Model Evaluation Tools (MET)

MET Development Team

Developmental Testbed Center

27 June 2008

MET Development Team

- Dave Ahijevych (scientist)
- Barbara Brown (scientist/statistician)
- Randy Bullock (software engineer)
- Eric Gilleland (scientist/statistician)
- John Halley Gotway (software engineer)

With thanks to the Air Force Weather Agency (AFWA) and NOAA for their support

Also thanks to Lacey Holland who helped lead much of the MET implementation so far

Outline

- MET Overview – Barb
- Technical details – John
- Point-stat and grid-stat science – Barb
- Confidence intervals – Eric
- Point- and grid-stat practical – John
- MODE science – Randy
- MODE Tool Practical – John
- VSDB and MODE analysis tools – Dave
- Future Plans – Barb

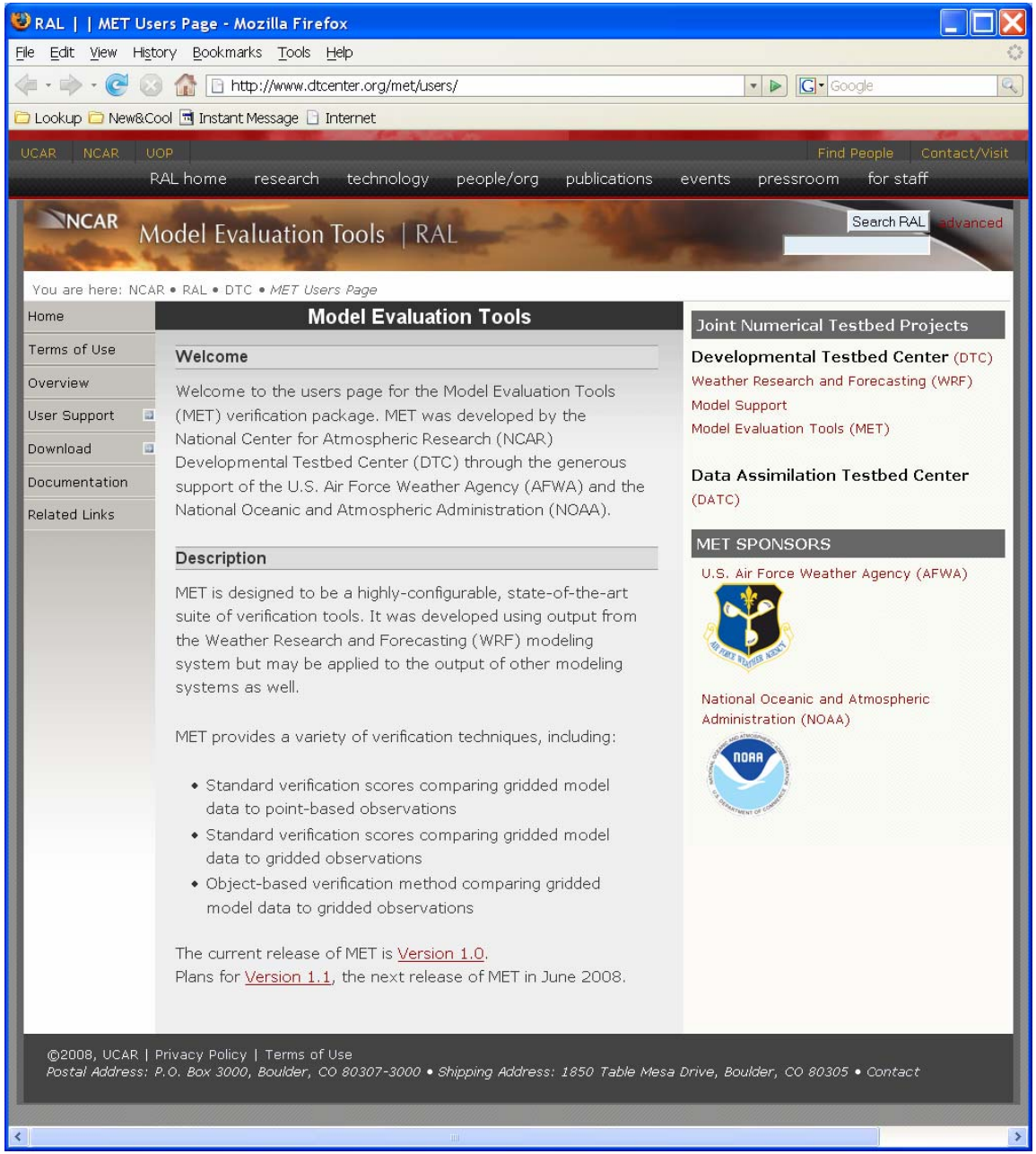
MET: A community tool

- Goal: to provide a set of forecast verification tools that are
  - Openly available to the community
  - “Created” by the community, through contributed methods and capabilities
    - Evaluation methods
    - Graphical methods
- Community includes diverse users
  - WRF Developers
  - Development Testbed Center (DTC)
  - University researchers
  - Operational centers

MET status

- MET implementation initiated in fall 2006
- Version 1.0 released in January 2008
- Version 1.1 to be released in July 2008
  - ASCII observations
  - Neighborhood methods
  - New confidence interval estimates for non-Gaussian measures
- Version 2.0 in early 2009

http://www.dtcenter.org/met/users/



MET is

- A modular set of verification tools that can be freely downloaded
- Fully documented
- Supported through an e-mail help address

Main MET components

- Data reformatting modules
  - Move data into the format(s) expected by MODE (ascii2nc; pb2nc)
  - Combine precipitation values across time periods (e.g., 24-h totals) (pcp_combine)
  - Subtract precipitation values to create values for finer sub-periods (pcp_combine)
- Statistics modules
  - Object-based spatial verification method (MODE)
  - Verification of grids (Grid-stat)
  - Verification at points (Point-stat)
- Analysis modules
  - Aggregate results across cases; stratify results by categories
  - VSDB analysis tool (for Grid-stat and Point-stat)
  - MODE analysis tool

Technical Information

- MET distributed as a tarball to be downloaded and compiled locally.
  - METv1.1 to be released in July, 2008.
  - Updated User's Guide and Online Tutorial.
  - Register and download: **www.dtcenter.org/met/users**
- Language:
  - Written primarily in C++ with calls to a Fortran library
- Supported Platforms and Compilers:
  - Linux machines with GNU compilers
    - g++ and gfortran or g77
  - Linux machines with Portland Group (PGI) compilers
    - pgCC and pgf77
  - IBM machines with IBM compilers
    - xlC and xlf

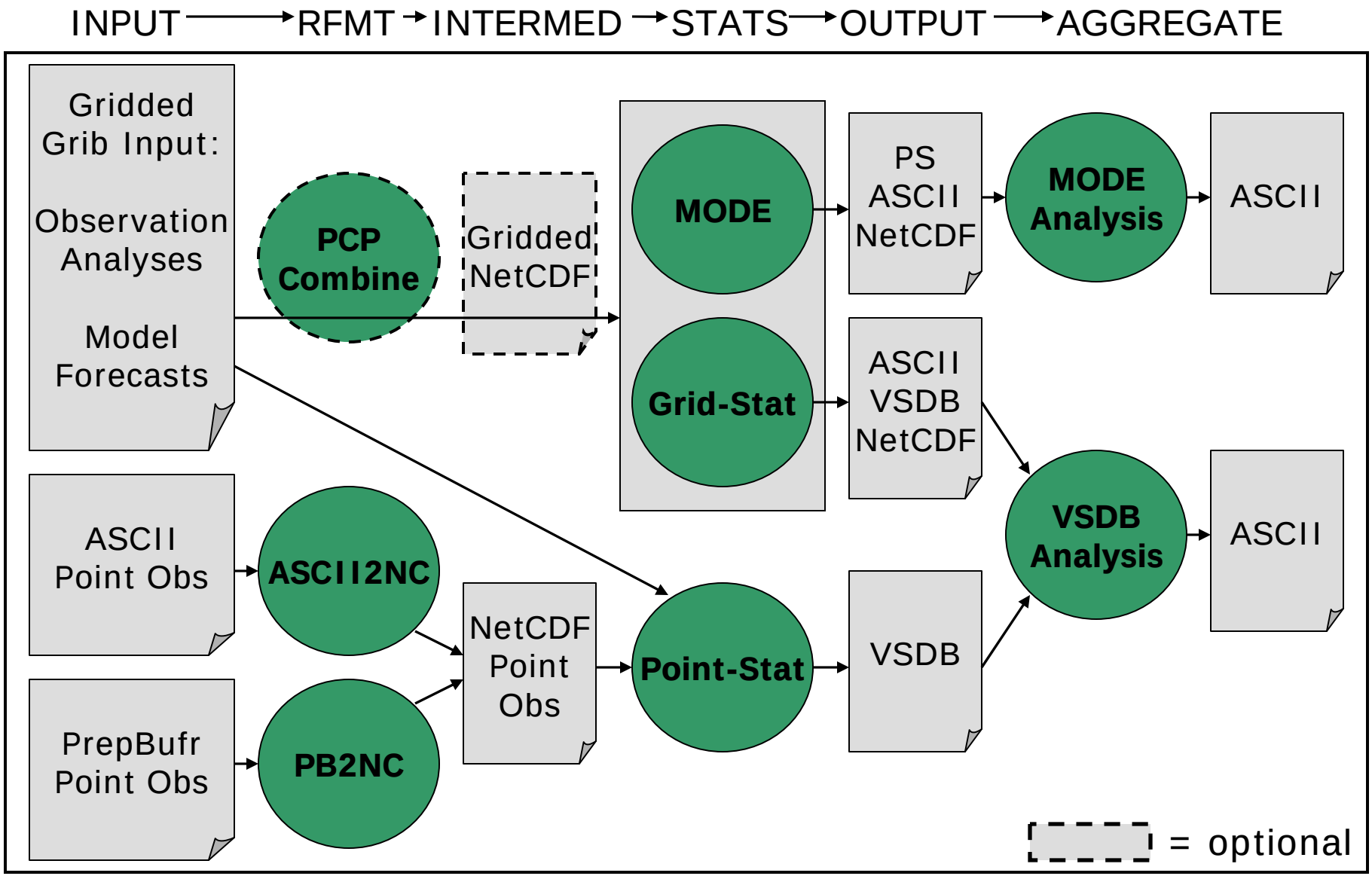
Dependencies

- Required to compile:
  - GNU Make Utility
  - C++ and Fortran compilers (GNU, PGI, or IBM)
  - NetCDF Library
  - BUFRLIB Library
  - GNU Scientific Library (GSL)
  - F2C Library (f2c or g2c)
- Recommended for use:
  - WRF Post-Processor
  - COPYGB (included with WRF-Post)
  - CWORDSH

Building MET

- Steps for building MET:
  1. Build the required libraries with the same family of compilers to be used with MET.
  2. Select the appropriate Makefile.
     - GNU, PGI, or IBM
  3. Configure the Makefile.
     - C++ and Fortran compilers
     - Paths for NetCDF, BUFRLIB, GSL, and F2C libraries
  4. Execute the GNU Make utility.
  5. Run the test script and check for runtime errors.
     - Runs each of the MET tools at least once.
     - Uses sample data distributed with the tarball.

