Strategic Priorities Fund:
ExCALIBUR
Met Office Science Plan and Activities

Purpose

This document articulates the overarching vision for research commissioned by the Met Office in the SPF "Exascale Computing Algorithms and Infrastructures Benefiting UK Research" programme (ExCALIBUR). This document, aimed at stakeholders including the SPF ExCALIBUR Programme Board and Steering Committee, has been informed by consultation within the Met Office, with UKAEA, and with domain experts from the UKRI community. This Science Plan is intended to complement the ExCALIBUR activities that are funded through the UKRI Research Councils. Also, while there is some reference herein to the plans of UKAEA, the detail of that work is captured in a separate Science Plan.

The joint UKRI and PSRE ExCALIBUR programme

The over-arching aim of the SPF UK ExCALIBUR Programme is to design cutting-edge algorithms and software for a range of scientific and engineering problems that will allow scientific simulations to fully harness the power of future supercomputers. It will do this by drawing together a multidisciplinary cohort of research software engineers and scientists who will work together to future-proof the UK against the fast-moving changes in supercomputer designs. Specifically, by exploiting the fundamental commonalities in the science of multi-scale, multi-component systems, ExCALIBUR will deliver:

1. A step-change in high-performance simulation capability for the UK's priority applications (use cases).
2. An interdisciplinary framework enabling application of cross-cutting knowledge to a broader range of use cases.
3. Development of capability in disciplines with emerging requirements for high-performance algorithms.
4. The ability to efficiently manage and fully benefit from the unprecedented data flow (data science) that will accompany the high-performance simulations.
5. A new, forward-facing, interdisciplinary approach to Research Software Engineer (RSE) career development, which positions the next generation of UK software engineers at the cutting-edge of scientific supercomputing.

ExCALIBUR will achieve this by building on the four pillars:

- **Separation of Concerns**: The algorithms that encapsulate the mathematics and physics of the problem are separated from the computational science of their implementation.

- **Co-design**: Holistic design of the software of the entire simulation system involving innovative collaborations between mathematicians, domain scientists and computational scientists.[1]
- **Data Science**: Research to design new workflows adapted to managing and analysing vast volumes of data ingested and produced by simulations.
- **Investment in People**: Improved RSE career development driven by professional forward-looking approach to scientific software design of simulation codes.

