Advancing the Frontiers of Earth System Prediction: Strategic Science Plan

Introduction
The Research Programme on Advancing the Frontiers of Earth System Prediction is a £30million 15-year investment by the University of Reading, in partnership with the European Centre for Medium-Range Forecasts (ECMWF), the UK Met Office and the National Centre for Atmospheric Science. It will deliver sustained investments to tackle some of the far-term (10–15 year) and difficult (high-risk, high-reward) research challenges in global Earth System Prediction

The research strategy comprises three scientific themes and a unifying cross-cutting theme. These have been chosen, via an open consultation process, as areas that require fundamental long-term cutting-edge research (in both physical and computational science), well suited to the skills and interests of the University, while also being of great relevance to operational centres such as (but not limited to) ECMWF.

Theme 1: Predicting the Earth System up to the sub-seasonal range
Theme 2: Challenges and opportunities in simulating the Earth System at the kilometre-scale
Theme 3: Data assimilation for the Earth system across a range of scales
Cross-cutting theme: Maximising the societal benefits of extended-range numerical weather predictions.

The choice of these scientific themes has also been informed by international developments (e.g. the World Weather Research Programme, World Climate Research Programme, and EU research programmes).

Vision
The research programme, developed in close collaboration with ECMWF (and involving other key partners), will be recognised internationally for:
- Making world-leading contributions to the advancement of next-generation, extended-range weather forecasts and Earth System predictions.
- Offering a unique research and innovation hub combining meteorology, physics, mathematics and computational science.
- Offering long-term career opportunities for exceptional scientists, able to move seamlessly between fundamental and applied research.
- Researching and enabling new applications and services, in partnership with public and private sectors.

The Grand Challenge
Our challenge is to realise the latent, albeit intermittent, Earth System predictability that exists in the medium to sub-seasonal forecast range. By strategically advancing capabilities in global data assimilation, simulation and analysis, the programme will deliver a new class of accurate, reliable, and usable forecasts to society, aiming to re-define the medium-range predictability limit, from two to at least four weeks, enabling a wide range of new applications.

Science Questions

Which specific Earth System processes and signals provide the foundation for enhanced near-surface prediction skill and reliability up to the sub-seasonal range, and what chains of mechanisms can be uncovered and understood, to be effectively exploited for the benefit of society?

Figure 1, a project mind map, shows how basic process and system understanding, focussed on phenomena that emerge at a resolution of ~1km, and on signals that can be leveraged by new approaches in data assimilation, can unleash the underexplored predictability existing in the Earth System. More detailed science questions, for each theme and across themes, are included in the mind map. Key to progress are three research cycles (see Figure 2) that combine physical understanding with continued development, implementation and testing of new algorithms for data assimilation and for model simulation. Assessment based on existing and new observations, as well as analysis of a wide range of simulations and predictions already existing at ECMWF and elsewhere will enable progress to be measured and communicated.