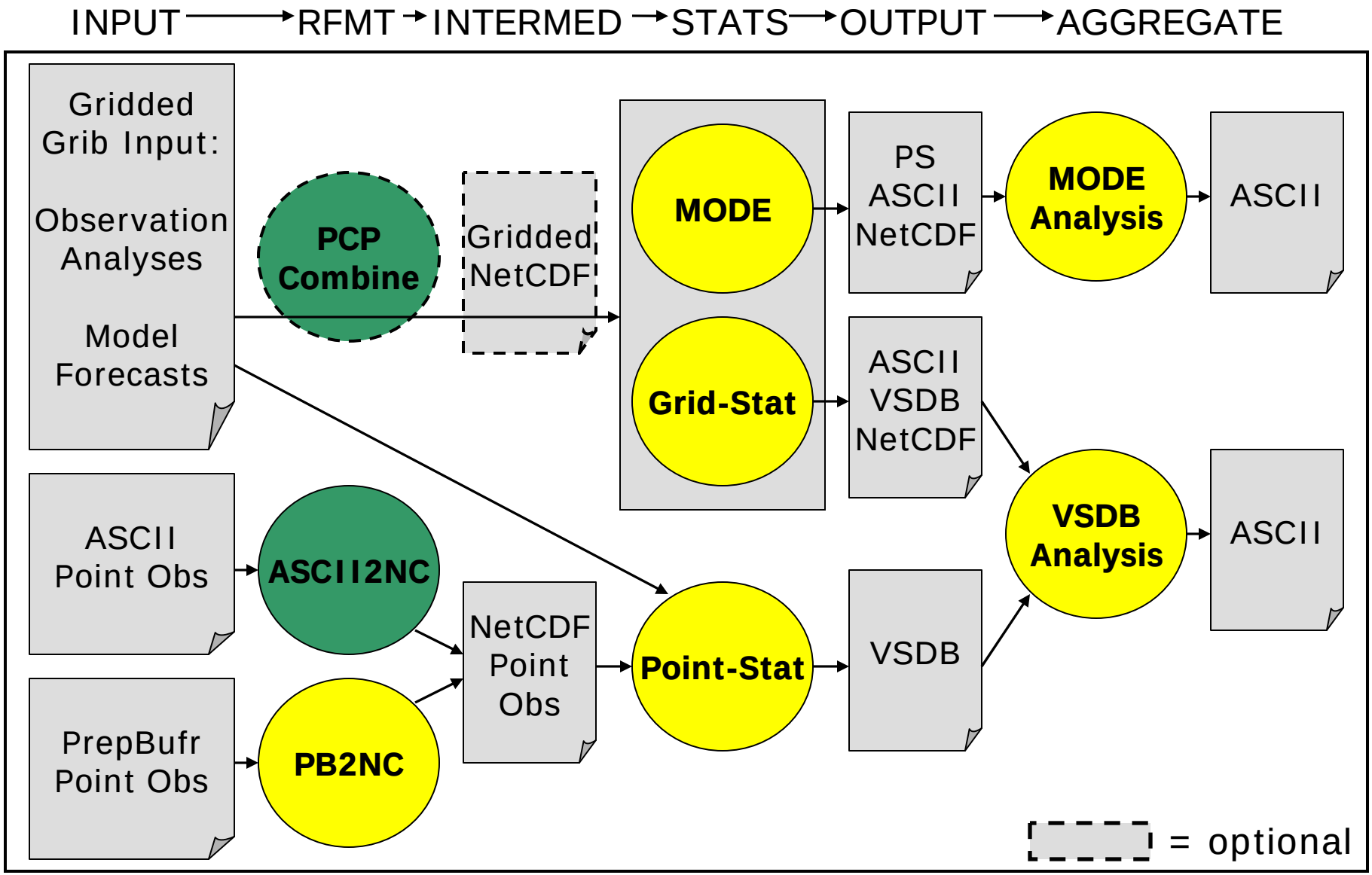
METv1.1 Flowchart



Configuration Files

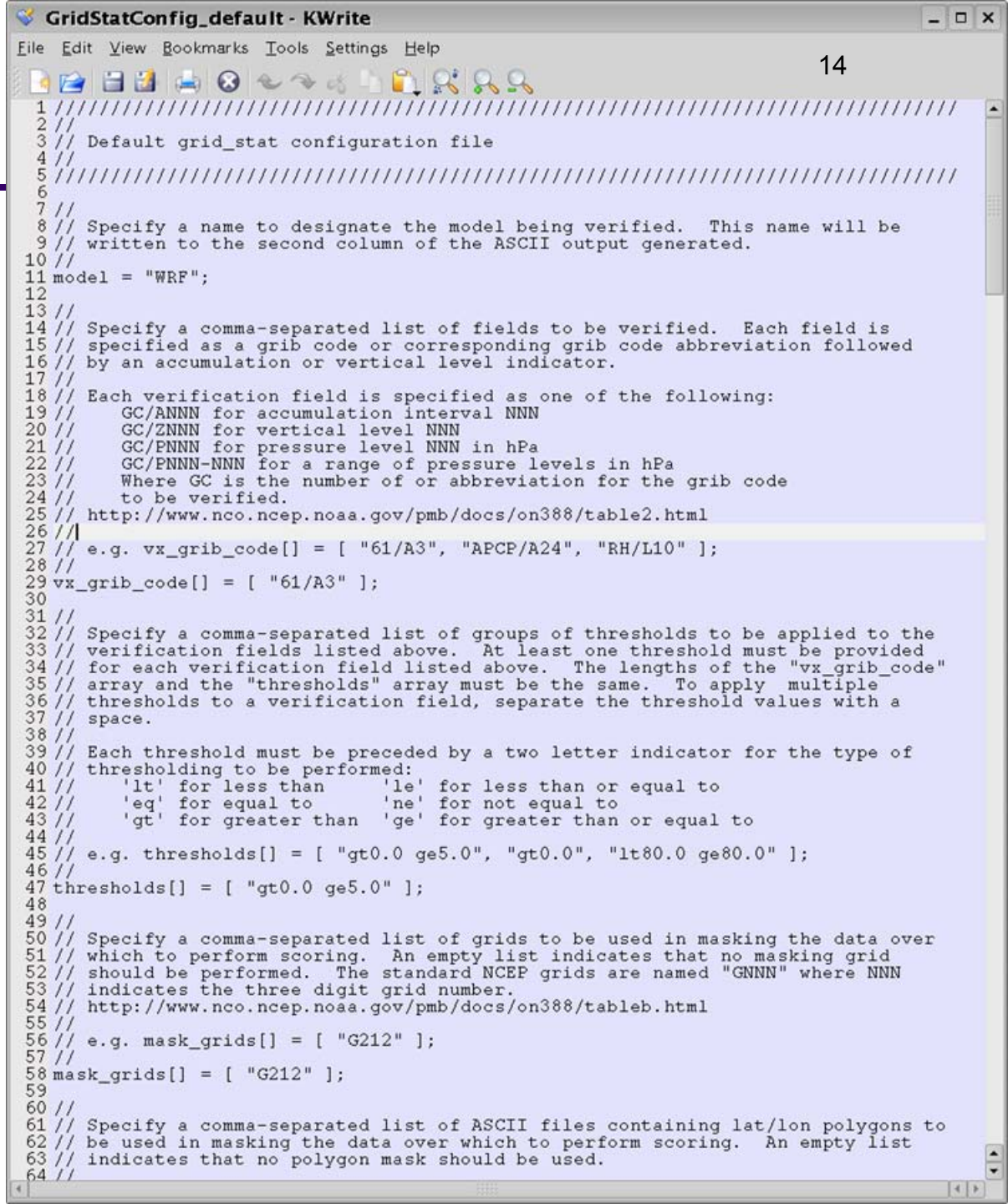
- MET is a set of command line tools which are controlled using ASCII configuration files passed to the tools on the command line.
- Configuration files control things such as:
  - Fields/levels to be verified.
  - Thresholds to be applied.
  - Interpolation methods to be used.
  - Verification methods to be applied.
  - Regions over which to accumulate statistics.
- Well commented and documented in User's Guide.
- Easy to modify.
- Use the version of the configuration files distributed with the tarball.

Use of Configuration Files

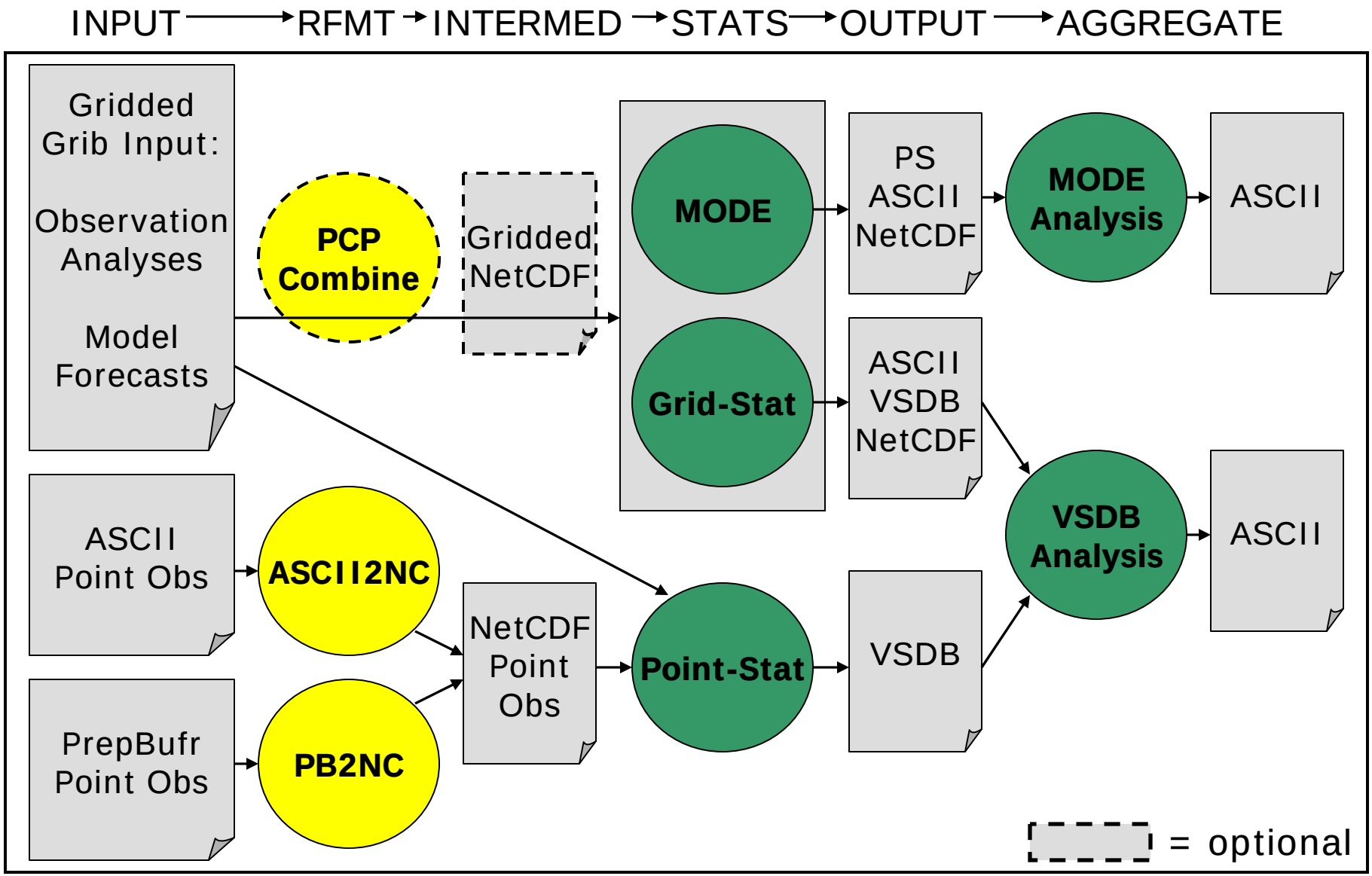


Sample File

- PB2NC (18)
- MODE (65)
- Grid-Stat (21)
- Point-Stat (20)
- VSDB-Analysis (23)
- MODE-Analysis (89)
- Only need to modify a few!



Data Reformatting Tools



PCP-Combine Tool

- Functionality:
  - Sum precipitation across multiple files.
  - New for v1.1:
    - Add precipitation in 2 files (i.e. NMM output).
    - Subtract precipitation in 2 files (i.e. ARW output).
  - Specify field name on the command line.
  - No configuration file.
- Data formats:
  - Reads GRIB.
  - Writes gridded NetCDF as input to stats tools.