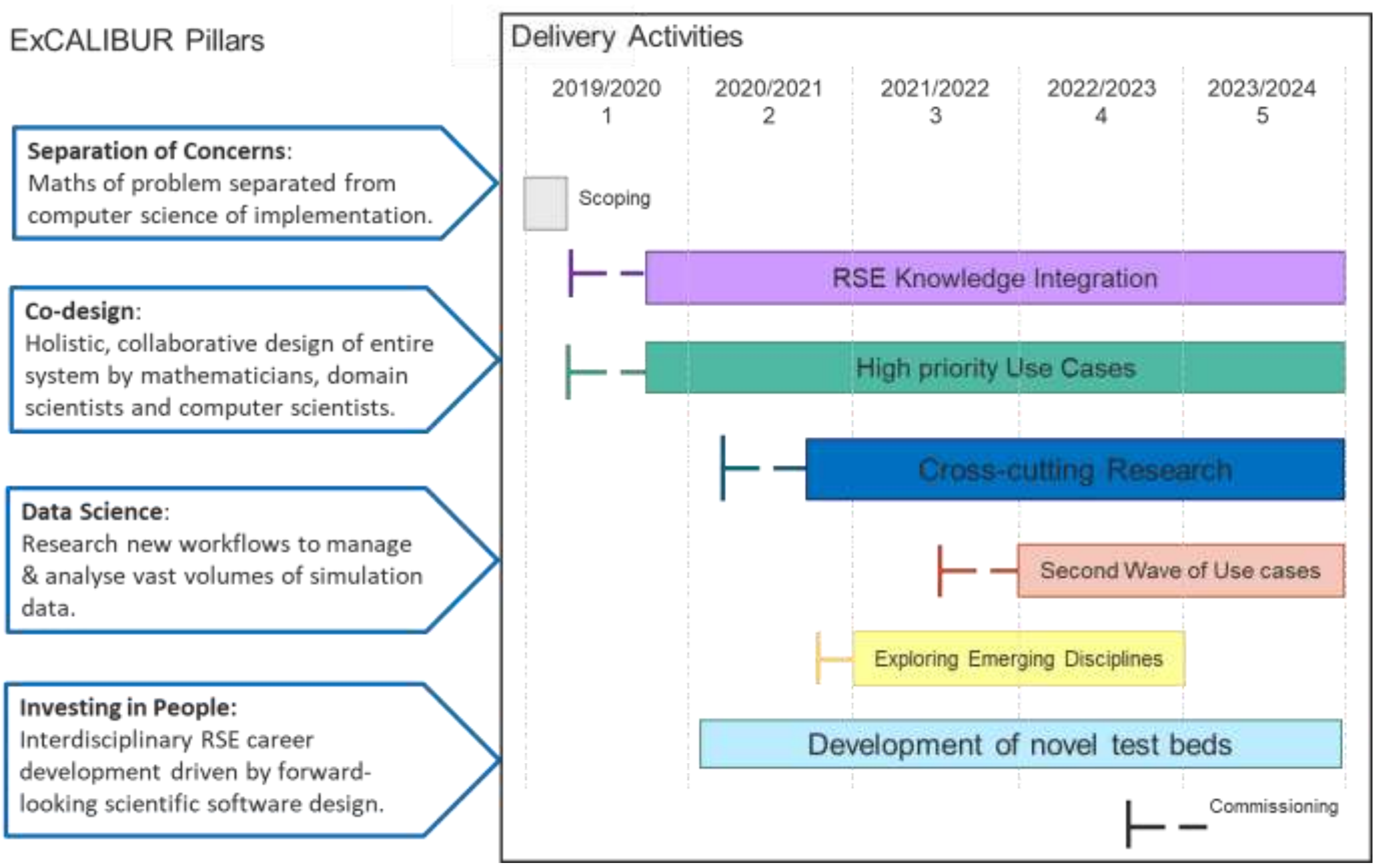

Figure 1: Outline of the ExCALIBUR programme in terms of the fundamental pillars of research and specific delivery activities.

The joint ExCALIBUR programme will be delivered through six main activities:

- Domain experts, mathematicians, and computational scientists from across the UK science base will determine a core set of simulation codes (***use cases***) chosen to span a wide range of science domains e.g. the simulation codes underpinning the benefit case studies above. Applying the principles of the four pillars, those experts will research methodologies to redesign the use cases for supercomputers of the near future and beyond.
- Knowledge will be integrated across the programme through widely applicable ***cross-cutting themes***. For example, ExCALIBUR will address the software necessary to support applications which need to store and/or access large datasets in a massively parallel environment. This will

[1] This definition of co-design does not include hardware design aspects (contrary to other common usage of the term). However, there is some opportunity for interaction with hardware designers via the novel test bed activity below. This interaction is referred to here as co-development.

involve close collaboration between industry and application domains, building on existing developments but stretching them towards exascale.
- Learning from these activities will be applied to a ***second wave of use cases*** to maximise benefits and demonstrate value.
- Exploratory research will be commissioned to identify and develop emerging high-performance algorithms (***exploring emerging disciplines***) in areas with significant potential impact, e.g. personalised medicine.
- The pressing need to advance the RSE profession will be addressed in an interdisciplinary ***RSE knowledge integration*** activity. By providing the impetus for expanding the number of RSEs and improving the two-way flow of trained individuals within and between academia and industry, the programme will support opportunities for professional development and communicate this learning more widely. Use cases will require RSE support and ensure host institutions have appropriate career paths which make RSE positions attractive and facilitate the requisite two-way flow with industry.
- Annual capital investment will permit ***development of novel test beds*** to enable co-development with industry. This provides space for innovative research that is essential in progressing future computing. It leverages further investment from industry, enables computing researchers to influence development, and prepares broader science communities to take full and immediate advantage of new commercial capability.