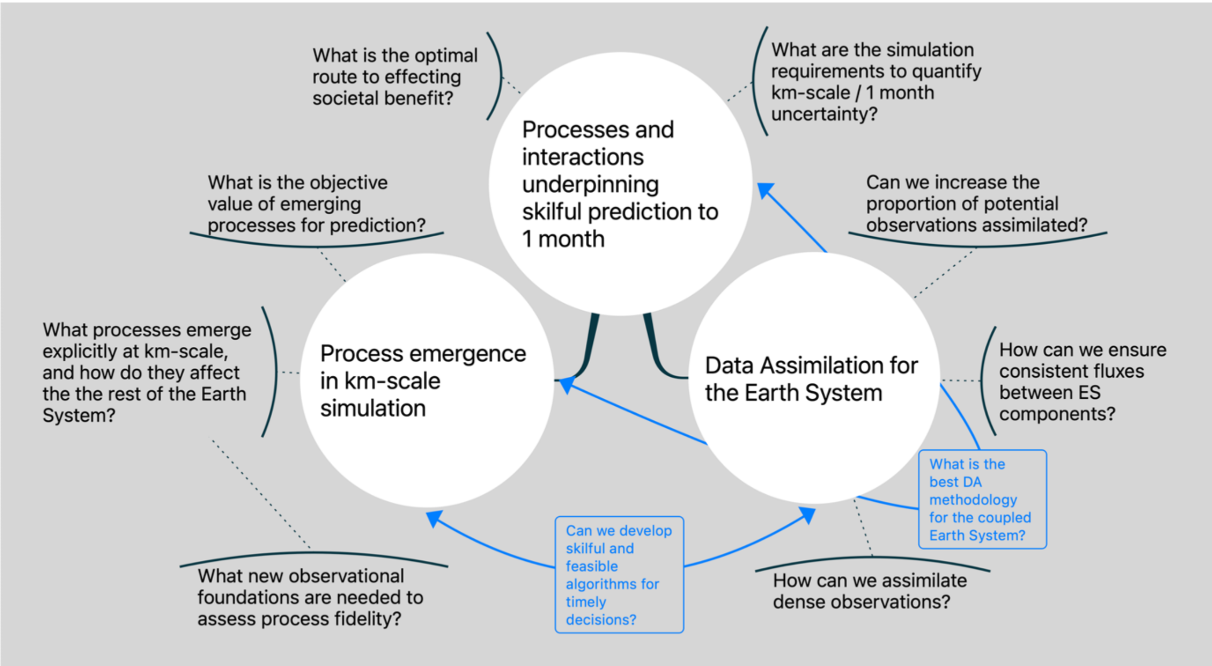

Figure 1. Mind Map of research questions around themes and their interactions

Requirements

Realising the overarching objective will require advances in many areas, including:

i. process understanding, including deep understanding of the underpinning chains of processes and interactions that shape and determine predictability.
ii. process fidelity in numerical models, providing the foundations for confidence that predictable signals are accurately reproduced in prediction systems.
iii. data assimilation methodologies, providing capability to incorporate observations into a coupled Earth system model at kilometre-scale.
iv. quantification of uncertainty, which is underpinned by an understanding of the value of each observation, and the limitations of sample size.
v. comprehensive and efficient use of computational and other resources, e.g. appropriate numerical algorithms, number of observations exploited, the resolution/complexity/ ensemble size conundrum, the lead time required for effective human decisions.
vi. optimised time and energy to solution.

vii. development and provision of (multi-variate) prognostics tailored for specific decision-makers, with clear and usable interpretation of uncertainty.

Figure 2 shows the rationale for three 5-year cycles in the project, with variable resource intensities in the areas of physical understanding, development of new algorithms and numerical methods, and work in partnership with sectoral users in order to effect societal benefits. (This illustration should be considered conceptual and indicative, not prescriptive.)

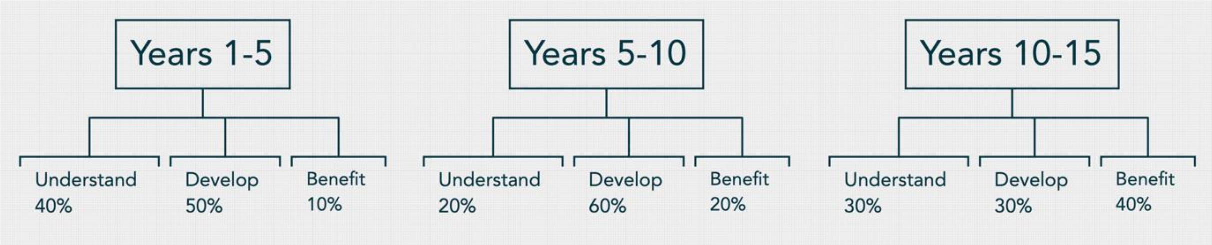

Figure 2. The three 5-year cycles in the project

Specific Thematic Challenges

Theme 1 Challenges

1A. Understanding predictability at monthly time scales

- Discover and understand the processes, and process chains, leading to predictability and uncertainty at weeks 3+4.
- Assess how well these processes are represented in current prediction systems, in order to underpin and constrain model improvement.

1B. From predictability to prediction

- Define the modelling capability that is needed to best realise the predictability that exists at weeks 3+4, and that realistically captures the related uncertainty.
- Identify improvements that are required in observations and data assimilation in order to best realise the predictability that exists at weeks 3+4, and its uncertainty.

1C. Sub-seasonal forecasting applications

- Design methodologies to best extract a forecast signal from weeks 3+4 and to best measure its value and limitations.
- Co-design best communication strategies to disseminate forecasts, their value and their uncertainties, for a robust decision-making process.

Major 15-year outcome: Significant improvements for week 3 and 4 prediction skill leading to wide use of ECMWF (and other) forecasts in decision making in a wide range of applications at these timescales.

Theme 2 Challenges

2A. Understanding Earth System processes at kilometre-scale to underpin model development

- Identify the processes that must be fully represented in globally coupled simulations at kilometre-scale, versus those that can be treated statistically (e.g. via ML), else simply downscaled.

- Understand scale interactions at kilometre-scale, including consideration of the need for 3D versus 1D representations of processes at key ES interfaces.

2B. Improving the process realism of Earth System models at kilometre-scale

- Develop methods to represent kilometre-scale processes that improve global prediction
    - Include suitable parametrisations of processes remaining at sub-grid scale, e.g. boundary layer, urban environment, land surface.
- Achieve high-fidelity representation of key scale interactions in kilometre-scale models, e.g. the impact of convection on the mesoscale, so that upscale effects are effectively included in global simulation.
- Develop efficient and accurate numerical methods, e.g. time-stepping, parametrisations (including scale-aware and ML approaches) and coupling methods.