PB2NC Tool

- Functionality:
  - Filter and reformat PREPBUFR point observations into intermediate NetCDF format.
  - Configuration file specifies:
    - Observation types, variables, locations, elevations, quality marks, and times to retain or derive for use in Point-Stat.
- Data formats:
  - Reads PREPBUFR using NCEP's BUFRLIB.
  - Writes point NetCDF as input to Point-Stat.
- CWORDSH utility for FORTRAN blocking

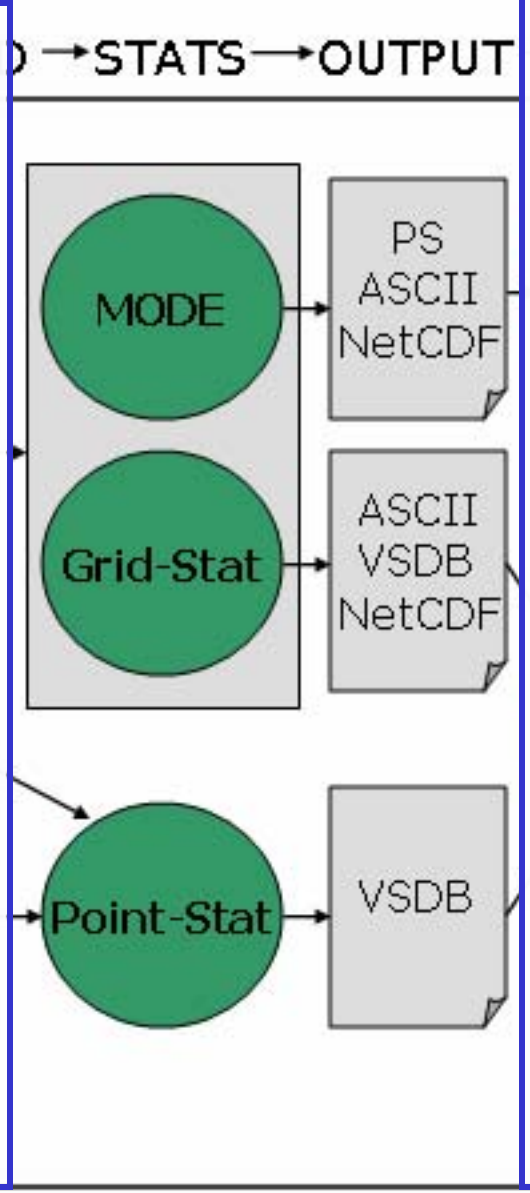
ASCII2NC Tool (new in v1.1)

- Functionality:
  - Reformat ASCII point observations into intermediate NetCDF format.
  - For v1.1, one input ASCII format supported (10 columns):
    - Message_Type, Station_ID, Valid_Time
    - Lat(Deg North), Lon(Deg East), Elevation(msl)
    - Grib_Code, Level, Height(msl), Observation_Value
  - No configuration file.
- Data formats:
  - Reads ASCII.
  - Writes point NetCDF as input to Point-Stat.
  - Support additional ASCII formats based on user input.

Grid-stat and Point-stat science

Stats tools

- *MODE*: Method for Object-based Diagnostic Evaluation
- *Grid-Stat*: Compares gridded forecasts and observations
  - Includes "neighborhood methods"
- *Point-Stat*: Compares gridded forecasts and point observations (e.g., rawinsonde output)
  - Includes several methods for matching forecasts to point obs



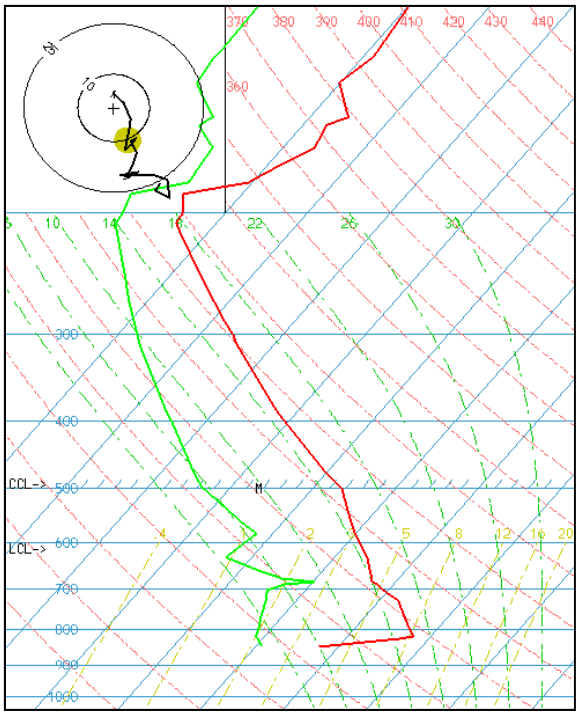
Statistics

Grid-Stat and Point-Stat:

- Traditional statistics
  - Contingency table statistics (POD, FAR, etc.)
  - Continuous statistics (RMSE, MAE, Bias, etc.)
- Confidence intervals
  - Parametric
  - Bootstrap

Point Stat: Grid-to-Point Verification

- Input Grib forecasts and NetCDF observations from PB2NC
- Select multiple…
  - Variables, levels, thresholds, masking regions, matching methods, confidence interval (CI) method and alpha values for CIs



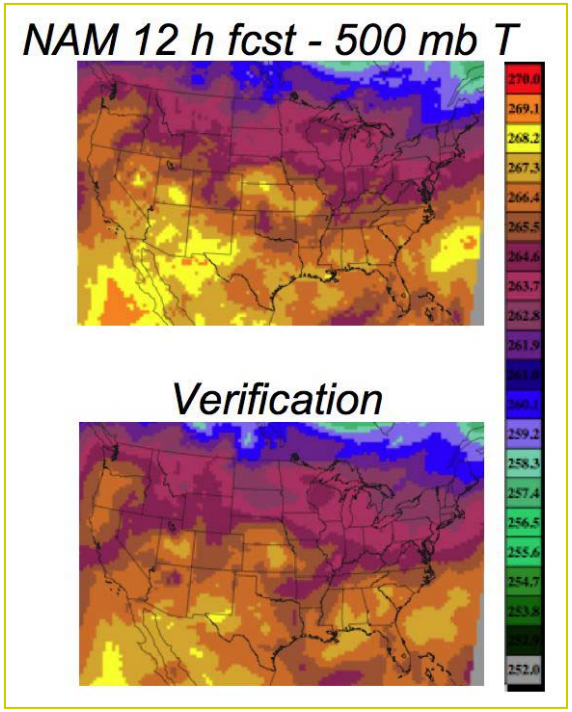
- Output VSDB and ASCII
  - Contingency table counts and statistics with CIs
  - Continuous statistics with CIs
  - Partial sums

Point-stat: Matching approaches

- User-specified region of gridpoints around the observation point:
  - 1 - nearest point
  - 2 - 2x2 box around point
  - 3 – 3x3 box around point
  - Etc.
- Several metrics can be used to create matched forecast value
  - Min, Max, Median, Unweighted mean, Distance-weighted mean, Least-squares fit
- User-defined min number of valid data points

Grid Stat: Grid-to-Grid verification

- Input Grib or NetCDF from PCP Combine
- Select multiple…
  - Variables, levels, thresholds, masking regions, smoothing methods, confidence interval (CI) methods and alpha values for CIs
- Output VSDB and ASCII
  - Contingency table counts and statistics with confidence intervals
  - Continuous statistics with CIs
  - Partial sums (L1L2, etc.)
  - Neighborhood methods
- Output NetCDF
  - Matched pairs and difference fields for each variable, level, masking region



Statistics for discrete variables

Measures for 2x2 contingency tables

Ex: based on applying a threshold to a continuous variable

- Number of observations
- FHO statistics
- Contingency table counts
- Contingency table proportions
- Accuracy
- Bias

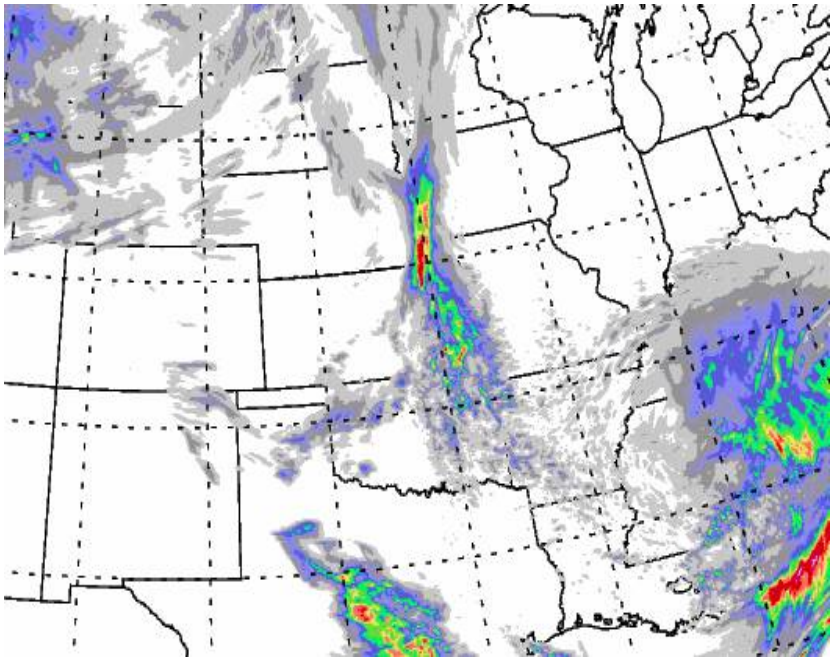
- Probability of Detection of Yes (PODy)
- Probability of Detection of No (PODn)
- False Alarm Ratio (FAR)
- Critical Success Index (CSI)
- Gilbert Skill Score (GSS = ETS)
- Hanssen and Kuipers Discriminant (H-K = TSS)
- Heidke Skill Score (HSS)
- Odds Ratio (OR)

Statistics for continuous variables

- Forecast/observation mean
- Forecast/observation standard deviation
- Correlation coefficients (Pearson, Spearman, Kendall's tau)
- Mean error (ME)
- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- Bias-corrected Mean Squared Error (BCMSE); also known as "standard deviation of the error" (ESTDV)
- Root-Mean Squared Error (RMSE)
- Error percentiles (10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th})
- Partial sums (1^{st} and 2^{nd} moments of the forecasts, observations, and errors)
  - Scalar
  - Anomaly
  - Vector
  - Vector anomaly

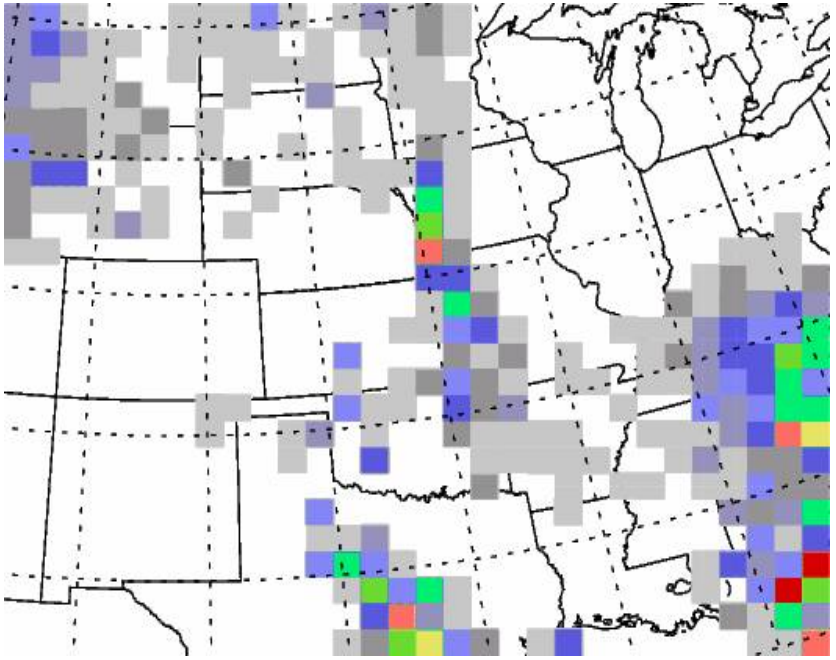
Neighborhood verification methods

- Also called “fuzzy” verification
- Upscaling
  - Put observations and/or forecast on coarser grid
  - Calculate traditional metrics
- Provide information about scales where the forecasts have skill



