WRF Reference Configurations – DTC Concept and Plans

Jamie Wolff[1%*], Beth Weekley[1%], Louisa Nance[1%], Ligia Bernardet[2&%], Barbara Brown[1%]

[1]National Center for Atmospheric Research, Boulder, CO
[2]NOAA Earth System Research Laboratory, Boulder, CO

I. Overview

DTC

The Developmental Testbed Center (DTC) provides a link between the development and operational communities in order to more efficiently transfer new technologies in numerical weather prediction (NWP) from research to operations. Extensive testing and evaluation must be performed to ensure that these new techniques are indeed ready for operational consideration. At the DTC, NWP research and operational communities interact to test and evaluate new developments in the Weather and Research Forecasting (WRF) model, for both research applications and operational implementation (Bernardet et al. 2008).

WRF

The WRF model receives contributions of new methods and features from both the operational and research communities. Development of the WRF system is led by NCAR, NOAA/NCEP/EMC, and NOAA/ESRL with partnerships at USAF/AFWA, USA/ARL, FAA, USN/NRL and USN/OON (see Appendix A for all acronym expansions), as well as collaborations with universities and other government agencies in the U.S. and overseas. In addition, through contributions to the WRF repository, a wide range of users from various arenas have the ability to share their new developments with the community, creating the potential for the most qualified techniques to advance into operations. The WRF system also provides a common framework that simplifies the implementation and testing of these new methods and capabilities.

The current WRF model has two dynamic cores; the Advanced Research WRF (ARW) and the Nonhydrostatic Mesoscale Model (NMM). In addition to dynamic core, several other elements of the WRF model are user configurable. These elements include the domain (i.e., location, resolution, and nesting), physical parameterizations, analysis nudging, and observation nudging, to name a few. Ultimately, the resulting model forecasts will vary depending on how the model is configured by the user.

[&]Contract with Systems Research Group, Inc. (SRG), Colorado Springs, CO
[%]Also affiliated with the Developmental Testbed Center, Boulder, CO

*Corresponding author address: Jamie Wolff, NCAR/RAL, P.O. Box 3000, Boulder, CO 80307, email: jwolff@ucar.edu

II. WRF Code Repository

The WRF code repository consists of community developed code from which different configurations, including operational ones, can be defined, as shown in Figure 1. These components are described in more detail in this section.

a. WRF Repository

The code contained in the WRF repository is developed by members of the WRF community as a whole, including the research community and operational centers. Newly developed methods or enhancements to existing methods offered as contributions to the WRF system must address a potential need or make improvements over current NWP techniques. Before code can be checked in to the WRF repository, it must be written in the WRF Software Framework, must meet WRF coding conventions and standards, and pass designated testing. More information about code contribution standards and procedures can be found at: http://www.mmm.ucar.edu/wrf/docs/contrib_info.pdf. Although the WRF repository is housed at NCAR/CGD and maintained by NCAR/MMM along with the DTC, the code contributions to the WRF repository are generally maintained and documented by the author.

Proposed changes offered by a developer to the WRF repository are reviewed at regular (e.g., bi-weekly or weekly) meetings of the WRF Developers' Committee. Once changes are approved, and before they are checked into the WRF code repository, each bundle of proposed changes must pass a basic regression test. The basic regression test consists of building the code without optimization for several different machine configurations (including serial, OpenMP and MPI). Each dynamic core is run out 10 time steps for a standard set of physics configurations. Regression tests are performed to ensure the results complete the 10 time-step runs without producing NaNs (pass/fail test), and that each output variable is identical ("bit-for-bit") when compared to a similar run with a different number of processors (current tests consist of runs using the serial build on 1 processor versus the OpenMP and MPI builds on 4 processors). When all of the qualified changes for the week have been checked into the repository, a final regression test is performed on an extensive set of runs (e.g., fully optimized, quilting, I/O in GRIB format). Bit-for-bit checks are again performed for each output variable of the runs using different numbers of processors, excluding the full optimization runs which are only examined for completion of 10 time steps without producing NaNs. Currently, regression tests performed as part of the pre and post check-in process do not examine computational performance (the time required to generate a forecast) or accuracy of the forecast.

b. Operational Configuration

An Operational Configuration is a stable, reliable, and robust configuration of the model that is being run operationally at one of the operational forecast centers (e.g., NCEP or AFWA). Each Operational Configuration is extensively tested for computational performance on its target hardware and for forecast accuracy. While developmental testing and evaluation of WRF codes may be done by partnering organizations, final responsibility for testing and acceptance of Operational Configurations lies with the operational prediction centers. The current Operational Configurations are maintained by the respective operational center.

