Why is there more than one dynamical core in WRF?
A technical perspective

William C. Skamarock
National Center for Atmospheric Research
Boulder, Colorado 80301
skamaroc@ucar.edu, 303-497-8161

9 September 2004; revised 7 June 2005

With the release of the ARW core in the WRF framework, with the porting of the NCEP NMM model into the WRF framework and its designation as a WRF core, with the ongoing discussions concerning how the Navy's COAMPS model might become a WRF core, and with our early technical investigations into dynamical core possibilities for a nonhydrostatic global WRF model, it is an appropriate time to reconsider the feasibility and desirability of the original goal of the WRF project – that of having a single WRF dynamical core within a single common modeling system. The collaborative development and use of a single common system by both research and operations would provide major benefits to both communities, including streamlining the incorporation of research advancements into operations, the focusing and leveraging of resources and scientists into the development of the shared modeling system, and perhaps most important, the development of collaborations and ties between the research and operational communities.

Neither the goal of having a single WRF core nor the obvious benefits have been realized. Additionally, the disadvantages of having multiple cores are becoming obvious, particularly the significant added expense of developing and supporting multiple cores and the lack of collaboration that is the result of groups working on their own cores and system components. These problems have had a major negative impact on community efforts to move the high-resolution NWP enterprise forward, and the negative impact is likely to increase given the parallel move to multiple data assimilation systems and, potentially, multiple frameworks. At the same time, operational results are suggesting that there is little difference between the performance of the dynamical cores, rather the significant differences between the WRF model core forecasts are more the result of different physics and not dynamical core designs. Given these two realities, instead of considering new cores, we should reconsider our original intent – a single system with a single core.

Presently there are two WRF dynamical cores in the WRF framework, the NMM and ARW cores. Given their existing formulations, these two cores cannot be merged – the two cores differ fundamentally in a few crucial aspects. The arguments and evidence supporting the choice of these different approaches indicate, however, that there is little of substance that prevents us from unifying our cores. Rather, our inability to unify the cores has arisen, on the technical side at least, from our not having collaboratively examined and resolved the crucial issues.

Given the major benefits of using a single shared system, we need to consider anew the original goals of the WRF collaboration. *Thus, I propose that we (the WRF operational and research community) collaboratively examine the technical issues that underlie the differences in the existing WRF cores, with the intent of determining, if possible, a single core formulation that would satisfy the needs of both operations and research.* This examination could take place in a small workshop (like the WRF planning meetings) in the near future. If consensus and compromise were to lead to an acceptable formulation, it is my belief that it could be quickly developed and implemented in a collaboration of the WRF developers. Towards this end, herein I briefly review the general features of the two WRF dynamical cores and I provide a brief outline of how the cores evolved to their current states. The crucial differences between the cores are outlined and a path forward is presented.

Core overview

An atmospheric dynamical core can be concisely defined by a relatively small number of attributes: form of the governing equations, prognostic variables, horizontal grid, vertical grid, terrain formulation, time integration method, spatial integration method. The NMM and ARW cores possess the following attributes:

1. Governing equations formulation

    NMM – advective form for the momentum and thermodynamic equations with a mass continuity equation.

    ARW – flux form for the momentum and thermodynamic equations with a mass continuity equation.

2. Prognostic variables

    NMM – velocities, temperature, pressure, and column mass (hydrostatic surface pressure).

    ARW – momentum (mass coupled velocities), dry entropy (mass coupled potential temperature), geopotential, and column mass (hydrostatic surface pressure).

3. Horizontal grid

    NMM – E grid

    ARW – C grid

4. Vertical grid

    NMM – Lorenz grid

    ARW – Lorenz grid

5. Terrain formulation

    NMM – terrain-following sigma coordinate that relaxes to a pressure coordinate at an intermediate level below the model top.

    ARW – terrain-following sigma coordinate.

6. Time integration method

   NMM – Eulerian formulation. Forward-backward scheme for inertia-gravity waves, 2nd order Adams-Bashforth for horizontal advection and Coriolis terms, Crank-Nicholson scheme for vertical advection. Acoustic modes are integrated explicitly for horizontal propagation (using the model timestep), and implicitly for vertical propagation.