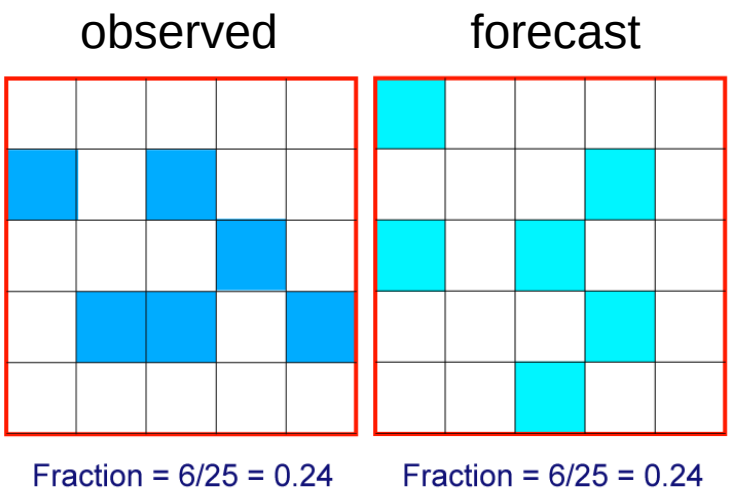
Neighborhood verification methods

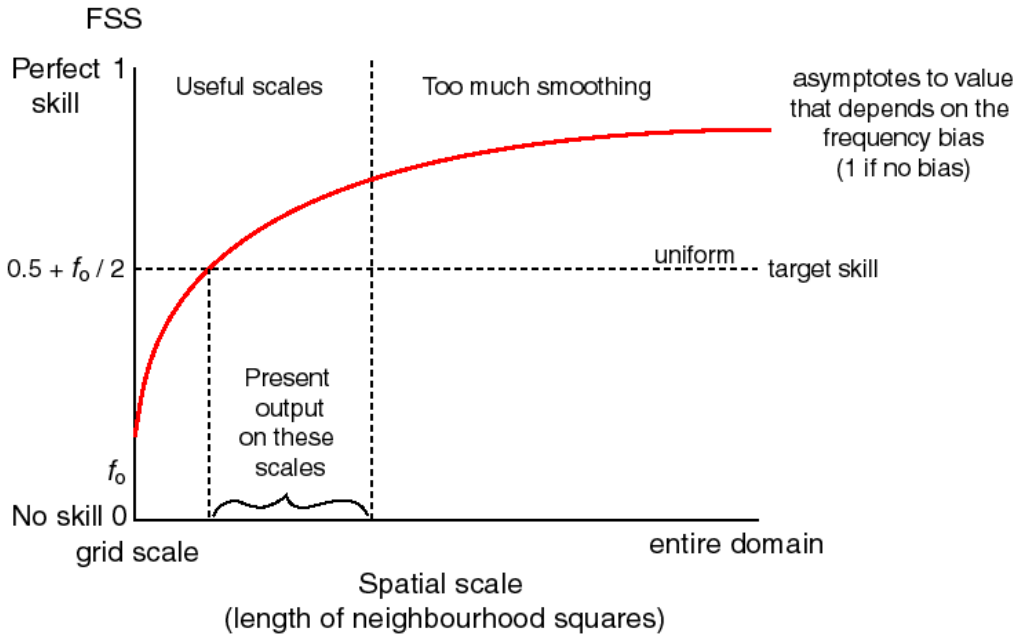
- Also called “fuzzy” verification
- Upscaling
  - Put observations and/or forecast on coarser grid
  - Calculate traditional metrics
- Provide information about scales where the forecasts have skill



Neighborhood verification methods

Example: Fractional skill score (Roberts and Lean, MWR, 2008)



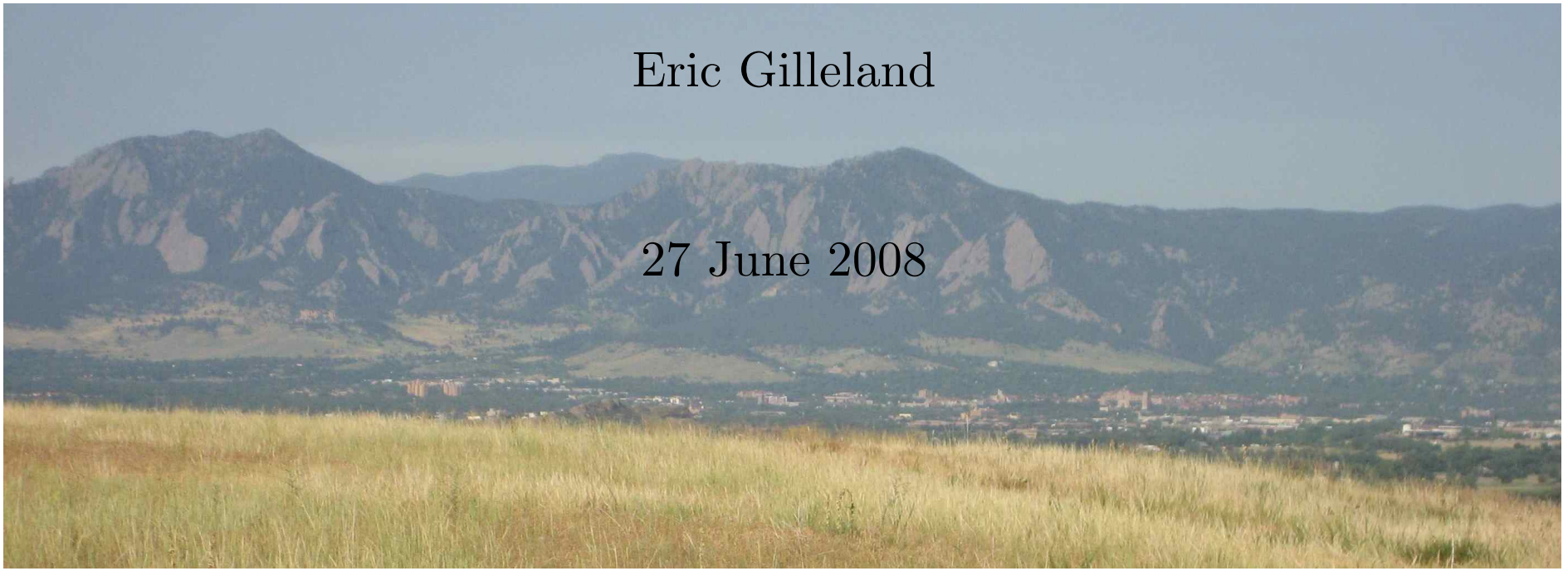


From Mittermaier 2008

Ebert (2008; Met Applications) describes the neighborhood methods in MET

28

Confidence Intervals and MET



Eric Gilleland

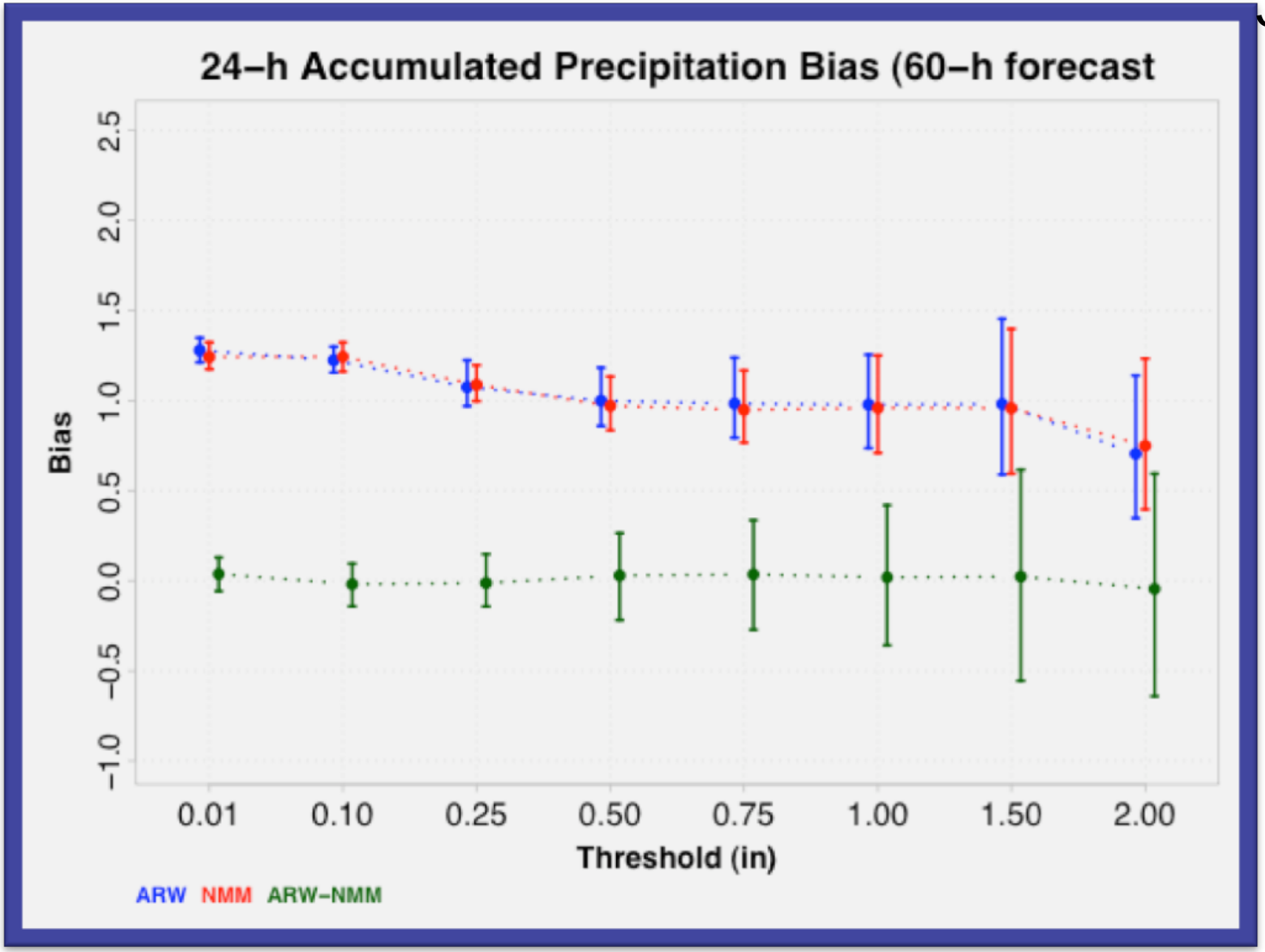
27 June 2008

Motivation: Why use confidence intervals?

$$\hat{\theta} \pm z_{\alpha/2} \cdot \mathrm{se}(\theta)$$

- Point estimate for (frequency) bias is 1.2
  - Is this significantly different from 1 (unbiased)?
  - A point estimate is a realization of a random variable.
- How much uncertainty is in the estimate?

30



Background

- A level α hypothesis test is related to a $(1-\alpha)\cdot 100\%$ confidence interval.
- Interpretation is that if the experiment were run 100 times (i.e., 100 CIs estimated), then the true parameter would fall within exactly $(1-\alpha)\cdot 100$ of those limits.
- For example, if $\alpha = 0.05$, then we expect the true parameter would fall inside the limits 95 times.

Normal Approximation

If the sample $\mathcal{X}_n = \{x_1, \ldots, x_n\}$ is independent and identically distributed (iid), then for large n, the distribution of $\mathcal{X}_n$ can often be approximated by a normal distribution. Then, for a given parameter, θ (e.g., mean, hit rate, odds ratio, etc.), a $(1-\alpha)\cdot 100\%$ interval for the estimate, $\hat{\theta}$ is given by

$$\hat{\theta} \pm z_{\alpha/2} \cdot \mathrm{se}(\theta),$$

where $z_\alpha = \phi^{-1}(\alpha)$, ϕ the standard normal distribution, and typically $\mathrm{se}(\theta)$ is replaced by an estimate, $\widehat{\mathrm{se}}(\theta)$.

Normal Approximation: Mean example

Must be able to find an estimate for se(θ).

For example, if $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$, then for $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, we have that $\widehat{\text{se}}(\bar{X}) = \sigma/\sqrt{n}$, and $(1-\alpha)\cdot 100\%$ CIs are given by

$$\bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},$$

and, of course, we replace σ (the standard deviation of $X_1, \ldots, X_n$) with its estimate $\hat{\sigma} = \sqrt{\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}}$.

Small n

When n is small (e.g., < 30), for the mean, it may be possible to use the t-distribution instead of the normal.

$$\bar{X} \pm t_{\alpha/2,n-1} \cdot \frac{\hat{\sigma}}{\sqrt{n}},$$

where $t_{\alpha,\nu}$ is the α quantile from a t-distribution with ν degrees of freedom.

Normal Approximation: Proportions

It is possible to estimate the standard error for most proportions (e.g., hit (false alarm) rate, POD, etc.

- A random sample of n items are observed, and X is the number of events (e.g., precipitation over a given threshold) in the sample.
- For n large, X has approximately a normal distribution.
- The sample proportion $\hat{p} = \frac{X}{n}$ subsequently also has approximately a normal distribution.
- The most straightforward interval is given by $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\hat{p}(1-\hat{p})/n}$ (Wald).
- For n small (and large), there is a better, slightly more complicated, interval (Wilson), which is used in MET.