III. Reference Configuration

a. Purpose

Currently, with the possible exception of Operational Configurations, there are no published baselines or standard statistical results for well-tested configurations of the WRF model. Moreover, there are no sets of tested WRF configurations that can help guide the selection of configurations to be considered for the next generation Operational Configurations. To create the infrastructure to address these and possibly other issues, the concept of WRF Reference Configurations has been proposed by the DTC.

Reference Configurations will serve both the operational and research communities. By conducting carefully controlled, rigorous testing, including the generation of verification statistics, Reference Configurations will provide the operational community guidance for selecting configurations with potential value for operational implementation. Reference Configurations will provide the research community with baselines against which the impacts of new techniques can be evaluated. Statistical results for a Reference Configuration may also aid researchers in selecting a configuration to use for their projects. Reference Configurations are anticipated to be limited in number, based on resources and the level interest of the modeling community, and will be maintained by periodic retesting as the WRF system evolves. Thus, the extensive tests performed on the Reference Configurations will provide the entire user community with specific information regarding whether the configurations remain robust and efficient, and have improved forecast skill over previous versions. The goal is to continually evolve the Reference Configurations as new NWP techniques are integrated into the WRF system and to sustain an optimal mix of testing between past Reference Configurations (to maintain statistics from previous years) and new Reference Configurations (which represent the latest promising developments).

b. Defining Reference Configurations

A Reference Configuration is defined as a particular collection of configuration files that are set at compile time (i.e., *configure.wrf*, dynamic core, Registry) and run-time (e.g., *namelist.input*). Each Reference Configuration will be uniquely defined and detailed in a design document that will be distributed to the user community. The details of this definition will then allow for a user to exactly replicate the results produced. The (set of) Reference Configuration(s) (to be determined) will be extensively tested, documented, and the baseline data and statistical verification results will be made available to the community.

The DTC will solicit feedback from operational centers and the research community on which configurations to establish and maintain. These suggestions will be obtained from the community at large during the WRF Users Workshop and NWP conferences, from the DTC Advisory Board and the DTC Executive Council. Operational Configurations obtained directly from operational centers will make up a subset of the Reference Configurations. Figure 1 shows how Reference Configurations fit with the WRF repository and Operational Configurations.

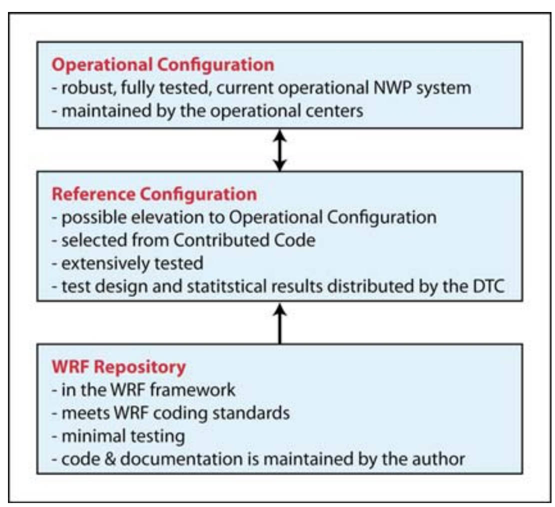

Figure 1: WRF Code and Configurations

Striking a balance between keeping a traceable history of results from year to year and testing new features and capabilities in NWP will be critical in considering which configurations will be tested. Ultimately, the candidate configurations will be selected based on their potential use and value added to the research and operational communities. If a Reference Configuration becomes out of date with current technology in the field of NWP, it may be moved to a "retired" status and replaced by another configuration shown to be more relevant. A prioritized list of candidate configurations will be drafted by the DTC and presented to the DTC Advisory Board for review and the DTC Executive Committee for final approval.

The path for a particular configuration in the WRF repository (in this case including Operational Code) to become a Reference Configuration is illustrated in Figure 2. After the prioritized list of candidate configurations is established, extensive testing and evaluation is performed to designate Reference Configurations, with results distributed to the user community.