   ARW – Eulerian formulation. Forward-backward for inertia-gravity waves and horizontally propagating acoustic modes (timesplit, using an explicit acoustic small timestep), 3rd order Runge-Kutta scheme for all other terms. Vertically propagating acoustic modes are integrated with an implicit scheme.

7. Spatial discretization

   NMM – The horizontal discretization follows the enstrophy and energy conservation principles of Arakawa. The vertical discretization is centered and second order. Mass and moist scalar conservation is maintained in the discrete formulation. Horizontal divergence damping is used. The full thermodynamic variables are used for pressure-gradient calculations.

   ARW – The horizontal discretization of the advection terms uses a 5th order upwind biased scheme. The mass-divergence and pressure gradient formulations are centered and 2nd order accurate. Vertical advection typically uses a third order upwind biased scheme. The discretization is in conservative form, with conservation of the first-order quantities – momentum, dry entropy, mass and moisture. Horizontal divergence damping is not used. The perturbation thermodynamic variables are used for pressure-gradient calculations.

A Little History

Both the NMM and ARW models represent advancements of earlier model formulations. These advances, taken as a whole, indicate that the formulations of the NMM and ARW models are moving closer together rather than further apart. Before considering the differences apparent in the previously listed model attributes, a little history is in order.

The NMM core

The nonhydrostatic NMM model evolved directly from the hydrostatic Eta model. Outside of the inclusion of the nonhydrostatic terms into the governing equations, the primary differences between the two are the use of a terrain-following coordinate in the NMM versus the step-mountain coordinate in Eta, and the removal of the time-split gravity wave integration from the NMM that is used in the Eta model. Additionally, the NMM uses a 2nd order Adams-Bashforth time integration scheme as opposed to the first order scheme used in the Eta model.

The configuration of the NMM has also evolved from that used for the Eta model. Specifically, the dissipation formulation and physics in the NMM have been modified and re-tuned to enhance and better retain the small scale structures in the forecasts. These small-scale structures are almost entirely missing from Eta forecasts and forecasts from early configurations of the NMM.

The primary features shared between the NMM and Eta models are the use of the E grid and the use of the energy and enstrophy conserving finite differencing.

The ARW core

The ARW core has its roots in the Klemp-Wilhelmson (KW) cloud model developed in the mid-late seventies. The ARW uses the KW time-splitting of the acoustic and gravity waves. It differs from the KW model in most other aspects; the ARW core is cast in conservative form and the discrete model conserves mass, momentum, dry entropy and scalars; it uses a 3rd order Runge-Kutta time integration scheme for the slow-mode integration; it uses a hydrostatic pressure (mass) vertical coordinate; it uses high-order upwind biased flux-divergence (advection) operators.

The KW formulation has been used as the basis for a number of other models, including MM5, COAMPS, ARPS, and the DWD-LM. These model formulations are very close in design to the original KW model. For example, all use an advective (non-conservative) formulation, and all use the KW leapfrog time-split integration method.

Core issues

Similarities

1. First and foremost The NMM and ARW cores share the hydrostatic pressure (mass) vertical coordinate – the NMM core retained it from its Eta model ancestry while the ARW model developed a new timesplit integration scheme for it.
2. Both models conserve mass exactly - to machine roundoff.
3. Both use terrain-following coordinates. The difference in the formulations (the NMM relaxes to constant pressure at an intermediate level whereas ARW uses the sigma coordinate throughout the atmosphere) is not significant because the NMM formulation is easily included in the ARW core (a similar formulation is presently being considered).
4. Both models use the forward-backward scheme for inertia gravity waves and horizontally propagating acoustic modes, and both use a vertically implicit formulation for vertically propagating acoustic modes. The details of the schemes differ because of the different variables and horizontal grid structures, but the philosophies behind their implementation and use are the same.
5. Both use the Lorenz grid for vertical discretization.