Verification statistics with CIs relying on normal approximation

Directly or indirectly

- Mean Error (ME)
- Mean Squared Error (MSE), Mean Absolute Error (MAE)
- Variance and standard deviation
- (Linear) Correlation
- (Rank) correlations
- Hit Rate, False Alarm Rate (*proportions*)
- Pierce Skill Score (PSS)
- Odds Ratio (OR)

Bootstrap Confidence Intervals

37

- Assume the sample is representative of the population
- Resample *with replacement* from the sample B times
- Estimate the parameter of interest for each resample in order to obtain a sample of the statistic of interest.
- Calculate confidence intervals using the distribution of parameters
  - Percentile method is simplest
  - BC_a (adjusted percentile) has better accuracy, but computationally inefficient

Verification statistics CIs relying on the bootstrap 38

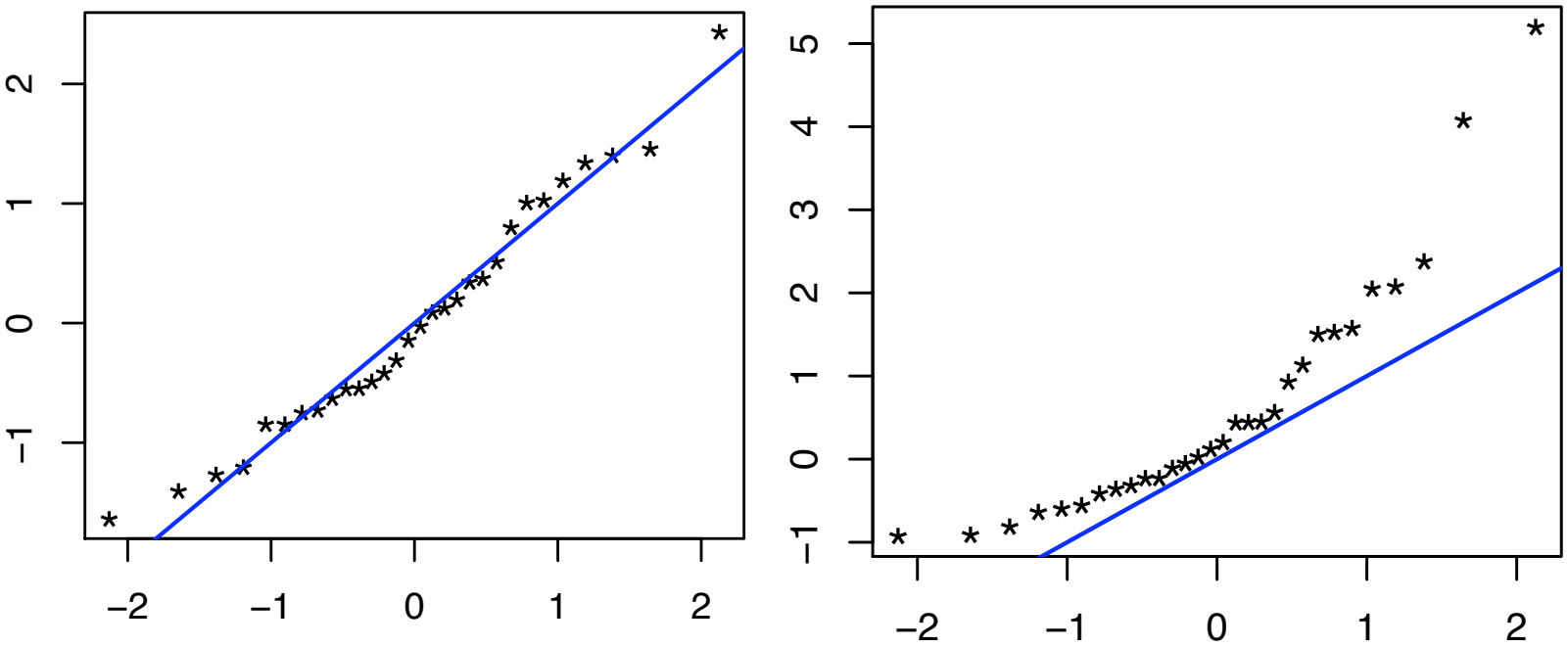
Some statistics cannot use the normal approximation, and often it is not known how to construct CIs without relying on resampling methods.

- Bias ($\bar{f}/\bar{o}$)
- Equitable Threat Score (ETS)
- For high percentiles, re-sample $m < n$ (e.g., $m = \lfloor \sqrt{n} \rfloor$)
- Others?

Some statistics (e.g., the mean) that can use the normal approximation will still have more accurate CIs via the bootstrap procedure.

Normal Approximation: Checking the Assumption of Normality

39



Empirical quantiles (ordinate) vs. Theoretical quantiles (abscissa)

Checking Independence Assumption, and accounting for Dependence

40

- Normal approximation bootstrap intervals assume iid samples
- Wildly incorrect intervals possible if these assumptions are not met
- Check plots!
- If the dependence can be modeled, best to model it and use normal approximation on errors or parametric bootstrap
- Inflate variance?
- Block bootstrap to be added soon
- Currently, uncertainty is only in the sample uncertainty
  - Observational uncertainty? (difficult, but maybe could be added)
  - Other sources of uncertainty?

Practical: Grid-Stat Tool

- Functionality:
  - Computes a variety of statistics for comparing a gridded forecast to a gridded observation.
  - One time step at a time.
  - Data must reside on a common grid.
    - Recommend COPYGB for regridding GRIB data.
  - Configuration file specifies:
    - Fields/levels, thresholds, vx regions (grids or polylines)
    - Smoothing options, neighborhood sizes
    - Vx methods, confidence interval options, output types
- Data formats:
  - Reads gridded GRIB and NetCDF output of PCP-Combine.
  - Writes ASCII statistics and NetCDF matched pairs.

Practical: Grid-Stat Usage

- Command Line:
  - Required: fcst_file, obs_file, config_file
  - Optional: -outdir, -v

Practical: Grid-Stat Line Types

- Statistics line types: 11 possible
  - Categorical - apply threshold
    - Contingency table counts (FHO, CTC, CTP, CFP, COP)
    - Contingency table statistics (CTS)
  - Continuous - raw fields
    - Continuous statistics (CNT)
    - Partial Sums (SL1L2)
  - Neighborhood – choose size (new for v1.1)
    - Neighborhood categorical (NBRCTC, NBRCTS)
    - Neighborhood continuous (NBRCNT)
- Ten header columns common to all line types.
- Data in remaining columns specific to each line type.

Practical: Grid-Stat Output

- Output files:
  - ASCII statistics file containing all line types (File ends with ".vsdb").
  - Optional ASCII files sorted by line type with a header row (File ends with "_TYPE.txt").
  - Optional NetCDF matched pairs and difference fields. (File ends with "_pairs.nc").
- Naming conventions:
  - grid_stat_HHMMSSL_YYYYMMDD_HHMMSSV [.vsdb | _pairs.nc | _TYPE.txt]
  - Ex: grid_stat_120000L_20050807_120000V.vsdb

Practical: Point-Stat Tool

- Functionality:
  - Computes a variety of statistics for comparing a gridded forecast to point observations.
  - One time step at a time.
  - Configuration file specifies:
    - Fields/levels, thresholds, vx regions (grids, polys, stations)
    - Vx message types, interpolation methods
    - Vx methods, confidence interval options, output types
- Data formats:
  - Reads gridded GRIB and NetCDF output of PCP-Combine.
  - Reads point NetCDF output of ASCII2NC and PB2NC.
  - Writes ASCII statistics.

Practical: Point-Stat Usage

- Command Line:
  - Required: fcst_file, obs_file, config_file
  - Optional: -climo, -ncfile, -outdir, -v

Practical: Point-Stat Line Types

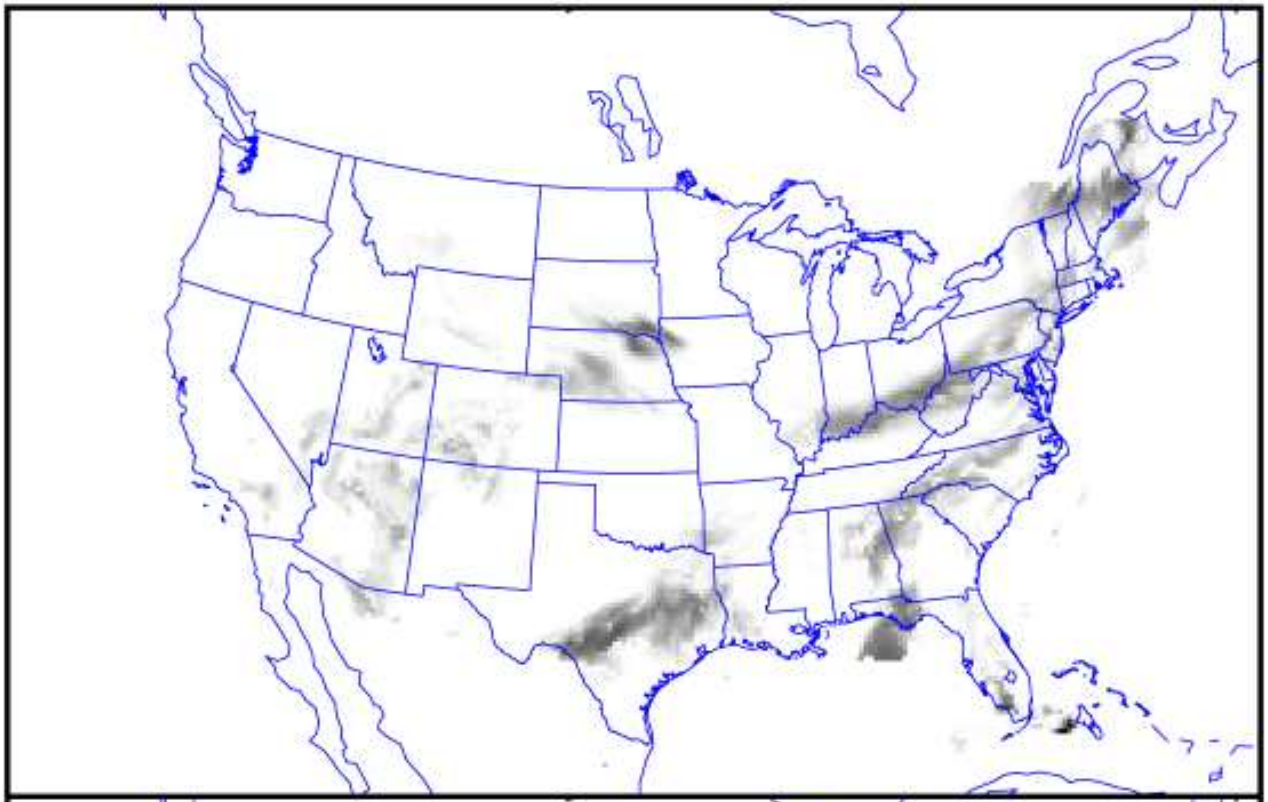
- Statistics line types: 12 possible
  - Categorical - apply threshold
    - Contingency table counts (FHO, CTC, CTP, CFP, COP)
    - Contingency table statistics (CTS)
  - Continuous - raw fields
    - Continuous statistics (CNT)
    - Partial Sums (SL1L2, SAL1L2, VL1L2, VAL1L2)
  - Matched Pairs (new for v1.1)
    - Raw matched pairs – a lot of data! (MPR)
- Ten header columns common to all line types.
- Data in remaining columns specific to each line type.

