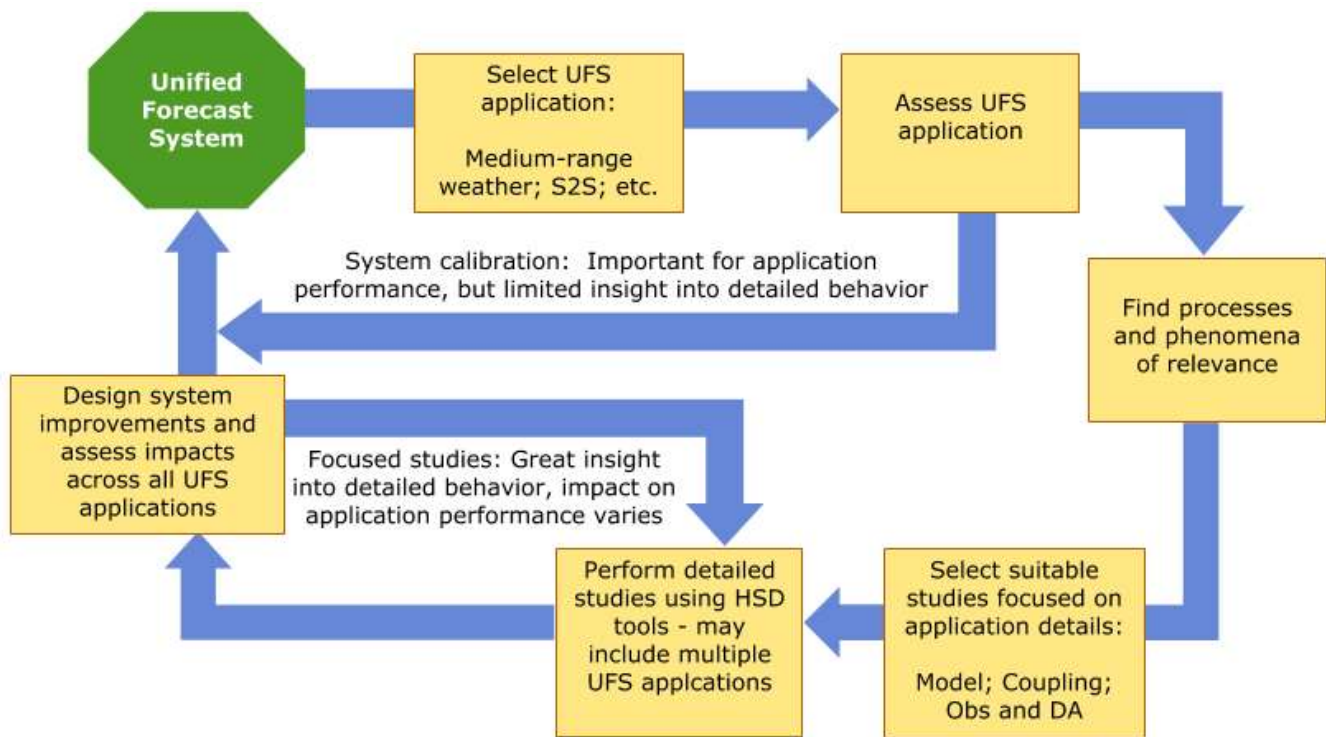


HIERARCHICAL SYSTEM DEVELOPMENT FOR THE UFS



HSD in the UFS Development Process (Figure adapted from Christian Jakob, BAMS 2010)

September 10, 2019

Michael Ek¹, Cecelia DeLuca², Ligia Bernardet², Tara Jensen¹, Mariana Vertenstein¹, Arun Chawla³, James Kinter⁵, Richard Rood⁴

¹National Center for Atmospheric Research

²NOAA Global Systems Division/Cooperative Institute for Research in the Environmental Sciences

³NOAA Environmental Modeling Center

⁴University of Michigan

⁵George Mason University

What is Hierarchical System Development?

Hierarchical system development (HSD) refers to the ability to engage in development and testing at multiple levels of complex prediction software such as the Unified Forecast System (UFS). It is critical for research because it enables the research community to have multiple entry points into development that reflect their interests. Those interests may include aspects of specific system components such as atmospheric physics, ocean and ice dynamics, or data assimilation for land models and other earth system components. Research interests may also include more integrative topics such as coupled model development or strongly coupled data assimilation. Research efforts such as the Community Earth System Model already have established many features of hierarchical development for modeling

HSD is likewise critical for operations because both localized and integrative, coupled processes must be improved to develop excellence in forecasts. For example, there is evidence that improved atmospheric microphysics and atmospheric coupling to ocean, land, and ice components, respectively, are necessary to achieve the best medium-range weather forecasts.