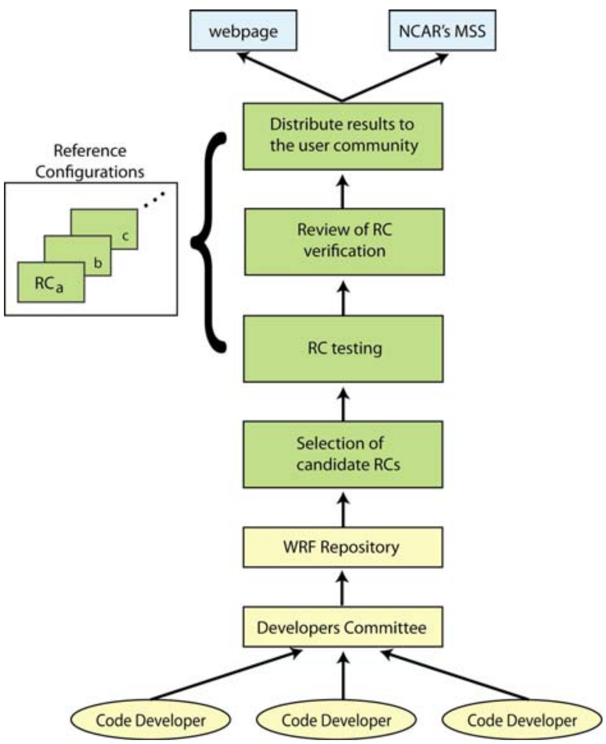

Figure 2: Path to Reference Configuration

c. Reference Configuration Testing

Using the prioritized list of candidate configurations established by the DTC Advisory Board, the DTC will test each configuration based on the resources available. This test will aid in providing a baseline for a particular Reference Configuration, so that future tests can verify timing and performance have not been adversely affected by code changes. The tests will be retrospective in nature and encompass a large number of cycles. The cycles will be selected in order to encapsulate the broad range of weather regimes from null, to weak or strong events. The exact periods and domains chosen will vary depending on the configuration being tested and the type of phenomenon that the configuration is designed to forecast. For some configurations, these cycles may be from all four seasons (e.g., extratropical domains for general predictions) while for others the cycles might come from a particular season (e.g., hurricane season, convective season).

Periodically, NCAR releases new versions from the WRF code repository, labeled as WRF release *n*, encompassing many individual changes for which regression testing has been performed. The first time a Reference Configuration is tested for WRF release *n*, a comprehensive set of cycles will be run and an extensive set of verification statistics will be computed. From this full set of cycles, a subset of cycles will also be evaluated in the same manner. The definition of this subset of cycles will depend on the number of cycles needed to obtain representative estimates of the verification statistics. Subsequent testing of a Reference Configuration following future releases of WRF will be done on the subset of cycles only, unless additional cycles are needed to further evaluate the performance of the configuration. This procedure will maintain a strong connection between current and prior testing, while conserving resources.

Before testing begins on a new Reference Configuration, a design document will be created detailing the specifications of the testing to be done. In addition to the details of the configuration of the model (compile-time and run-time settings), this document will also include specific information on the domain size and location, grid spacing, geophysical fields, forecast length and cycles being run for the comprehensive set of cycles, as well as the subset of cycles to be run during subsequent testing periods. The document will also identify post-processing and verification methodologies.

Since the WRF code repository changes on almost a weekly basis and new WRF releases typically occur annually, the schedule of Reference Configuration testing must be designed accordingly. A diagram showing when Reference Configuration testing is expected to occur relative to the evolution of the WRF code repository and WRF code releases is shown in Figure 3. The horizontal line represents the WRF repository changing over time. The vertical hash marks represent WRF code repository updates; the green hash marks represent check-ins related to on-going development of WRF repository, while the red hash marks depict bug-fix check-ins during the period when the repository is frozen for pre-release testing.

For results to be most relevant to the user community, Reference Configuration testing (i.e., RC_a, RC_b in Figure 3) will be based on official releases of the WRF code with the inclusion of any bug fixes that have been posted before the retrospective testing begins. However, the DTC reserves the right to serve the needs of its sponsor agencies by testing at any time, a newly contributed technique (a "tagged" version from the WRF code repository) not yet available in the official release (i.e., RC_c in Figure 3).