Practical: Point-Stat Output

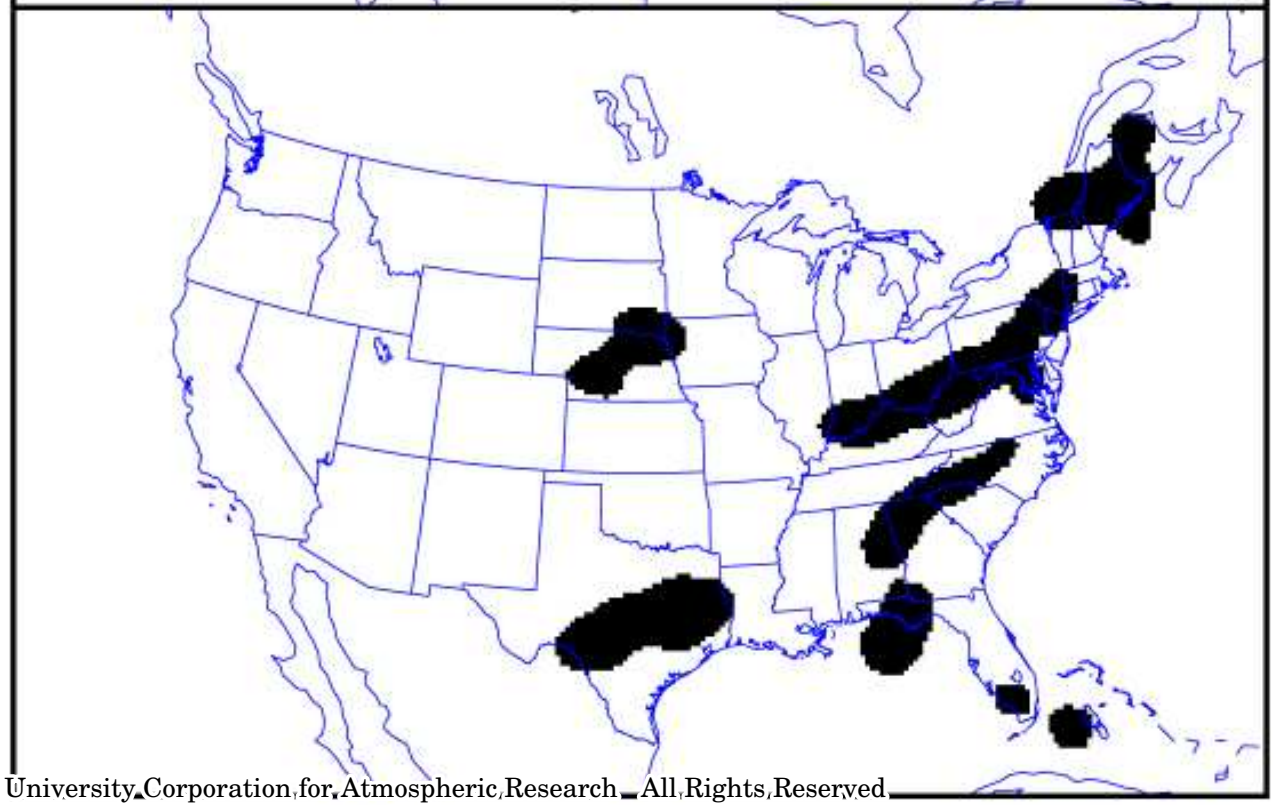
- Output files:
  - ASCII statistics file containing all line types (File ends with ".vsdb").
  - Optional ASCII files sorted by line type with a header row (File ends with "_TYPE.txt").
- Naming conventions:
  - point_stat_HHMMSSL_YYYYMMDD_HHMMSSV [.vsdb | _TYPE.txt]
  - Ex: point_stat_120000L_20050807_120000V.vsdb