The pillars and activities are shown schematically in Figure 1.

The Met Office activities

The aims of the allocation managed by the Met Office are:

- To apply the principles of ExCALIBUR to deliver the benefits outlined above to the two use cases "Weather & Climate Prediction System" (shown schematically in Figure 2) and "Fusion Modelling".
- To develop and deliver cross-cutting research in themes that are aligned with the UKRI Research Council contribution.

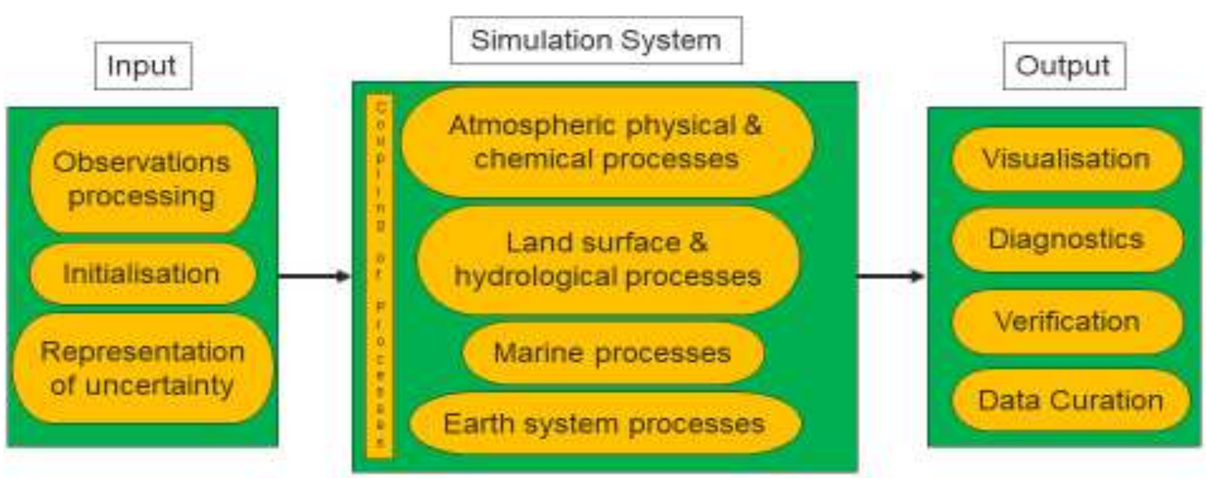

Figure 2: Schematic of the Weather & Climate Prediction system

The specifics of the cross-cutting themes have been determined by the joint ExCALIBUR community and are covered by the two over-arching themes of "Common Approaches and Solutions" and "Potential Disruptors".

The Met Office will work with the SPF ExCALIBUR Programme Board and Steering Committee to ensure alignment with UKRI funded research activities, which could include: participation in workshops and knowledge exchange activities; Met Office commissioning research aligned with research proposals in UKRI led work packages, e.g. cross-cutting themes; and participation in stakeholder engagement exercises.

As part of this Science Plan the Met Office will commission five work packages, i.e. WC-WP1/2/3 and XC-WP1/2 (see Figure 3, which for completeness includes the proposed UKAEA work packages that will be commissioned by the Met Office):

1. Component model co-design;
2. System co-design;
3. System integration;
4. Common approaches & solutions;
5. Potential Disruptors.

WC-WP1 will research the application of the principle of the separation of concerns to critical components of the weather and climate prediction system (i.e. relevant parts of the orange components shown schematically in Figure 2).

