The I-WRF Framework: Containerized Weather Modeling, Validation, and Verification

Richard Knepper
Sara C. Pryor
Bennett Wineholt
Xin Zhou
rich.knepper@cornell.edu
sp2279@cornell.edu
bmw39@cornell.edu
xin.zhou@cornell.edu
Cornell University
Ithaca, New York, USA

Melissa Bukovsky
Jared Lee
bukovsky@ucar.edu
jaredlee@ucar.edu
National Center for Atmospheric Research
Boulder, Colorado, USA

ABSTRACT

We describe the I-WRF project for the NSF Cyberinfrastructure for Sustained Scientific Innovation program, which provides a framework for application containers that allow the Weather Research and Forecasting Model (WRF) software and accompanying MET and METplus validation software to be run on a wide range of resources with minimal installation requirements. I-WRF will support three major science use cases that quantify impacts of environmental or human developments together with climate change on critical outcomes. I-WRF is also intended to facilitate outreach by making it easier to provide training and demonstrations to build understanding and interest for new potential atmospheric scientists.

CCS CONCEPTS

• **Applied computing → Earth and atmospheric sciences**; • **Computing methodologies → Modeling and simulation**; • **Information systems → Computing platforms**; • **Computer systems organization → Cloud computing**.

KEYWORDS

weather modeling, application containers, orchestration

ACM Reference Format:

Richard Knepper, Sara C. Pryor, Bennett Wineholt, Xin Zhou, Melissa Bukovsky, and Jared Lee. 2023. The I-WRF Framework: Containerized Weather Modeling, Validation, and Verification. In *Practice and Experience in Advanced Research Computing (PEARC '23), July 23–27, 2023, Portland, OR, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3569951.3597547

PEARC '23, July 23–27, 2023, Portland, OR, USA

ACM ISBN 978-1-4503-9985-2/23/07…$15.00
https://doi.org/10.1145/3569951.3597547

1 INTRODUCTION

In this paper, we introduce I-WRF [7], a framework to support the development and deployment of an application container suite that supports the Weather Research and Forecasting Model (WRF)[5], Model Evaluation Tools (MET)[2], and METPlus, developed by the National Center for Atmospheric Research (NCAR), National Oceanic and Atmospheric Administration, the U.S. Air Force, the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration. I-WRF is being developed to support the broad portability of weather modeling software, making deployment of the software easier for researchers and students, reducing the reliance on local IT staff, and providing the capability to demonstrate the benefits of weather forecasting with WRF. The I-WRF project started in September of 2022[1] and is supported through the NSF Cyberinfrastructure for Sustained Scientific Innovation program, with partners at Cornell University Center for Advanced Computing (CAC) and NCAR. Below, we describe the motivations for creating WRF, MET, and METplus containers in detail; followed by a description of the sample science use cases that will be supported by this work. We describe the initial steps to improve upon existing solutions and lay out the future activities that will be supported by I-WRF.

2 MOTIVATIONS OF THE I-WRF PROJECT

The "WRF software" is used by more than 50,000 users worldwide and provides facilities for weather prediction based on observations or conditions provided by the researchers. WRF is used for operational weather forecasting as well as the simulation of weather outcomes based on potential inputs. WRF has been extended with a considerable number of additional models that provide capabilities for additional conditions, including atmospheric chemistry, fire, hydrological, solar, and urban models. The WRF software in its general sense includes a complex set of interacting components, including the pre-processing system, model extensions, post-processing, visualization, and verification components, with multiple data sources that include observations, terrestrial conditions, and gridded data systems, and output file formats. This makes the installation and maintenance of WRF on shared systems a complex responsibility for cyberinfrastructure professionals. For new users intending to try out WRF with the resources at hand, or a current user that wants to extend their use of WRF to cloud resources or a new HPC environment, the complexity of the software requires considerable experience with compiling and managing software, and the ability to analyze generated outputs with confidence.