MODE

Method for Object-Based Diagnostic Evaluation

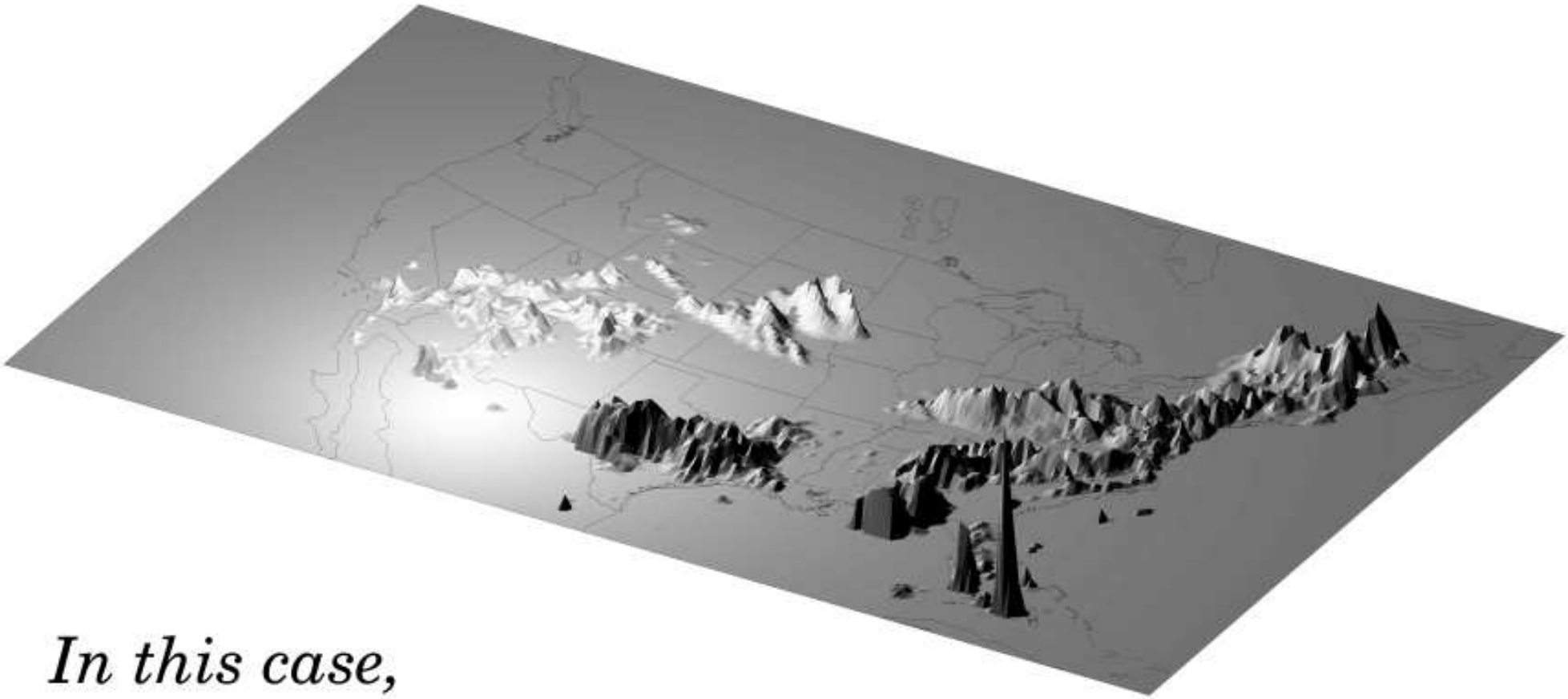


Raw Field



Objects

Step # 1: Raw Data

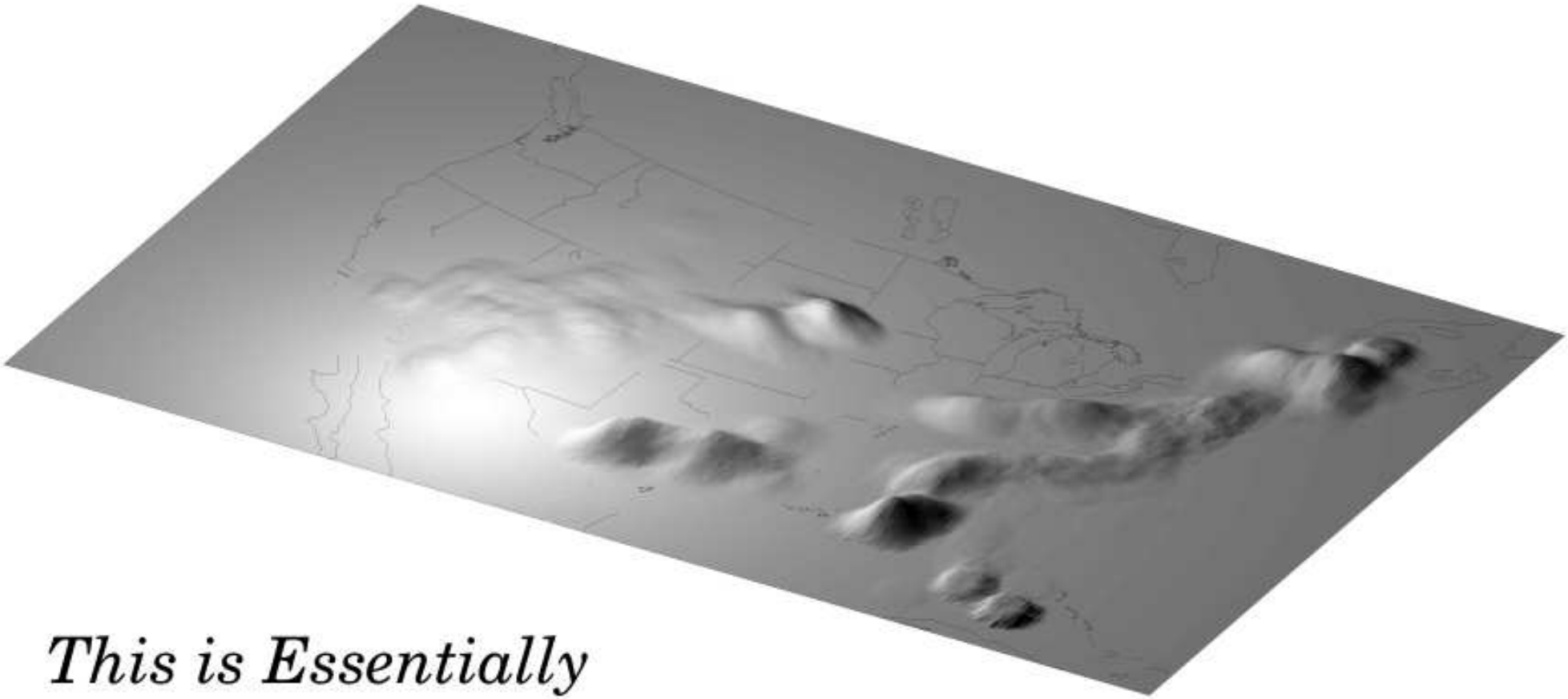


In this case,

Precipitation Data over the

Continental United States

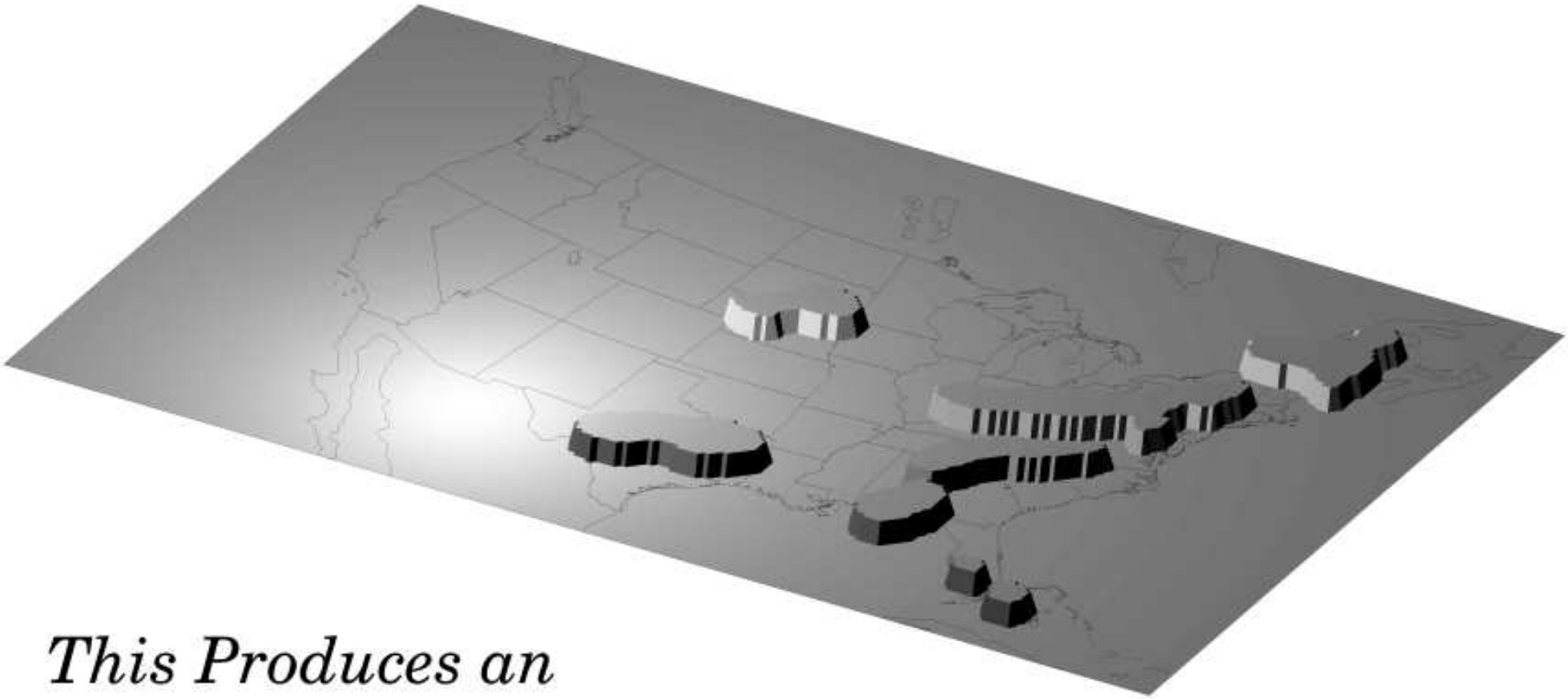
Step # 2: Convolution



This is Essentially

a Smoothing Operation

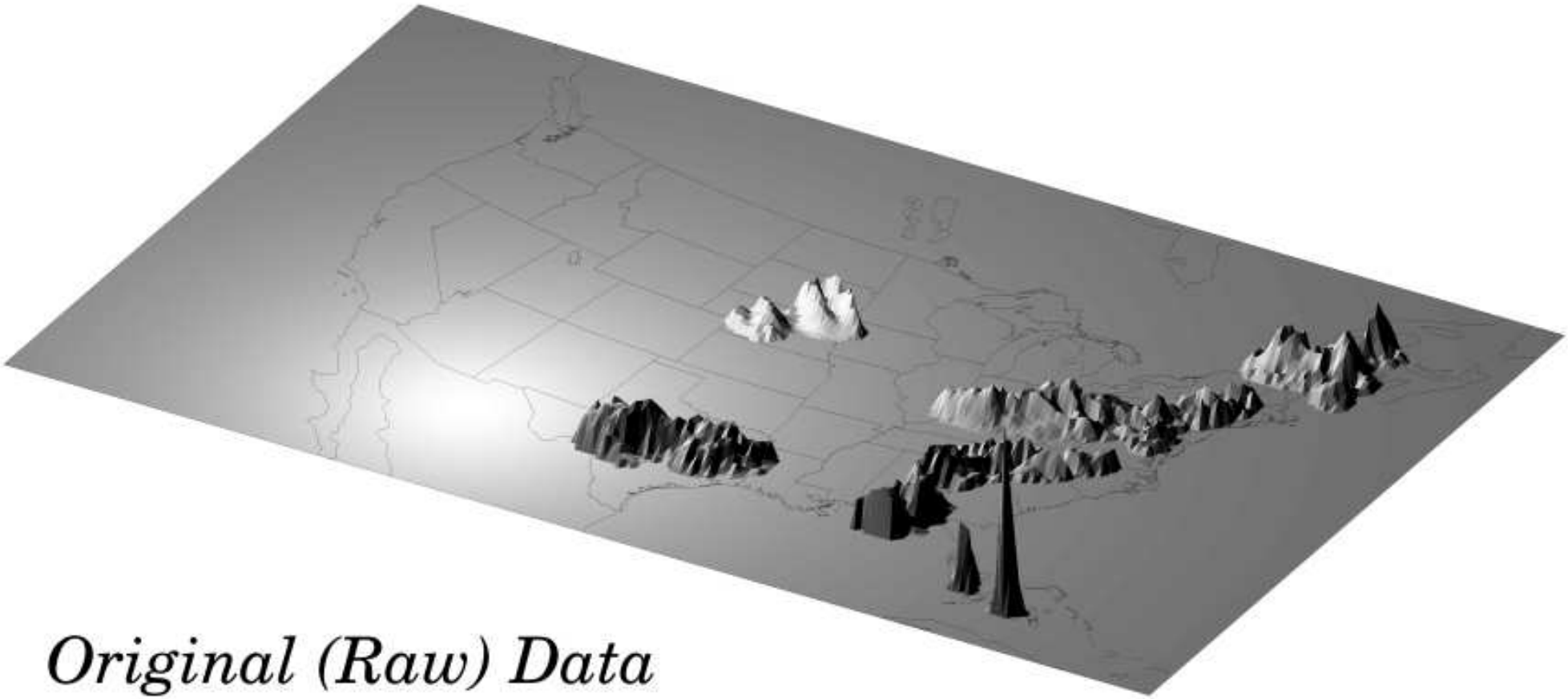
Step # 3: Thresholding



This Produces an

On / Off Mask Field

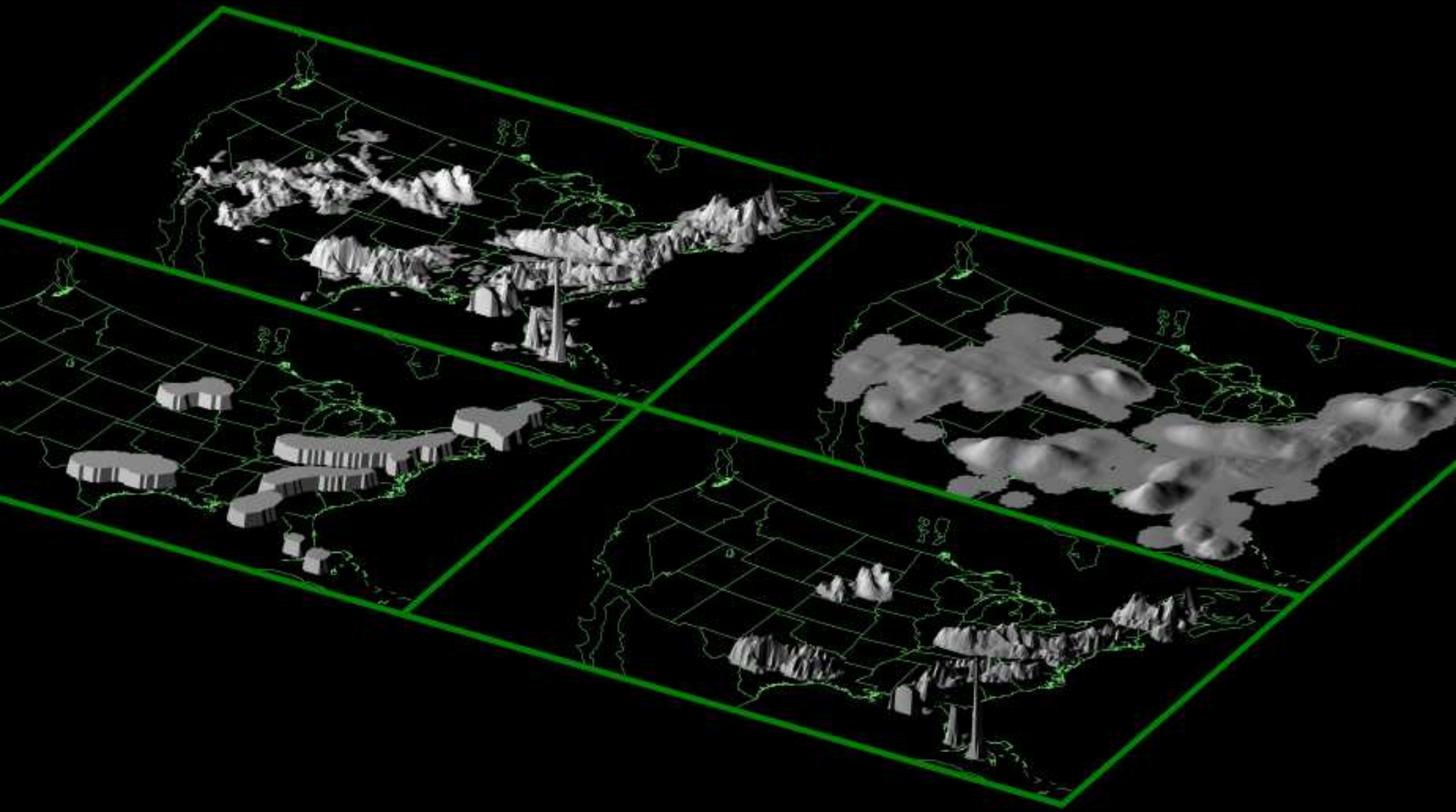
Step # 4: Restoration



Original (Raw) Data

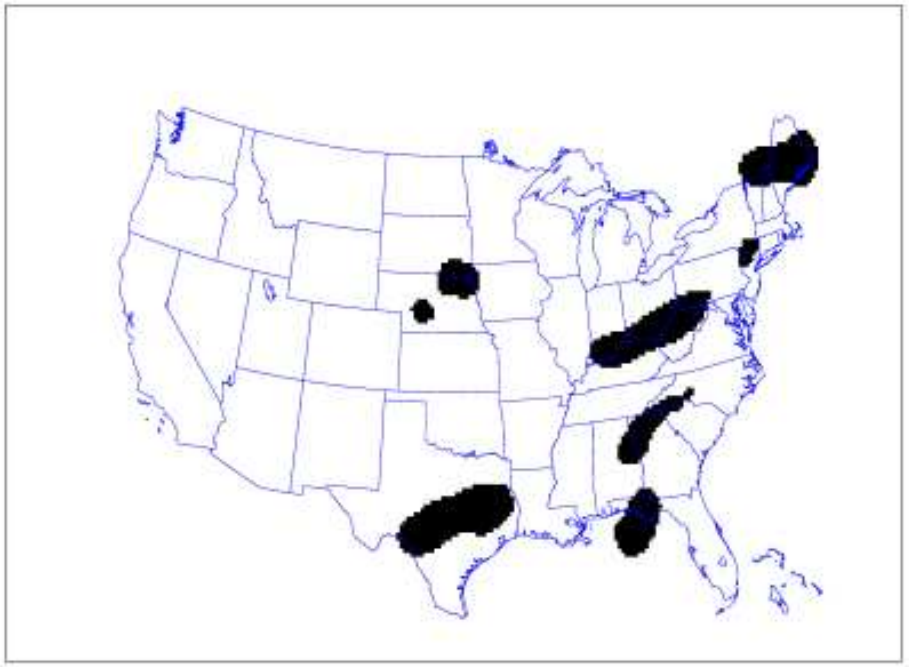
is Restored to

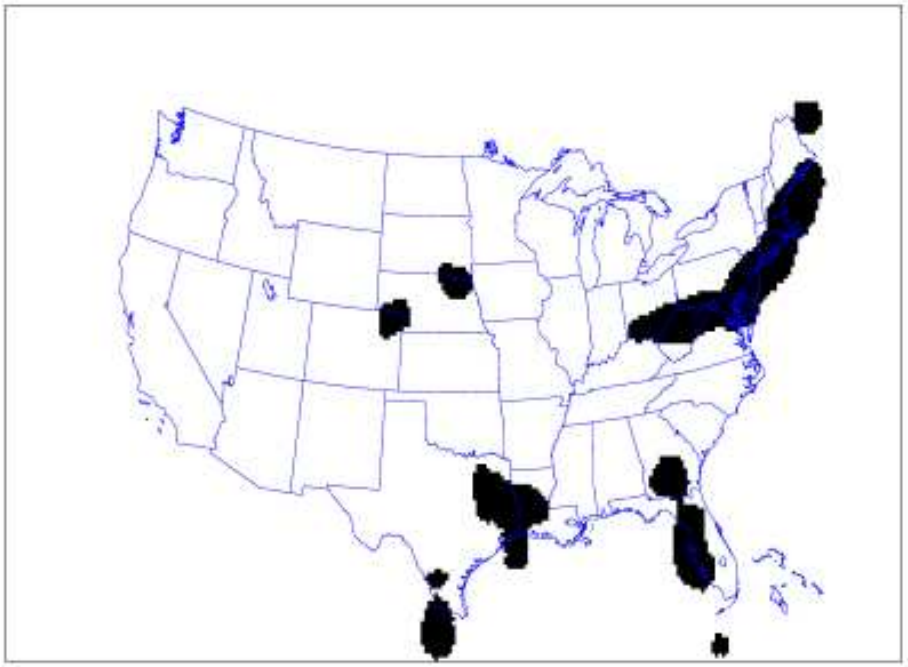
Object Interiors