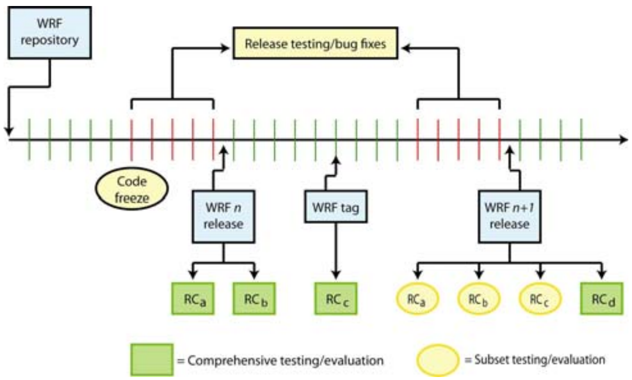

Figure 3: DTC Reference Configuration testing relative to the evolution of the WRF code repository and releases.

d. Verification

Once a Reference Configuration has been run for the specified set of cycles, extensive verification statistics will be computed. These statistics will be clearly identified in the design document and explained in the report on the Reference Configuration evaluation. Verification approaches may include, but are not limited to, current standard verification statistics in wide use by the NWP community (e.g., RMSE, bias, ETS). In addition, new verification techniques relevant for the particular configuration and its intended applications may be applied in the evaluation. Confidence intervals will be provided as appropriate. A range of variables and levels will be evaluated, including surface and upper-air predictions for temperature, humidity and wind, as well as precipitation. Additional quantities may be verified as appropriate for particular applications of a Reference Configuration (e.g., cloud ceiling and visibility for an aviation-forecast configuration). If the Reference Configuration has been tested for a prior WRF release, results from current testing will also be compared to previous results.

e. Dissemination of Results

Once the verification results have been completed, the results will be distributed to the user community via a webpage and all the input and output datasets will be archived on NCAR's Mass Storage System (MSS) for at least one year. The webpage will provide instructions on how to set up a particular Reference Configuration, including where to find the raw data and configuration files on NCAR's MSS. It will contain verification statistics from each Reference Configuration for each testing period and, thus, provide a traceable history to the user community as the model code changes.

IV. Summary

The DTC has detailed the new concept and plans for defining and testing WRF Reference Configurations. Statistical verification of a variety of configurations will serve both the operational and research communities by evaluating the impacts of a new technique and helping to guide the selection of the next generation Operational Configurations from these.

Details regarding the most effective forum for obtaining proposals for candidate Reference Configurations from the community at large are still under discussion. Suggestions from the user community on this issue, as well as other aspects of this concept are encouraged.

Acknowledgments: The authors would like to thank the following people who participated in developing the concept of Reference Configurations through earlier discussions: Tom Henderson and Steven Koch of NOAA ESRL, Laurie Carson, Robert Gall, Dave Gill, Lacey Holland, Joe Klemp, and Jordan Powers of NCAR, Nelson Seaman of NOAA, Steven Rugg of AFWA, Geoff DiMego, Steve Lord, and Naomi Surgi of NOAA NCEP, and Jack Kain of NSSL.

V. References

Bernardet, L., L. Nance, M. Demirtas, S. Koch, E. Szoke, T. Folwer, A. Lought, J. L. Mahoney, H.-Y. Chuang, M. Pyle, and R. Gall, 2008: The Developmental Testbed Center and its Winter Forecasting Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 611-627.

Appendix A: List of Acronyms

AFWA	Air Force Weather Agency
ARL	Army Research Laboratory
ARW	Advanced Research WRF
DTC	Developmental Testbed Center
CGD	Climate and Global Dynamics Division
EMC	Environmental Modeling Center
ESRL	Earth Systems Research Laboratory
FAA	Federal Aviation Administration
MMM	Mesoscale and Microscale Meteorology Division
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic & Atmospheric Administration
NMM	Nonhydrostatic Mesoscale Model
NRL	Naval Research Laboratory
NSSL	National Severe Storms Laboratory
NWP	Numerical weather prediction
OON	Office of the Oceanographer of the Navy
USA	United States Army
USAF	United States Air Force
USN	United States Navy
WRF	Weather Research and Forecasting