These barriers to access make the considerable benefits of WRF software accessible only to researchers who invest considerable amounts of time into learning how to manage this complex software or who have local expertise that can manage this for them. Researchers at institutions with limited resources face a considerable barrier to adopting and successfully using WRF software, and students are unlikely to encounter WRF until they are fairly far into career progression, severely limiting the utility of WRF as a demonstration of the impact of atmospheric analysis and specifically the flexibility of WRF modules to forecast and model conditions for weather events that have significant outcomes in a local context.

3 I-WRF CAPABILITIES

The I-WRF framework is intended to address some of these barriers to adoption by providing a set of application containers that will encompass base use cases for researchers and allow WRF execution on multiple environments with a minimum requirement for configuration and compilation. An application container that utilizes Docker and Singularity capabilities removes the requirement to compile and install software for each system, although different execution environments require changes in storage and interfaces that can be delivered as part of an orchestration system or by layers within the application container. I-WRF will allow the orchestration of multi-node application containers for large-scale models with a common namespace and output and will provide verification and visualization capabilities through MET and METplus.

Based on the execution/orchestration configurations provided, I-WRF will provide a portable software installation that is easier to initiate, use, and maintain than the standard software distribution. This simpler environment will be extensible by varying execution scripts that provide alternative execution scales, data sets, and namespaces. As the project continues, the project team intends to evaluate WRF-CHEM and WRF-Smoke for inclusion into the framework to extend the research capabilities towards broader data sets and areas of inquiry.

The concurrent development of MET/METplus containers will allow the evaluation of outputs from WRF analyses to ensure that results from the WRF container conform to expectations. The planned evaluation container will also include some facilities for visualization of results so that a researcher can execute quickly and see a relatively good approximation of the simulation that can be exported to other tools for further manipulation and improvement. This also provides the facility to demonstrate the capability of short-term weather modeling to new students and show the potential for the simulation of weather elements with real-world outcomes.

4 I-WRF SCIENCE USE CASES

The I-WRF project team selected three science use cases to demonstrate the effective capabilities of a framework with weather simulations that investigate phenomena whose outcomes can impact U.S. citizens. These use cases examine weather events across a changing landscape affected by agriculture, development, and growth; variance in renewable energy generation due to "droughts" of productive conditions; and how climate change affects the improvements to air quality created by recent regulations. By pairing the new containerized application framework with real-world weather simulation applications, the project team will demonstrate the viability of the containerized workflow for complex simulations, likewise the intricacies of actual simulation problems will pull forward the functionality of the container suite so that it supports detailed sets of science requirements.

4.1 Use Case 1: Land Use/Land Cover

Recent simulations have demonstrated that the structures on and usage of the landscape affects weather patterns and the sensitivity of human infrastructure to changes in the extremity of weather events[3]. The Land Use/Land Cover model will create high-resolution simulations of areas that include buildings, parking lots, and agricultural usage of land areas. In order to provide high-resolution simulation, a "storyline" approach will be used that identifies historically important weather events. Those important events are simulated at high-resolution using the WRF models as well as other initial and lateral boundary conditions. The resulting simulation is examined for fidelity and if it meets requirements is run again with extrapolated conditions inferred from changes in building coverage, land use, or general climate to understand the impact of new development and changes in climate. The use case seeks to answer questions about how areas with different characteristics exhibit different sensitivity to extreme weather events, how the usage of land affects the development of extreme weather, and how social vulnerability to climate change is impacted by ongoing changes in land use and coverage. The following analysis describes a LULC simulation used to characterized the needs of the application for the I-WRF framework.

To characterize a typical implementation of Use Case 1 prior to incorporation into the containerized framework, Pryor and Zhou selected two historically important events with deep convection based on analyses of RADAR and other remote sensing data. The first case is March 28, 2017 (0328) and is a well-organized squall line under strong synoptic forcing with a south-north organized line of deep convection that moved over DFW from west to east. The second case is July 03, 2018 (0703) and represents more locally forced convection and comprises a line of thunderstorms that advected over DFW moving from northwest to southeast. The simulations run for 60 hours and nested grids are used that have a grid spacing (dx) of 9 km (d01), 3 km(d02) and 1 km (d03) to ensure the urban region is highly resolved (Figure 1).