<table>
<tr><th colspan="6">PSRE Use Cases</th></tr>
<tr><td>Weather & Climate prediction system:</td><td colspan="2">WC-WP1:
Component model co-design</td><td colspan="2">WC-WP2:
System co-design</td><td>WC-WP3:
System integration</td></tr>
<tr><td>Fusion Modelling system:</td><td>FM-WP1:
Numerical representation</td><td colspan="2">FM-WP2:
Gyro-kinetic model</td><td>FM-WP3:
Neutral gas model</td><td>FM-WP4:
Code structure & coordination</td></tr>
<tr><th colspan="6">Cross-Cutting Themes</th></tr>
<tr><td colspan="2">XC-WP1:
Common approaches & solutions</td><td colspan="4">XC-WP2:
Potential Disruptors</td></tr>
</table>

Figure 3: SPF ExCALIBUR Met Office commissioned Work Packages

WC-WP2 will co-design the entire end-to-end prediction system (essentially the green aspects of Figure 2), including the redesigned components of WC-WP1 and, in time, any emerging ideas from the cross-cutting themes of XC-WP1 & 2.

WC-WP3 will implement the design of WC-WP2 using the components from WC-WP1.

Both work packages WC-WP1 and WC-WP2 will start in year 1. Work package WC-WP3 will need to closely coordinate with WC-WP1 & 2. Noting though that those work packages will not start until late in year 1, work package WC-WP3 will not start until year 2.

XC-WP1 will explore and accelerate the development of solutions and approaches to problems that are common to more than one use case and that will thereby enhance the ability of the community to better exploit exascale computing power.

XC-WP2 will explore and accelerate the development of such new, potentially disruptive, technologies that will enhance the ability of the community to better exploit exascale computing power.

Scoping work for the cross-cutting themes, represented by **XC-WP1 & 2**, will be undertaken early in year 2 so that these work packages will start later in that year.

All five work packages will involve a close working arrangement with UKRI-led work packages.

Research plans

Component model co-design (Weather & Climate prediction system, work package 1)

Weather and climate prediction systems are highly complex and comprise many different component models. These include:

- Dynamical cores that integrate the equations governing how fluids (such as the air of the atmosphere or the water of the oceans) evolve;
- Physical parametrisations which estimate the influences of those processes that are not resolved by the dynamical cores;
- Non-fluid dynamical processes such as the heating and cooling effects of radiation and the micro-physical processes that govern the formation of clouds and rainfall;
- Chemical and aerosol solvers which determine how the composition of the atmosphere evolves and how constituents are injected into, and removed from, the atmosphere;
- Land and hydrological processes which predict the fluxes of heat, water, and other constituents, between the atmosphere and the land;
- Sea-ice models which predict the evolution of ice packs;
- Wave models which determine the state of the sea surface;
- Various Earth system components such as biogeochemical solvers which determine how the composition of the oceans evolves and land ice models which predict the evolution of glaciers.

Initialisation procedures are required for many of the above components to determine the starting conditions for each of the component models from observations of the system (data assimilation). This itself needs the observations to be processed and quality controlled in a way that is consistent with the system's characteristics. Allowance also has to be made for the uncertainty in those starting conditions as well as in the model components themselves. And then the data that the system produces needs to be managed, or curated, to permit products such as forecasts to be visualised, verified, post-processed for products, and archived.

Not all of the above elements and components of the system need to be completely redesigned to remain performant into the future. Some, those for which the processes are, by nature, local processes, are already relatively agnostic to the architecture of future supercomputers; some are relatively cheap and so do not need to exploit the full power of future supercomputers (and might be amenable to the flexibility of architecture that cloud computing offers).

Therefore, this work package will focus effort on the larger, most vulnerable components. It will also seek to exploit commonalities of design in existing families of models. The foci of research in this work package could include:

- The atmospheric dynamical core.
- The physical parametrisations and non-fluid dynamical processes.
- The chemical and aerosol solvers.
- The NEMO family of codes (the ocean model, ocean initialisation, and the sea-ice model).
- The wave model.
- The atmospheric observation processing and data assimilation.
- The verification and post-processing systems.
- The coupler used to couple together different executables.