Differences

Listed in order of importance with respect to core unification, the differences between the NMM and ARW cores are:

1. Grid staggering, the NMM uses the E grid and the ARW uses the C grid.
2. Choice of equations, variables and conservation properties: The NMM uses pressure

and temperature so as to facilitate an energy and enstrophy conserving discretization, while the ARW core uses potential temperature to facilitate conservation of first order quantities (momentum, dry entropy).

3. Time-integration methods: The NMM is a fully explicit model (except for the treatment of vertically propagating sound waves and vertical advection). The ARW core is a split-explicit model – acoustic and gravity waves are integrated explicitly with a small timestep (using the forward-backward technique) while slow modes are integrated with a large timestep (using 3rd order Runge-Kutta).

Resolution of the Critical Differences

From a practical standpoint, the different grid staggerings represents the largest obstacle to unifying the cores because the grid staggerings permeate all aspects of the dynamics in the cores and in model pre- and post-processors. The technical arguments supporting the different choices are relatively simple. The E grid (and the related B grid) is well suited to large-scale rotational modes because the horizontal velocities are co-located and the pressure and mass are staggered in such a way as to preclude the existence of a non-zero null mode in the pressure solution (the non-zero null mode is the primary problem with the unstaggered A grid). In contrast, the C grid is designed to provide the best resolution for divergent modes (pressure-gradient – divergence computations). The velocities are not co-located on the C grid, and this leads to less accuracy for the rotational modes. The choice between the E grid or C grid rests on evaluations of the relative importance of the strengths and weaknesses of the two grids with respect to the rotational and divergent modes. A more complete explication of the trade-offs needs to be considered, and this would include both a complete outline of the theory (most of which already exists scattered throughout the literature), practical examples – in both simple and full models, and implications on code complexity.

The choice of variables and many aspects of the discretizations rest largely upon the choice of conservation properties for the dynamical core. While mass conservation is an obviously desirable and perhaps necessary component in a state-of-the-art research and NWP system, the need for exact conservation of other first or second order quantities is not easily discerned. An additional complication is that the nonhydrostatic models typically do not exactly conserve quantities other than mass or scalars. For example, the ARW core does not exactly conserve momentum because of the acoustic mode treatment; the NMM does not exactly conserve enstrophy or energy (although it should be noted that the governing equations themselves do not conserve enstrophy). As with the grid staggering, the trade-offs between the different conservative and nonconservative formulations needs to be considered, using as guides both theory and practical examples overlaid with the intended uses for the core.

The time integration methods are perhaps the most straightforward issue to address if the previous issues are resolved. Multiple time integration approaches could exist in a given core if needed; the ARW core supports both a second and third order Runge-Kutta time integration, in addition to a leapfrog time integration for the dry/no-physics dynamics.

Other differences between the NMM and ARW cores are somewhat less important to the central issue of core unification.

The Path Forward

As stated at the outset, *I propose that we (the WRF operational and research community) collaboratively examine the technical issues that underlie the differences in the existing WRF cores, with the intent of determining, if possible, a single core formulation that would satisfy the needs of both operations and research.* The issues most in need of consideration and resolution are

1. the horizontal grid staggering, and
2. the need for conservation of first and/or second order quantities (i.e., the choice of variables), and the discretization approaches dictated by these choices.

These issues should be examined in a small workshop that should be preceded by the production of a detailed outline of the trade-offs of the various approaches. Such an outline would essentially be a expanded version of parts of this paper and could be compiled using input from the WRF community. As noted in several points, analyses and test results addressing some of these issues already exist in the literature or as unpublished results, although they are scattered. Thus the outline could be put together quickly. A period of comment on the outline, along with suggestions for further analyses or tests, would allow workshop participants to arrive having considered the material and ready for substantive discussion. I believe that this pre-workshop preparation would take no longer than a few months, and the workshop could follow soon thereafter.

The workshop should focus on the technical issues only, with the goal of producing a report on these issues and on the possibilities for a single core. The success of a workshop would depend on participation by all members of the WRF project.

If agreement on the formulation for a single core can be reached, then the process for constructing the core in a fully collaborative manner, and its subsequent use and support, can be considered. If agreement cannot be reached, then the technical issues and arguments will be clear to all.