Singletrack: Data Assimilation Requirements to Support Application/Roadmaps

Four science roadmaps support the benefits of a new atmospheric modeling capability at NCAR. All four roadmaps explicitly require forecasting capabilities for the atmosphere. Any state-of-the-art research or applications of atmospheric forecasting require a modern data assimilation (DA) capability and the Geospace and Chemistry roadmaps explicitly note this. DA and resulting ensemble forecasts are a key tool for evaluating predictability, uncertainty, and for sensitivity analysis that can be used to diagnose model deficiencies and estimate appropriate parameter values. All roadmaps include model development and tuning activities, the Weather roadmap explicitly requires ensemble forecasting, and the Hydrology roadmap explores predictability. These elements of the roadmap motivate the following requirements analysis for DA-related capabilities in the atmospheric model. State-of-the-art atmospheric DA for the Hydrology, Chemistry and Weather roadmaps will also require the effective assimilation of radiance observations from a variety of satellite-borne instruments, but this gap is not addressed here.

<u>Model forecasting capabilities</u>:
An atmospheric model to support the science would require the following forecasting capabilities, independent of the details of the DA methodology:

- Ability to generate forecasts as part of a suite of earth systems models including (but not limited to) global/regional ocean, land surface, chemical constituents (MICM), ice, ionosphere,
- Configurable to be efficient for range of space/depth/time scales including limited-area,
- Useful forecast skill for research across wide range of space/time scales including global to cloud scale, and troposphere to exobase (500km). A capability for large eddy simulations is also *desirable*,
- Support for stochastic forecasts including stochastic physics and noise addition.

<u>Design characteristics to support DA</u>:
There are a number of required design characteristics of the atmospheric model that are specific to its efficient use by the DA system, independent of the details of the DA methodology:

- Precise, accessible definition of model's prognostics state,
- Easy to invoke range of damping and smoothing to stabilize model in presence of DA increments,
- Ability to access and modify model/physics parameters,

The following *highly-desirable* capabilities would provide additional DA flexibility and reduce computational cost for DA:

- Easy-to-use interface to compute/output diagnostic quantities,
- Computation of forward operators efficiently at high frequency.

These additional capabilities for DA could be implemented outside of the atmospheric model, however, it would be *desirable* to include them in the model design:

- Incremental analysis update capability,

- Digital filter initialization.

<u>Model computational characteristics for efficient DA</u>:
The atmospheric model must have the following computational characteristics to support research demands using DA and ensemble forecasts, independent of the details of the DA methodology:

- Sequence of short integrations has minimal computational overhead relative to single long integration,
- Efficient/scalable ensemble forecasts (including large numbers of tracers) on available high performance computing,
- Can stop and restart exactly.

The following highly-desirable capability would facilitate exploration of parameter sensitivity and provide access to some hybrid ensemble DA methods:

- Concurrently executed ensemble forecasts can have different resolution and physical configurations.

<u>Support for 4-dimensional variational assimilation (4DVar)</u>:
Today, the best global atmospheric forecast systems employ 4DVar for DA. The model adjoint used in the 4DVar optimization also has applications in computing the sensitivity of forecast quantities to model initial conditions. On the other hand, a variety of other DA techniques, from variational schemes to purely ensemble approaches, now perform nearly as well as 4DVar. Thus, it would be *highly desirable* but not essential for the model development to provide:

1. Tangent linear and adjoint versions of the nonlinear model.

If they were included in the model, it would also be *desirable* to include:

2. Tangent linear and adjoint capabilities for model computed forward operators.

<u>Support for DA applications</u>:
It is important the DA capabilities be viewed as an integral part of the model system to support ongoing science advances. DA places particularly stringent demands on the usability, flexibility and robustness of the model code: typical experiments may easily require tens of thousands of short model simulations and, because diagnosing DA results often depend on understanding deficiencies of the forecast model, involve reconfiguring the model grid spacing, number of levels, and physical parameterizations. Thus, DA requires:

- Ongoing support for DA users,
- A stable interface, updated infrequently, for configuring and running the model.

In addition, it is *highly desirable* that

- DA users should be consulted before support is discontinued for legacy versions of model.

<u>Note on the use of observations</u>:

Many of the applications in the roadmaps will require assimilation of specific types of new observations. The development of forward operators for these observations and the development and testing of a useful assimilation system are non-trivial tasks and will require dedicated resources. Particular challenges are associated with the use of remote sensing observations, especially orbital radiometers, that can have complicated observational error characteristics.