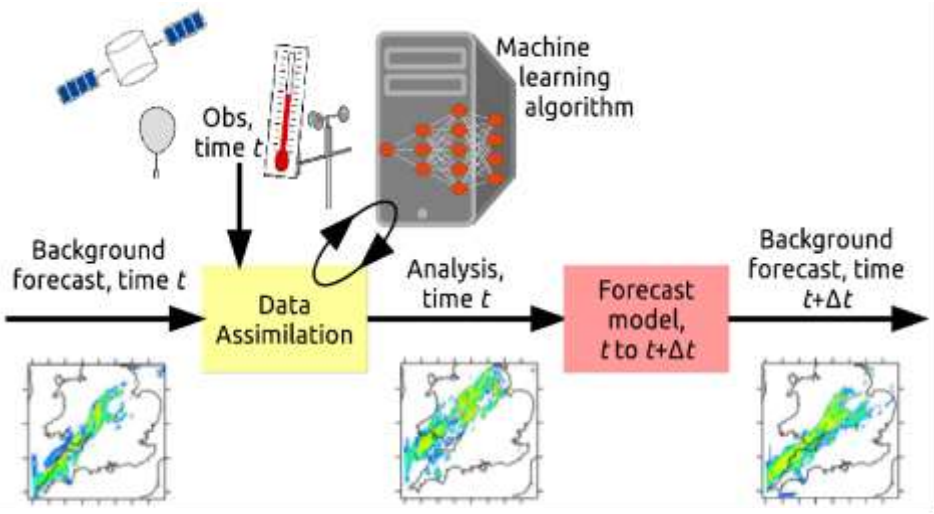
Are you fascinated by the complex models and systems that are used to produce weather forecasts? Are you a physicist/engineer/mathematician/meteorologist/computer scientist who would like to work towards a PhD with the Met Office in this important area of scientific endeavour?

The Met Office is experimenting with very high resolution models covering regional domains (e.g. a ~100m grid length model over London, UK). Models with such high resolution are right at the leading edge of weather forecasting, but are able to produce useful forecasts only if they are initialised close to the current weather and in a way that is consistent with the model dynamics. Even with modern observing systems, including radar and satellites, this process is currently far from optimal and so a lot of fundamental research is required to improve it.

Integrating dynamical models and observations is a formal process called data assimilation, which makes up a significant portion of the numerical weather prediction computation burden. Data assimilation attempts to use the latest (but limited) observational information to infer the initial conditions of the model so that useful forecasts can be made. This process projects reality (as observed) to the model's state, but this projection must not be allowed to excite inappropriate instabilities in the ensuing forecasts, so the assimilation must also be done in a way that leaves the model balanced.

Ensuring appropriately balanced forecasts requires statistical information about how different components of the model's field should be made to co-vary as the observations are assimilated. Such information is highly dependent on the flow itself and is not easily available in real time. The hypothesis of this PhD project is that machine learning techniques (like neural networks) can be used to diagnose useful covariances between different positions and variables in the model state so that unwanted instabilities can be avoided. The challenges will be to understand the covariances that should be imposed, to select and apply an appropriate machine learning method, and then to show how these covariances can impact the data assimilation and subsequent forecasts.

This is a stimulating task at the intersection of meteorology, data science, and computer science.



Data assimilation incorporating machine learning

Training opportunities:
As a student on this project, you will be supervised by and work with academics at the University of Reading and the Met Office. Training will be offered in data assimilation, machine learning, meteorology, and programming in the form of masters-level modules and summer schools. The studentship is supported by SCENARIO and the Met Office and as such you will spend some of the time working at the Met Office. The student will primarily be a member of the Department of Meteorology at the University of Reading, but also part of the Department of Computer Science, the Data Assimilation Research Centre (DARC, research.reading.ac.uk/met-darc), and the National Centre for Earth Observation (NCEO, www.nceo.ac.uk) and so will have access to facilities offered by these organisations.

Student profile:
The ideal student would have an excellent degree in physics, engineering, mathematics, meteorology, computer science, or another highly quantitative and analytical discipline. The student will be willing to learn the mechanics of data assimilation and machine learning and show creative and innovative thinking to demonstrate new knowledge in this challenging area. The project will require complex coding, including with the Met Office's new JEDI-based system (www.jcsda.org/jcsda-project-jedi).

Funding particulars:
SCENARIO is the primary sponsor of this project, but the studentship is confirmed to have CASE sponsorship from the Met Office, and support from the NCEO is also being sought.

research.reading.ac.uk/scenario