To investigate the urban influence on these two representative deep convection events, Pryor and Zhou performed simulations using current day LULC and then perturb the surface conditions such that we increase DFW to twice (DFWx2) and four times (DFWx4) of its original size while keeping the morphology of the urban expanse and avoiding urbanization of water bodies. Specifically, the WRF "met_em" files are modified before creating initial and boundary conditions. The land use category (LU_INDEX) and land use fraction (LANDUSEF) in the peripheral areas of DFW are replaced by the urban counterparts while maintaining the lake grids. Figure 1

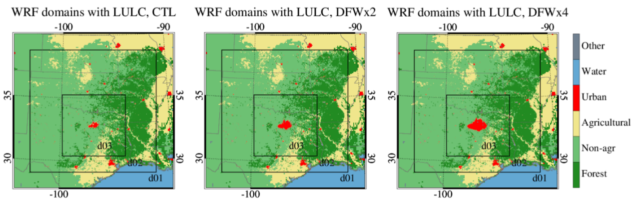

Figure 1: WRF simulation domains and land use land category (LULC) used to simulate deep convection. The urban land use category is shown in red and the panels from the left to right show LULC in the control run (CTL) with the original size of DFW, plus perturbed runs; DFWx2 and DFWx4.

shows the WRF domain with land use land categories (LULC) for the control (CTL), DFWx2, and DFWx4 simulations.

An example of the urban influence on one aspect of deep convection; rainfall rate (RR, i.e. the accumulation of water at the ground expressed in depth units per time interval) is shown in Figure 2 for a 3-hour period after the line of convective cells passed over DFW. The left hand panels show the mean rainfall rate in this 3 hour period in the CTL simulations. The right hand panels show the difference in RR from the LULC perturbation experiments. As shown, for the event with regional forcing (0328), the urban environment has a highly heterogeneous impact on RR, and the impact is small when spatially averaged. However, with weaker atmospheric forcing (0703), the change in RR is much more spatially coherent and scales with the size of DFW both in terms of magnitude of the change in RR and the spatial extent of the perturbation.

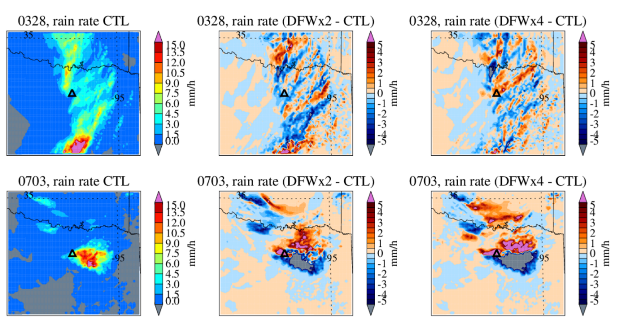

Figure 2: Rain rate (mm/hr) in domain d03 (note centroid of DFW is shown by the triangle) in the CTL run (1st column) and the differences in rain rate in the DFWx2 run (2nd column) and DFWx4 run. The rain rate and perturbations are calculated as the mean during the 3-hours after the line of convection passed over DFW.

This science is preliminary to containerization but represents the type of analysis that would be applicable under the I-WRF framework.

4.2 Use Case 2: Climate Change Impacts on Wind and Solar Resources

As the electricity supply continues to decarbonize and renewable resources are added to the overall generation capacity across grids, weather dependency of energy generation becomes increasingly important, and variability of generation has the potential to increase due to climate change[4]. The primary investigation in this use case is to quantify variability of generation to identify frequency and quantity of production droughts and excesses that can impact grid-wide generation capabilities. To identify these highs and lows of energy generation, the continental US weather patterns will be simulated based on a set of initial and lateral boundary conditions over a 40-year domain. This analysis will improve understanding of the impacts of changes on these renewable energy resources and provide useful information for grid system planners to integrate carbon-free sources of electricity into power production systems.

