

## ***Interactive comment on “WRF4G: WRF experiment management made simple” by V. Fernández-Quiruelas et al.***

**Anonymous Referee #2**

Received and published: 2 September 2015

General comments:

The paper introduces the WRF management tool WRF4G, which facilitates the execution of simulation experiments. WRF4G covers the entire work flow for different types of experiments (parameter/ensemble, hindcast, long-term climate) and allows users to use multiple computational resources in a flexible way. The authors thereby address the needs of a large community and provide a detailed and well-structured explanation of the tool WRF4G. The examples contained in the software package make it possible to reproduce the steps discussed in the paper.

A number of alternative workflow managers exist for modeling with WRF or other numerical weather prediction/regional climate models, which aim either at NWP or a more general application. The focus of WRF4G on the above mentioned type of experiments

C1850

therefore distinguishes it from other workflow managers. This difference, however, is not highlighted enough in the paper.

Further, the use of English language in the paper needs to be improved. British English and American English are mingled in the text and fairly casual language is used in some parts. For details and further corrections of typos etc. see below ("technical corrections").

Specific comments:

(a) It is left unclear how the pre-processors (before ungrib) and post-processors (after wrf) are embedded in WRF4G. Which kind of pre-/post-processors have been used/tested, how are they configured and triggered? Section 5.2.1, lines 16-19 (WRF4G\_APPS), mentions that the data repository by default contains "some pre- and post-processors". Further reference is given in section 4.2, lines 7-8. Please provide more details on which kind of software, which kind of data can be pre-processed and what the post-processing software does.

(b) One major issue that is not addressed or discussed here is the comparability of results when running NWP/RCM experiments on different hardware. It is widely known that the choice of hardware (e.g., architecture), software (e.g., OS, compiler, compiler optimizations) and parallelization (e.g., number of tasks, MPI vs. OpenMP vs. hybrid MPI+OpenMP) can have significant influence on the results, in particular for mid- to long-term simulation experiments. While a scientific discussion of this issue is beyond the scope of the paper, I encourage the authors to mention this problem and provide simple solutions (e.g., running appropriate, identical test cases on all machines involved in the experiment to assess the reproducibility across different systems).

(c) The list of recent examples of WRF multi-physics ensembles and other parameterization changes (page 6557, last paragraph) is incomplete and should include at least: Brown and Sylla (2012) analyze the "Regional Climate Model Sensitivity to Domain Size for the Simulation of the West African Summer Monsoon Rainfall", while Klein

C1851

et al. (2015) study the "Variability of West African monsoon patterns generated by a WRF multi-physics ensemble". Related to that point are also experiments using different land-use classifications, adjustments of the land-surface model parameter tables etc. These experiments require different geogrid files, which should be mentioned on page 6561, lines 10-13.

(d) The differences between other workflow managers, cited in the text, and WRF4G could be stated more clearly (for example on page 6554). I suggest to also mention the missing or existing (?) capability of data assimilation, observation nudging and running nested domains using ndown. The former two are often used for hindcast experiments, while the latter finds its application in long-term climate simulations.

(e) The section 7 "Conclusions" could summarize in one short paragraph at the end where WRF4G has been used thus far and what distinguishes it from other workflow managers.

Technical corrections:

(1) Choice of flavor of English language: British English and American English are used concurrently in the paper: modelling (UK), modeling (US), parameterize (US), analyse (UK), etc. Please choose one flavor and revise the spelling.

(2) Consistent spelling of pre-/post-processor or pre-/post-processing. In parts of the text, the "-" is omitted. Likewise, consistent spelling pre-compile(d) or precompile(d) should be used.