With domain experts working with computational scientists and algorithmic experts, the research will explore:

- Co-design: Whether (and if so how) the existing algorithms need to be redesigned to make those algorithms more scalable (so that the cost of the algorithm increases linearly with any increase in the number of degrees of freedom). This will require exposing as much as possible those elements of a process that do not depend on other elements (removing unnecessary serial operations and replacing them with parallel ones) and implementing more inherently local algorithms.
- Separation of concerns: How to ensure that any changes needed to apply the algorithm optimally on alternative supercomputer architectures can be implemented in an automated manner. This requires the application of a Domain Specific Language (DSL). The desire is to have a DSL that is generic enough to encompass the requirements of this use case whilst remaining as specific as possible. The research will determine how achievable that desire is and, if it is not, then how limited can its relaxation be made (e.g. what is the fewest number or types of DSLs needed? How branched does one DSL need to be?).

System co-design (Weather & Climate prediction system, work package 2)

Work package 1 focuses on ensuring that the component models of the weather and climate prediction system will be performant on future generations of supercomputers. This work package will focus on research to ensure that when those component models are assembled (see Work Package 3 below) the system as a whole is performant.

A key element of this work package will therefore be the strategy for how to couple different components together: how and when should that be done; and how should the coupler itself be designed to be optimal for future architectures.

Many of the principles that apply to the individual components apply to the system as a whole. It will be important to ensure that any serial operations are absolutely necessary and that all other operations can be applied in parallel (thus allowing for as much task-parallelism as possible). Any communication between different components needs to be minimal which places constraints on how the system is designed and in particular how diagnostics and data output are handled. Since the relative costs of disk access versus the overhead of passing memory around is unclear (and possibly architecture dependent) it will be essential to have a system that is flexible in this regard. Similar considerations apply to the archiving of data from the system and to access of that data; it is inefficient if users have to manually optimise their code for different types of hardware; therefore a separation of concerns between the data and its storage medium should be researched. This might be characterised as ‘data location aware work flows’.

Many of these aspects are not specific to this use case; other use cases are likely to face the same challenges with the same solutions. Such aspects would be likely candidates for the cross-cutting themes (e.g. XC Work package 1). This work package will therefore focus on redesigning those elements that are specific to this use case. Such elements could include the design and development of:

- The specific system (the data workflow) for this use case, in particular one that addresses any challenges presented by the use of the automatic code generation required by the DSL(s) and also the data workflow challenges presented by the ensemble system and cycling data assimilation.
- An optimal strategy for the coupling of the component models together that minimises the data dependencies between components.
- An optimal diagnostic system, and associated visualisation system, that retains flexibility and a good user interface while optimally manipulating data across vast numbers of processors and possibly combining data from different components.
- As few DSLs as possible while covering as many components of work package 1 as practicable.

System integration (Weather & Climate prediction system, work package 3)

This work package will bring together all the elements of work packages 1 and 2 to deliver a working weather and climate prediction system that is ready for the supercomputers of the mid-2020s. As well as the complex testing and trialling of the new system, an important element of this work package will be to stage the integration as much as possible so that where possible components of the existing system can be replaced in an incremental way. This approach will enable the benefits of the ExCALIBUR programme to start to be realised as soon as possible and ahead of its end date. Where that is not possible it will be important to devise a testing strategy (perhaps by developing appropriate sub-systems) to minimise the risk when various new components are finally brought together.

Common approaches and solutions (Cross-cutting themes, work package 1)

The issues addressed by the Weather and Climate Use Case work packages WC-WP1, 2 & 3 are specific to the weather and climate prediction system. However, it is anticipated that there are a significant number of issues that are common to two or more of the use cases considered by ExCALIBUR, i.e. “Common solutions and approaches”. It is therefore more efficient for these cross-cutting themes to be explored more generally than explicitly as part of a use case.

The activities of this work package will explore and accelerate the development of solutions and approaches to problems that are common to more than one use case and that will thereby enhance the ability of the community to better exploit exascale computing power.