4.3 Use Case 3: Air Quality in the Northeast Urban Corridor

With the introduction of the Clean Air Act, air quality in the U.S. has improved dramatically, however changes in atmospheric dynamics and emissions linked to global climate change may substantially erode these improvements[6]. This problem includes both air movement as well as particulate matter in the air. The project team will containerize the more complex WRF-Chemistry module to simulate historically important periods of poor air quality under current conditions as well as possible future climate conditions, including land use/land cover conditions. The computational demands of WRF-Chem are roughly twice that of standard WRF. The technical requirements for containerizing and validating the chemistry model will represent the most challenging aspect of the project. The analysis result will be the quantification of the degree to which climate change will offset the improvements in air quality due to emission controls.

5 I-WRF ARCHITECTURE OVERVIEW

The overall I-WRF framework as envisioned operates as follows: researchers identify the I-WRF container suite through the project website, a research software portal, or a container repository; and identify a resource that will run the I-WRF container to model desired factors. The resource can be a personal system proximal to the researcher, such as a laptop; public cloud resources, national resources such as the Jetstream2 research cloud, or campus clouds that support container runtimes; or HPC resources running Singularity. The container will access Initial Conditions and Lateral Boundary Conditions desired for the model, meteorological, observational, or terrestrial data that support specific conditions to be modeled, and eventually chemistry input. The researcher will make use of execution scripts (Terraform, Ansible) created by the project team that will facilitate the desired scale of resources to be used, supporting modeling on a single proximal resource, on a single node with many processors, or multiple nodes for extremely large scale modeling.

The I-WRF container will include post-processing scripts that will permit the output of WRF to be ingested by a second container which provides METplus in order to validate the WRF output. Validation is critical to ensure that model results are conformat to the specifications of the area of inquiry within the model. A final container will be able to receive validated WRF output and provide facilities for visualization and further analysis. Once the researcher is satisfied with the model fidelity, data can be exported to local storage. An overview of the I-WRF framework is provided in Figure 3.

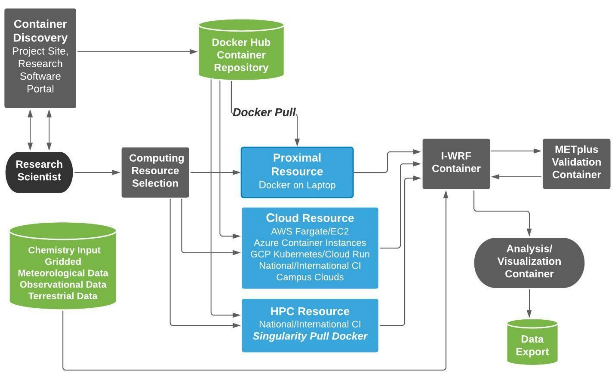

Figure 3: I-WRF top-level overview

5.1 I-WRF Framework Distribution and Execution

The I-WRF containers will be made available through the project website (https://www.i-wrf.org), the NCAR WRF GitHub site, and the Docker Hub Registry. The project team will work to make the containers and associated execution scripts discoverable and available through resources such as the ACCESS Research Software Portal (https://software.xsede.org).

The initial I-WRF containers to support science use cases will be run on the NCAR resources Cheyenne and later Derecho. These will utilize Singularity for running Docker containers within an HPC context and will be supported with Slurm execution scripts that will manage the number and scale of WRF containers. The project team will also deploy containers on the Cornell region of the Jetstream2 research cloud, as well as the Red Cloud private cloud at the Cornell CAC. The project team will also prepare execution scripts for running I-WRF containers on the Stampede2 resource at Texas Advanced Computing Center (TACC).

