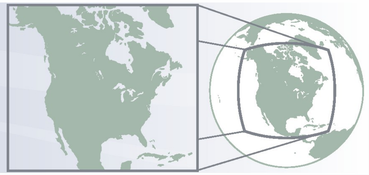
A33A
0209

NARCCAP



NCAR

Standardization of NARCCAP Data for Automated Processing, Quality Control, and Usability

Seth A. McGinnis, Linda O. Mearns, and Larry McDaniel — National Center for Atmospheric Research, Boulder, CO

email: mcginnis@ucar.edu website: http://www.narccap.ucar.edu

ABSTRACT

The North American Regional Climate Change Assessment Program (NARCCAP) is an international program to produce high resolution climate change scenarios and investigate uncertainties in regional scale projections of future climate by nesting multiple RCMs within a collection of driving AOGCMs. The resulting 60+ TB of output is intended for a wide variety of uses, including regional analysis, impacts studies, and further downscaling.

This poster discusses the importance of data standardization in supporting these goals, usability insights from the 2nd NARCCAP Users Workshop, and techniques for detecting and correcting common data errors encountered in the quality-control process, as well as providing an update on NARCCAP progress and some preliminary climate-change results.

QUALITY CONTROL

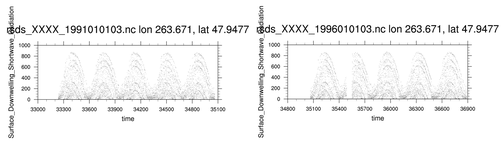
60 TB is a very large volume of data to check. The NARCCAP QC process was designed with the intent of automating every task that does not require human judgment. The process evolves: when new types of errors are discovered, new tests are added to detect them. Current tests include:

- Run Rosalyn Hatcher's `cfchecks.py` script to check whether data structure and metadata follow CF conventions
- Check for missing variables and files
- Print date/time of first and last timestep to check temporal coverage, time coordinate metadata
- Plot single timestep in file to check validity of coordinate system and map projection metadata
- Plot climatological average to check that spatial pattern and value range is reasonable, test for presence of bad data
- Plot cross-sections through data cube to look for missing or scrambled data
- Sanity-check headers for meaning

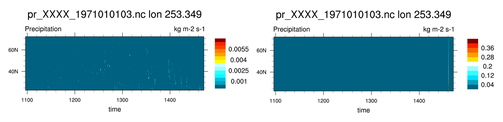
SEEING ERRORS...

The human eye is an excellent tool for finding differences between patterns. Many common errors become obvious when visualized in the right way

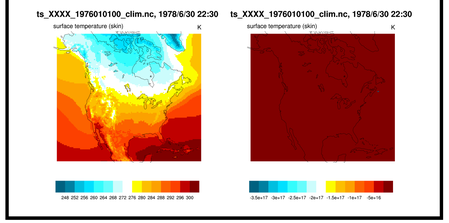
Gaps are easily visible in a time-series plot for a single gridcell



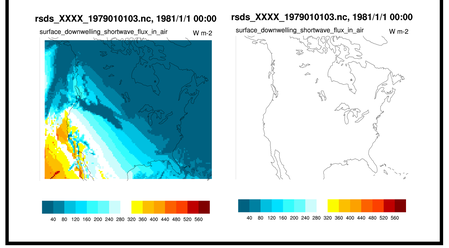
Time-space cross-section plots reveal bad timesteps



Climatology plots expose isolated anomalous values



Bad projection information results in an empty plot



...AND FIXING THEM

The modelers must spend considerable time after the model has finished running post-processing the output into publishable form. Where possible, when errors are found during QC it is better to correct the error than to redo the post-processing. Whether an error can be fixed depends on the type of error.

Metadata: These errors are easily fixed by hand.

Coordinate Variables: Incorrect coordinates can be fixed if the correct values are known. Often they can be copied from another file.

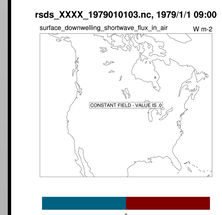
Data: Structural errors such as extra degenerate dimensions or incorrect ordering are usually fixable with some effort. Isolated bad values can be replaced with missing_value. Other errors, such as missing data, require the modelers to regenerate output files — or possibly to re-run the model.

STANDARDIZATION AND USABILITY

Standardization is a key factor in usability because it creates structural regularity and predictability that enables automation, which decreases investment in the creation of tools to handle the data, leaving more resources for understanding it. The ultimate in reusability is standards-aware software, where someone else has already written the code for you. For example, ensuring that your NetCDF data follows the CF metadata standard gives you GIS compatibility for free.

At the 2nd Users Meeting, we decided to increase uniformity by disaggregating 3-D data by pressure level, treating each vertical level like a separate 2-D variable. As a result, the files for 3-D variables now have the same structure as those for 2-D surface variables, allowing us to re-use existing tools for post-processing and QC instead of developing new ones. The same re-use advantage accrues to users as well. In addition to savings in time and effort, the new organizational scheme also increases usability by allowing us to prioritize standard pressure levels for earlier processing and publication.

PUZZLE: This plot is empty, but there's no error. Why?



Answer: it's right.

NARCCAP AT A GLANCE

- 6 downscaling RCMs
- 4 driving AOGCMs + NCEP
- Domain: North America
- 50 km spatial resolution
- 3 hour temporal resolution
- 57 output variables
- Current period: 1970-2000
- Future period: 2040-2070
- Emissions scenario: SRES A2
- CF-compliant NetCDF output
- FREE

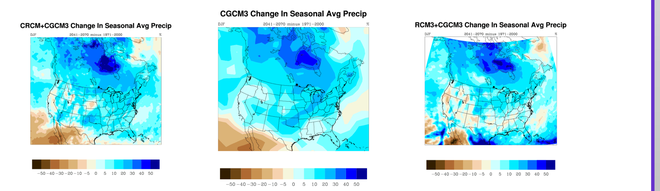
http://narccap.ucar.edu

PROGRESS AND PRELIMINARY CLIMATE CHANGE RESULTS

NARCCAP is designed to allow apples-to-apples comparison of model outputs. Data from complementary pairs of models is now available, allowing for the evaluation of uncertainty due to model choice and the study of how the driving GCM influences the downscaling RCM. Status of the modeling runs is shown in the table below.

RCM \ GCM	Phase I	Phase II			
	NCEP	CCSM	CGCM3	GFDL	HADCM3
CRCM	finished	running	finished	—	—
ECPC	finished	—	—	running	planned
HRM3	finished	—	—	planned	finished
MM5I	finished	running	—	—	planned
RCM3	finished	—	finished	finished	—
WRFP	finished	finished	running	—	—

One global model downscaled by two different regional models



Two different global models downscaled by the same regional model