These themes have been identified by the ExCALIBUR programme as a whole. This work package is for the Met Office commissioned contribution to those “Common solutions and approaches”.
The identified themes are:

- Input/output (IO), storage & file format: How to efficiently get exabytes of data into and out of the simulation system and into and out of whatever storage medium is used to archive that data. This would include the aspect of data location aware workflows.
- Data workflow: Having got the data into the simulation system, the passage of that data through the system needs to be managed efficiently, particularly when large ensembles are used to simulate the

uncertainty. Approaches to diagnostics and coupling between components will be key elements of this theme.

- Visualisation: How to effectively manipulate exabytes of data to visualise an appropriate level of detail from the simulations.
- Mixed-precision: Most physical simulation codes use 64-bit arithmetic throughout their calculations. Appropriate use of lower precision arithmetic can speed up calculations and, more importantly, make communication between processors more efficient. To make optimal use of this technique will require research into how, in general, to make algorithms less precision dependent.
- Fault tolerance: Although the failure rate for processors is small, when hundreds of thousands of those processors are used, the likelihood of failure during any given simulation can become significant. It is important therefore to develop approaches for handling such failure in as efficient a manner as possible.
- Couplers: Most large simulation systems need to couple different models together. It is therefore appropriate to explore the generic issues around how to optimally achieve that coupling between models running on large and different numbers of processors.
- Separation of concerns and Domain Specific Languages: By building on the experience from the use cases it is anticipated that a number of common themes around how to achieve the separation of concerns can be identified and then codified. This will then facilitate application of this approach to the second wave of use cases.

For each of these themes, the implementation is likely to be case specific (and might therefore be managed within the use case work packages); this work package is more concerned about exploring generic questions and solutions. In part this will be done by learning from experiences across the ExCALIBUR community, and in part it will require fundamental research.

Potential Disruptors (Cross-cutting themes, work package 2)

The issues addressed by the Weather and Climate Use Case work packages WC-WP1, 2 & 3 are specific to the weather and climate prediction system. However, it is anticipated that there are new, potentially disruptive, technologies whose benefit and application will be common to two or more of the use cases considered by ExCALIBUR. It is therefore more efficient for these cross-cutting themes to be explored more generally than explicitly as part of a use case.

The activities of this work package will explore, and accelerate the development of such new, potentially disruptive, technologies that will enhance the ability of the community to better exploit exascale computing power.

These themes have been identified by the ExCALIBUR programme as a whole. This work package is the Met Office commissioned contribution to those technologies.
The identified themes are:

- The use of machine learning to emulate appropriate physical processes, thereby optimising simulation time while retaining a physically based approach.
- Research into methods that expose further levels of parallelism in the simulated systems (e.g. parallel-in-time methods).
- The use of containers, object stores and other tools that will help workflows migrate to cloud environments.

- Future supercomputing paradigms such as exploration of how quantum computing can be applied to the simulated systems.
- Verification, Validation and Uncertainty Quantification in the context of exascale simulations.
- HPC horizon scanning to ensure that the programme remains aware of and is agile in its reaction to developments in supercomputer architectures.

Monitoring, Evaluation and Learning

Benefits of the programme will be defined in terms of:

- Increased science **productivity**.
- Transformational change in **capability**.
- Increased **skills** base of interdisciplinary Research Software Engineers (RSEs).

Measures of the success of the programme are expected to include:

- Improvements in the speed, resolution and/or complexity of the first and second waves of use-case simulation systems (productivity).
- New frontiers reached by those simulation systems that were previously unachievable (capability).
- The number of scientific publications produced from ExCALIBUR research, including details of collaborations and, as time progresses, citation of that body of research (skills).

Due to the recognised skills shortage in this area, knowledge integration is an important element of the programme and the ***RSE knowledge integration*** activity will run for the duration of the programme. This will support interdisciplinary working, capture lessons learned, and communicate findings to the wider community. Measures will need to be developed to capture the:

- Expanded number of RSEs working in the ExCALIBUR disciplines.
- Improved two-way flow of trained individuals within and between academia, PSREs and potentially also industry.
- Enhanced and increased uptake of opportunities for professional development.