The I-WRF project team consists of members of both NCAR teams in the Mesoscale and Microscale Meteorology Laboratory, members of CAC, and Cornell Earth and Atmospheric Sciences faculty (Pryor) as well as a post-doctoral researcher. The NCAR team will provide support on WRF compilation and configuration for the WRF, MET/METplus, and visualization containers, as well as post-processing scripts which allow for data to pass between the containers without modification. The CAC team will create execution scripts and orchestration that allow the containers to run on the targeted execution environments described above, as well as testing and integration suites that will prepare the application containers themselves for distribution. Project scientists (Pryor and Sue Ellen Haupt at NCAR) will provide the required WRF namespaces and initial and lateral boundary conditions for analyses, feedback on the container fidelity and veracity as well as the ease of use of execution, and the post-doctoral researcher will test and execute weather modeling based on the science use cases from section 4.

The project team will also jointly develop outreach materials to demonstrate the capability of the I-WRF framework, and weather forecasting models in general, that can be incorporated into NCAR outreach materials.

6 INITIAL PROJECT GOALS

The first few months of the project have consisted of assembling the project team, making connections, and establishing cadence for team meetings. The team has taken stock of the current state of the existing WRF container developed by NCAR as a technology exercise and mapped out the requirements for expanding the capability of the containers to further execution environments with intervention-free interoperability between WRF and MET. Users will want the ability to provide their own data, namelists, and build files for WRF, which requires developing interfaces that allow this upon instantiation. The Cornell team has demonstrated multi-container execution for larger-scale models, and execution scripts for multi-node need to be refined and adapted for more execution environments and to include verification capabilities.

7 DISCUSSION

The I-WRF project is intended to broaden capabilities for weather modeling, tested by demonstrating three timely science use cases. Simplifying the installation and execution of WRF software and accompanying tools will reduce barriers to weather modeling on a broader set of systems and place fewer demands on cyberinfrastructure professionals for complex compilation, configuration, and maintenance of the WRF software. By making weather modeling more accessible, researchers can demonstrate the utility of weather forecasting tools for supporting critical, relevant inquiry to a new generation of future researchers.

ACKNOWLEDGMENTS

This work is supported by NSF Award 2209711 and through the cooperation of the Cornell Center for Advanced Computing and the National Center for Atmospheric Research.

REFERENCES

[1] 2022. Frameworks: Large Scale Atmospheric Research Using an Integrated WRF Modeling, Visualization, and Verification Container Framework (I-WRF). https://www.nsf.gov/awardsearch/showAward?AWD_ID=2209711&HistoricalAwards=false NSF Award # 2209711.
[2] Barbara Brown, Tara Jensen, John Halley Gotway, Randy Bullock, Eric Gilleland, Tressa Fowler, Kathryn Newman, Dan Adriaansen, Lindsay Blank, Tatiana Burek, et al. 2021. The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society* 102, 4 (2021), E782–E807.
[3] Melissa S Bukovsky, Jing Gao, Linda O Mearns, and Brian C O'Neill. 2021. SSP-based land-use change scenarios: A critical uncertainty in future regional climate change projections. *Earth's Future* 9, 3 (2021), e2020EF001782.
[4] Sara C Pryor, Rebecca J Barthelmie, Melissa S Bukovsky, L Ruby Leung, and Koichi Sakaguchi. 2020. Climate change impacts on wind power generation. *Nature Reviews Earth & Environment* 1, 12 (2020), 627–643.

[5] William C Skamarock, Joseph B Klemp, Jimy Dudhia, David O Gill, Zhiquan Liu, Judith Berner, Wei Wang, Jordan G Powers, Michael G Duda, Dale M Barker, et al. 2019. A description of the advanced research WRF model version 4. *National Center for Atmospheric Research: Boulder, CO, USA* 145, 145 (2019), 550.
[6] Ryan C Sullivan, Robert C Levy, Arlindo M Da Silva, and Sara C Pryor. 2017. Developing and diagnosing climate change indicators of regional aerosol optical properties. *Scientific reports* 7, 1 (2017), 1–13.
[7] I-WRF Project Team. 2023. I-WRF Project Website. https://www.i-wrf.org https://www.i-wrf.org.

Received 21 July 2023