Model Evaluation Tools: an Overview
MET Team

- David Ahijevych (MMM/NCAR)
- Barbara Brown (DTC/NCAR)
- Randy Bullock (NCAR)
- Eric Gilleland (NCAR)
- John Halley Gotway (DTC/NCAR)
- Lacey Holland (DTC/NCAR)

In alphabetical order
Objectives

- Develop a verification package to be used at the Developmental Testbed Center (DTC) to support transfer of technology from research to operations
- Build on capabilities of existing verification packages
- Incorporate state-of-the-art new verification methodologies from the community
Requirements

- Toolkit must include:
  - standard verification approaches
  - confidence intervals
  - initial capability for spatial verification
- Replicate most existing capabilities of other systems
- Documentation
- Code maintenance
- Ability to implement additional capabilities
- Freely available
Technical Dependencies

**Required to compile:**
- C++ and Fortran compilers (GNU or PGI)
- BUFRLIB
- NetCDF
- GSL (GNU Scientific Library)
- F2C Library

**Recommended:**
- WRF Post-Processor
- COPYGB (included with Post)
- CWORDS
Confidence intervals

Confidence intervals take into account various sources of error (e.g., sampling).

Computation of confidence intervals for verification stats is not always straight-forward

- Parametric vs. non-parametric methods
- Take into account spatial and temporal correlations

Toolkit includes both parametric methods (e.g., assuming Gaussian distribution, v1.0) and re-sampling methods (e.g., bootstrapping, v1.1)
Grid-to-grid verification (grid_stat)

- Includes “standard” stats package
- Includes confidence intervals
- Used for verifying forecast field with a gridded verification dataset (e.g. Stage II precipitation)
- Provides continuous and discrete statistics
Method for Object-based Diagnostic Evaluation (MODE)

- MODE (Method for Object-based Diagnostic Evaluation) Tool
- Advanced spatial verification methods
- Defines objects in both observation and forecast fields
- Other advanced spatial verification methods included in the future
Grid-to-point verification 
*(point_stat)*

- Uses point-based observation types (e.g. rawinsondes, surface stations)
- Matched pair data (*v1.1*)

- Provides continuous and discrete statistics with confidence intervals
- User selects observation types
- First release uses prepbufr format
- Demand for ASCII (new utility in *v1.1*)
- User can use other data sources via NetCDF interface
Analysis Tools

- Filters a collection of output files from *point_stat*, *grid_stat*, and *mode* into user-specified subsets.
- Summarizes statistic values over time or combinations of regions.
- Aggregates statistics over time or combinations of regions.
- Calculates additional metrics that require a combination of statistics at different lead times.
http://www.dtcenter.org/met/users

~325 users from 37 countries

Online survey + met_help to receive feedback

Identify community needs and demands!
Community Feedback

- WRF User’s meeting on the system at AMS in San Antonio (Jan 2007)
- Local RAL-wide Townhall (Jan 2007)
- WRF Verification Workshop (Feb 2007)
- Two online surveys to date
- Visitors through the DTC visitor program
Current and Upcoming Programs

- Coretest 2007
- SPC Spring Program 2008
- CoSPA (Consolidated Storm Prediction for Aviation)
Helping Users…

- Two WRF Tutorial talks (already!)
- List of current bugs and issues actively maintained on webpage
- Met_help@ucar.edu:
  - Actively respond to user requests for assistance
  - Maintain list of commonly asked questions to identify trouble spots
  - Learn where users are having problems
  - Get feedback from users
Future Work

- Version 1.1 due out late spring
- Identify and implement techniques based on feedback
- Continue development of the verification system with controlled successive version releases
- In future, will have capabilities for database interface
- Intercomparison of spatial verification techniques is underway
- More data formats:
  - ASCII
  - Grib2
  - Improve compatibility of NetCDF
Future Work (con’t)

Additional techniques to be included:
- Neighborhood verification methods (including FSS in v1.1)
- Intensity-scale
- CRA
- Ensemble methods
- Hurricane Tracking
- Others
Considerations & Concerns

Prioritizing demands
- Lots of possibilities for additions/improvements
- Data formats
- Forecast types
- Verification methods

Growing list of libraries
- How many is reasonable?
- Is it worth time and resources getting rid of some of them?

Identify additional methods for inclusion
- Would like to include many, but have time and budget constraints

Need for education and training on new verification methods
Acknowledgments

- U.S. Air Force Weather Agency
- NOAA

Thanks also to Barbara Casati for help developing an Intensity-scale method version for MET!