Research-to-operations-to-research (R2O2R) refers to the interactions between these communities that are needed to support their different goals and ways of working. The UFS strategy for fostering positive interactions is through participatory governance and shared infrastructure. In this kind of community approach, priorities can be developed and communicated, the research community can access and test aspects of the operational system, and the needed scientific and technical advances from research can be transitioned into operations. HSD is at the heart of R2O2R because it ensures that every advance proposed by the research community has a clear development and test path that includes integration into end-to-end applications.

What are the functions of HSD, and what are some design considerations?

Some of the basic functions that an HSD for UFS must perform are:

- Ability to test atmospheric physics using a single column model (SCM). A SCM is easy to use and requires few resources to run. It is an important feature for supporting UFS participants who are engaging in development as part of a student project, or who may not have access to supercomputer facilities. Part of the SCM is the ability to turn on/off given physics parameterizations or sets of parameterizations in order to test them in isolation for better understanding of processes and to more easily identify systematic biases at the process level.
- Ability to turn model component feedbacks off using data components. For applications where there are multiple coupled components (e.g. atmosphere, ocean, ice, ionosphere, etc.), complexity is managed through controlled experimentation. A critical part of this is the ability to turn off feedbacks by replacing a prognostic component (one that is taking inputs, evolving in time, and providing updated information) with a data component that simply provides a constant realistic data set as its output.
- Ability to use model component configurations that have simplified scientific algorithms. An example of this is running in “aquaplanet” mode for a coupled system, in which Earth has no land mass. It is also routine to use simplified “slab” ocean models as part of development and testing. These sorts of substitutions enable scientists to better understand the interactions among components and the causes of the results they are seeing.
- Ability to run a model with pre-assimilated initial conditions and cycling, but no data assimilation. Running a system with data assimilation involves preparing and processing observations, and supporting more complicated workflows where the results of one run are used in the next. In this mode, the effects of using data assimilation are simulated. This may involve generation of ensembles as well.
- Ability to run a quasi- or full operational workflow, including cycling with data assimilation. If a researcher wants to see exactly how their change affects an end-to-end operational application, running the full operational workflow is desirable. This has been difficult in the past because of the resources required and the complexity of the end-to-end applications. It may be that the Earth Prediction Innovation Center (EPIC) project, which plans to extend available resources using cloud computing, is able to provide this.

Many of these HSD capabilities exist independently, but must be integrated together in a holistic way in order to achieve an HSD that accelerates the R2O2R process for UFS collaborators. The HSD must include appropriate metrics and benchmarks of technical and scientific performance for UFS applications, their subcomponents, and the processes within these subcomponents, plus the necessary forcing data sets. These can be obtained from observations/field programs and model output, and may need to be idealized for “stress-testing” model parameterizations and other code. It may be the case that slightly different sets of metrics are needed for each aspect of the HSD.

Testing within the HSD is expected to be concurrent and iterative, i.e. more complex steps can provide information “backwards” to be used at simpler HSD steps, and vice versa. This also includes understanding the spatial and temporal scale dependencies in UFS, and the need for consistency in solutions between higher-resolution/regional short-range applications versus global applications for the medium- and extended-range, S2S and longer climate time scales, e.g. “scale-aware” physics. The figure above shows what development processes associated with HSD typically entail.

Prerequisites, current work and next steps for a plan forward

A key prerequisite for evolving HSD is development of a long-term strategic plan for UFS that spans its constituent applications and includes federal, academic, private and international community partners. This will enable the UFS community to define projects that address agreed-upon goals, and will lay a foundation for R2O2R success. The Strategic Implementation Plan (SIP) for the UFS, collaboratively developed with many of these partners, extends out about three years and is a start at this kind of planning.

The SIP includes several activities that could be entrained into a comprehensive HSD framework. One of these is the atmospheric SCM being developed by the Global Model Testbed (GMTB), which is based on a Common Community Physics Package (CCPP). Another is delivery of a flexible workflow called the Common Infrastructure for Modeling Earth (CIME). CIME was developed by the National Center for Atmospheric Research in collaboration with DOE. It includes a simple way to run a multi-component Earth system model, along with a set of data components. Additionally, the Joint Effort for Data assimilation Integration (JEDI) is making it easier to develop and run different data assimilation algorithms and more complicated workflows. Lastly, the extended Model Evaluation Tools (METplus) verification, validation and diagnostic system needs to integrally tie into each aspect of the HSD. A next step in this effort is to define an overall project and begin to define how the various infrastructure and science elements will be integrated to realize the HSD vision.