(3) Page 6553, line 12: "computationally-efficient framework, which"

(4) Page 6554, line 2: "time-consuming task, which requires"

(5) Page 6556, line 1: "this domain and stores the information in"

(6) Page 6556, line 9: "are interpolated horizontally to the model domain"

(7) Page 6556, lines 13-14: "in the format required by the numerical core of the model

C1852

(NetCDF with WRF metadata)"

(8) Page 6557, line 5: "This section illustrates"

(9) Page 6557, line 16: "soil-atmosphere" (the - is too long in the paper)

(10) Page 6559, line 4: "use a computer resource for months to produce"

(11) Page 6559, lines 19-21: This sentence is unclear, please revise: "In the case of ... errors in the output"

(12) Page 6560, line 3 and page 6563, line 10: "become familiar"

(13) Page 6560, lines 6-7: "In WRF4G, a set of individual simulations, which try to answer ... questions, are called "experiment" "

(14) Page 6560, line 12: "WRF is able to produce restart files, which"

(15) Page 6560, line 19-20: ", which is the case for climate change projections"

(16) Page 6560, line 24: "dependent WRF simulations that read the last"

(17) Page 6561, line 8 and line 16: "execution on the computing node"

(18) Page 6561, line 18: "track the events occurring during the model execution"

(19) Page 6561, lines 25-26: This sentence is very casual, please revise: "And the ... also an issue"

(20) Page 6562, line 5: "can configure the WRF4G Computing Resources (WCR)"

(21) Page 6562, lines 10-11: "chunks can be run on any of the configured resources or just on a subset of them"

(22) Page 6562, line 14: This sentence is too casual and can be deleted entirely: "Nothing else needs to be installed or configured."

(23) Page 6562, line 20: This sentence is too casual, please revise: "The precompiled

C1853

... OpenMPI."

(24) Page 6563, lines 8-9: "boundary datasets, needed to run ... version of WRF, is included"

(25) Page 6563, line 11: "WRF4G runs a database that retains" - persists cannot be used here

(26) Page 6563, lines 13-15: This sentence is grammatically incorrect, please revise: "Although the database ... could be used."

(27) Page 6563, lines 4-5: Please consider omitting part of this repetition, in particular the URL in the footnote.

(28) Page 6564, lines 16-17: This sentence is incomplete, please correct: "These starting dates ... and end\_date." Also, the following sentence should start with "Further two variables ..." and end with "of each realization, respectively (see examples in Sect. 6)"

(29) Page 6565, line 21: "before running an experiment"

(30) Page 6566, line 4: "can be located on a different server"

(31) Page 6566, line 15: "grids; all of them being configured"

(32) Page 6567, line 7: "selects the best-suited"

(33) Page 6567, line 17: "against OpenMPI libraries"

(34) Page 6567, line 24: "to the data repositories referred to by the"

(35) Page 6568, line 1: Too casual, suggestion: "Then, users can submit the entire experiment (by default) or a subset of it by running"

(36) Page 6568, lines 3-5: "These chunks are efficiently scheduled on the computing resources, taking into consideration the dependencies between them."

C1854

(37) Page 6568, line 7: "when the second chunk of an experiment"

(38) Page 6569, lines 3-6: Replace "running scheme" with "approach" or "method" and revise the following sentence (grammatically wrong and too casual): "They found that ... cheapest"

(39) Page 6569, lines 8-9: "Fig 8. Setting the variable "multiple\_dates=1", "simulation\_interval\_h = 24"" and "simulation\_lenth = 42" tells WRF4G"

(40) Page 6570, lines 2-3: "execution framework WRF4G that" - omit the entire footnote (repetition of the URL etc.)

(41) Page 6570, line 4: "data management frameworks, and provides"

(42) Page 6570, lines 7-8: "This leads to an enormous saving of time ... with little additional overhead."

(42) Page 6570, lines 13-14: "such as grids and supercomputers allow the WRF community"

(43) Page 6576, Figure 3, entire caption: "... three sections, each of them describing a WCR. The first resource ... accessed through SSH. The remaining two clusters have been ... max\_jobs\_in\_queue. The my\_workstation section only contains a max\_jobs\_running entry due to the fact that ..."

(44) Page 6578, Figure 5, second sentence of caption: Too casual, please revise.

(45) The abbreviation "WRF4G" = WRF for Grid(s?) is not explained in the paper.

(46) The typesetting of binaries, files, parameters in configuration files, variables etc. is inconsistent (italic, texttt, uppercase, or combinations of these). Please revise and unify.

---

Interactive comment on Geosci. Model Dev. Discuss., 8, 6551, 2015.

C1855