2C. Improving the design and computational performance of global simulation and prediction systems, to enable usable and efficient predictions at kilometre-scale, suitable for exascale computers

- Achieve best value for money in terms of combined speed, portability and flexibility, to enhance productivity at a given level of fidelity
    - Measures of success to include:
        - Ability to exploit evolving forms of parallelism
        - Strongly reduced IO requirements, e.g. via online diagnostics or via data streaming (in synergy with Digital Twin approaches)
        - Development of internationally deployable workflows
        - Reduced time and cost to scientific discovery and publication.
- Develop and adopt, where beneficial, the use of computational science advances such as Machine Learning, reduced precision, stochastic physics, DSLs.

2D. Improving simulation Fidelity at a level of Reliability defined by co-design with users

- Identify the sources and growth rates of errors at kilometre-scale.
- Define the data assimilation requirements for initialisation at kilometre-scale.
- Determine the ensemble size, and design, that are required in order to realistically represent uncertainty at kilometre-scale.

Major 15-year outcome: Scientifically proven and computationally scalable global prediction systems, capable of efficiently running across a range of exascale machines, underpinning operational 1-month/1kilometre NWP at ECMWF (and others), enabling a class of user applications that leverage the latest weather and climate research, similar to that trialled in EU's Digital Twins.

Theme 3 Challenges

3A. The Path to Global kilometre-scale Data Assimilation

- Research fast, scalable ensemble 4D-Var or alternatives
    - Novel, flexible multiscale methodology to deal with nonlinear processes across a range of time and space-scales in the coupled Earth-system
    - Approaches for inhomogeneities at the land-surface.
- Reduce observation gaps, particularly for the planetary boundary layer and land- surface.
- Develop methodological approaches for assimilation of dense observation datasets, using consistent uncertainty estimates and efficient numerical techniques.

3B. Develop Earth System Coupled Data Assimilation

- Develop methods to initialise coupled ES assimilation systems, ensuring consistent flux exchanges between different components.
- Develop new and efficient observation operators that enable prediction systems to exploit the information content of observations that affect more than one ES component. Relevant examples: atmosphere-land, atmosphere-ocean, land surface-hydrology.
- Identify the most suitable coupled data assimilation techniques, and test new approaches for model adjoints inspired by deep learning.

3C. Maximising the Benefit of Observations

- Significantly increase the proportion of satellite observations used in operational DA, i.e. from the current 5-10%, to 20-40%, with a view to constrain:
    - Model state variables relevant to 1-month and/or kilometre-scale prediction
    - Model parameters (or processes) associated with model biases.

Major 15-year outcome: Computationally scalable DA, capable of efficiently running on exascale machines, suitable for 1-month/1kilometre NWP at ECMWF (and others), and leading to broader decision-making capability, due to new information content in forecasts. A new class of Earth System and Climate reanalyses are enabled by this programme and shall be delivered by synergistic projects attracted by this programme.

Cross-cutting theme: Integrating Achievements and Maximising the Benefits to Society

Each of the three themes aims to deliver world-leading advances in prediction – achieving this will immediately have large societal benefits through improved predictions and services from operational centres. However in addition, activities will be required both to communicate developments, their new benefits and limitations, and to develop novel approaches to extract the most value to users from enhanced prediction systems.

It is important to recognise that communicating and exploiting these advances will be challenging. For example, the emergence of previously underrepresented processes, despite new assimilation approaches, has the potential, at least during interim model development cycles, of increasing uncertainty, before the benefits of process realism can be fully realised.

To maximise the impact of the programme, important links across Themes must be created and sustained, aiming for an **integrated understanding** of research outcomes, particularly understanding the relative impact of the different sources of uncertainty.

Specific coordination and integration activities need therefore to be built within and across each theme in order to i) maximise the overall impact of the programme and ii) to attract wider funding from a variety of sources. These will receive technical and programme management support.

Examples of cross-cutting activities:

- In partnership with sectoral users, develop new forecast and/or observation impact verification methods that are more strongly geared towards decision-making
    - For example, in providing extended predictions of Tropical Cyclones, what is the optimal balance of DA, resolution and ensemble size when impacts are integrated in the planning.
- Develop and evaluate what new paradigms of information production/extraction can

maximise the value of kilometre-scale/1-month simulations
    - For example, identify new prognostic variables, enabled by processes emerging at kilometre-scale, and quantify their value for:
        - improving the usefulness of predictions and
        - for underpinning innovative societal services, particularly in the area of compound risk.
- Assess the added value of the new simulation/analysis capabilities in improving our understanding of weather in the context of a changing climate, and thus in consistently shaping adaptation/mitigation policies worldwide.

Major 15-year outcome: Key sectors (e.g. energy production, civil protection, agriculture, insurance) define and use specific surface-based prognostics in their decision-making suites, in synergy with similar programmes worldwide (one example is the approach in the EU's Digital Twins). Demonstrated economic benefits: (1) savings from reduced losses; (2) profits from added value services from private service providers; and (3) capture extra resource for research from wider funding opportunities.