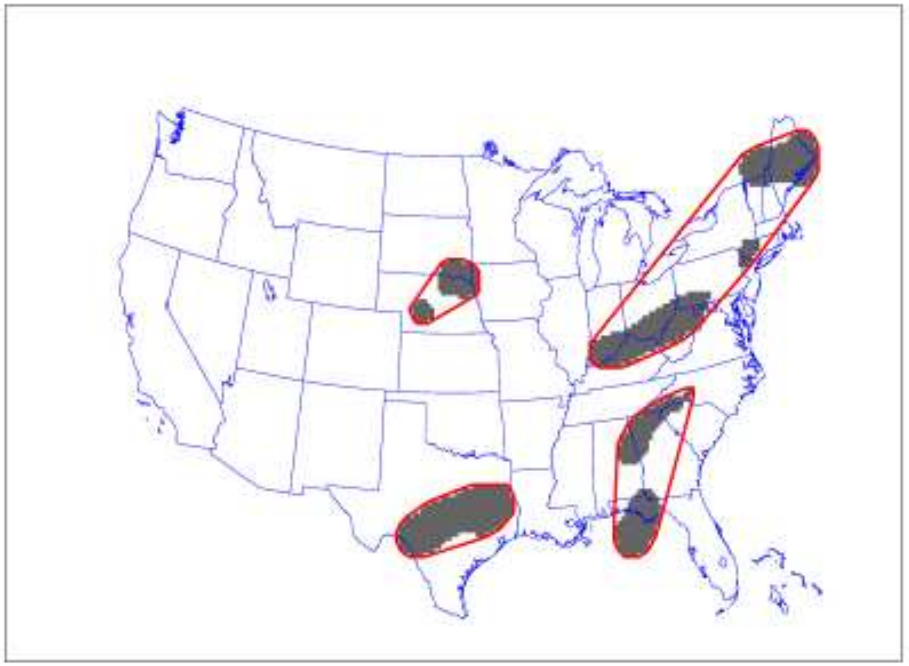


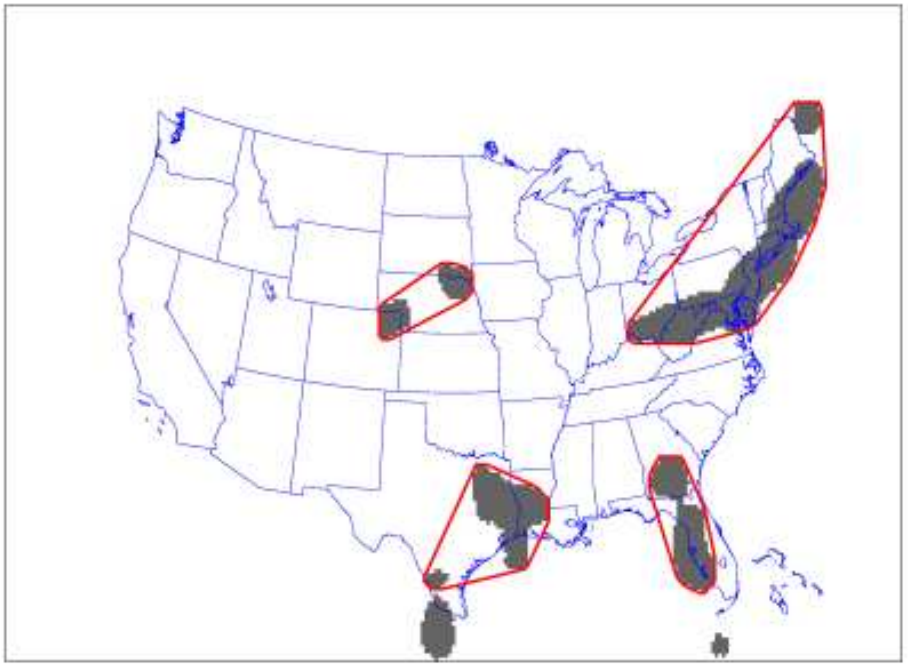
Forecast

Observed









Object Attributes

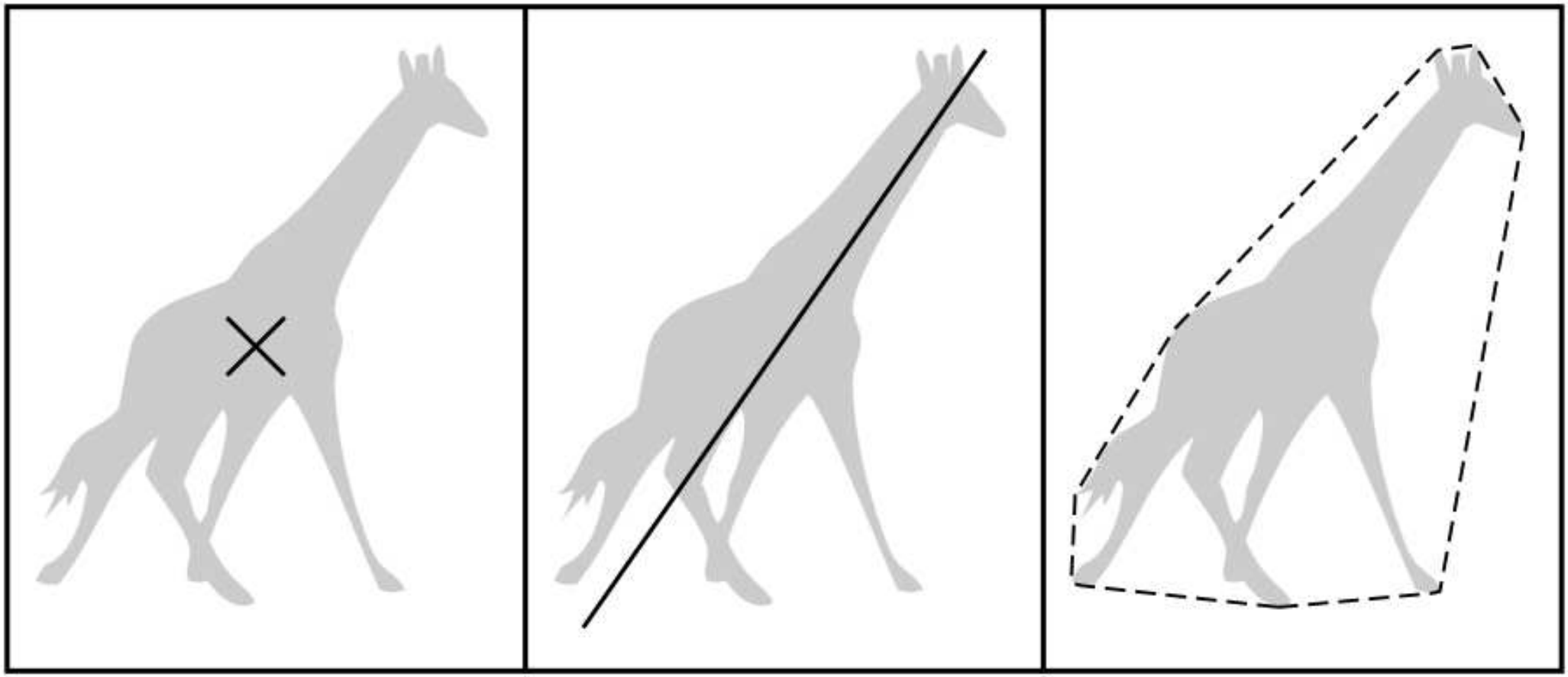
Simple or Composite

Single

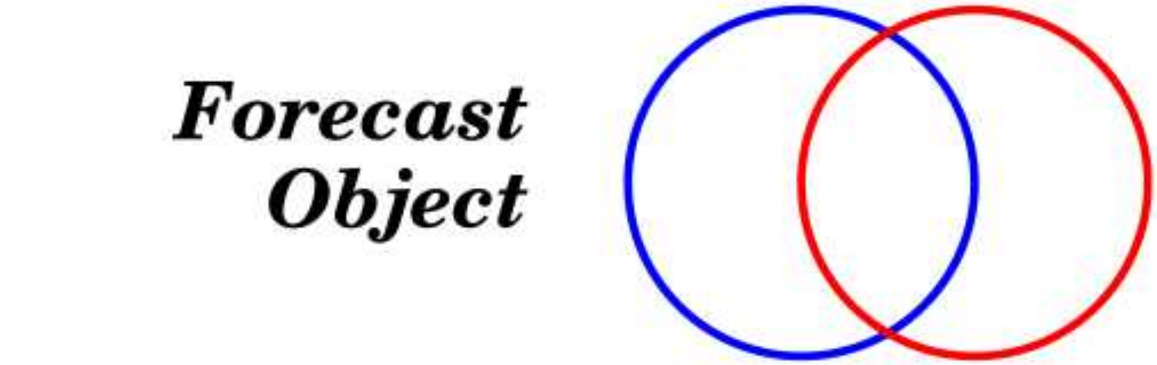
- *Area*
- *Centroid*
- *Axis Angle*
- *Angle Confidence*
- *Median Intensity*

Pair

- *Intersection*
- *Union*
- *Centroid Distance*
- *Angle Difference*
- *Area Ratio*
- *Intensity Ratio*

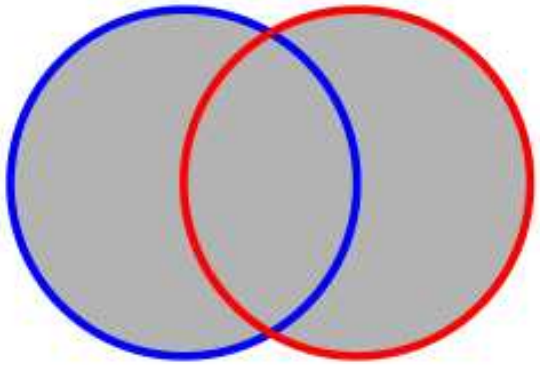


Centroid *Axis* *Convex Hull*

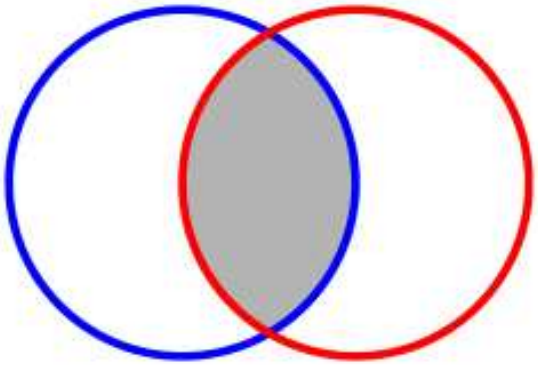


Forecast Object

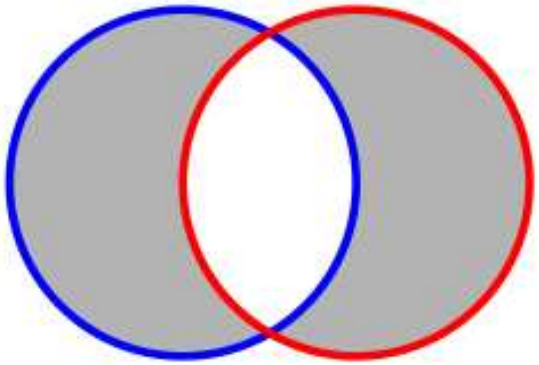
Observed Object



Union



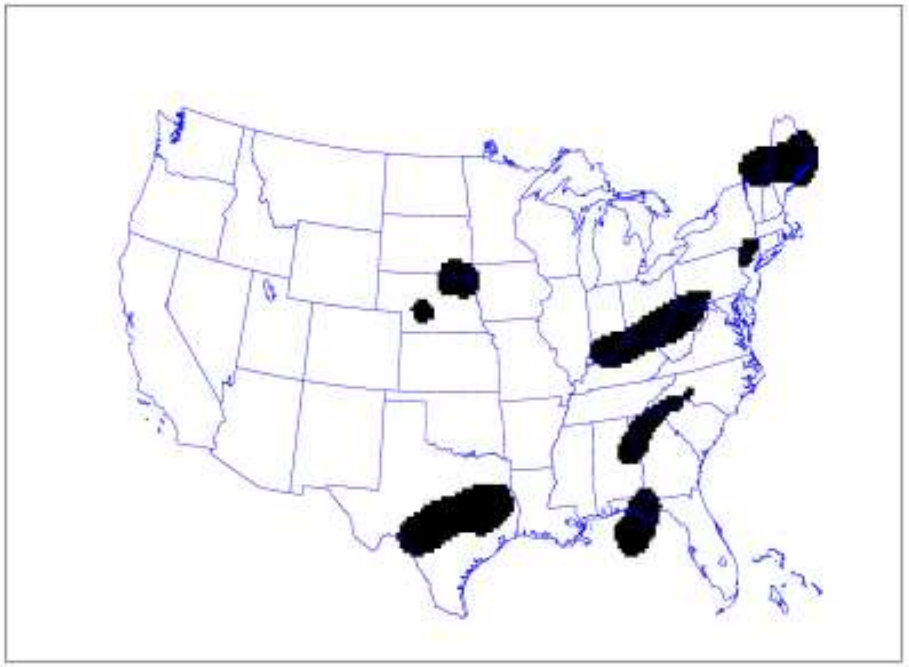
Intersection

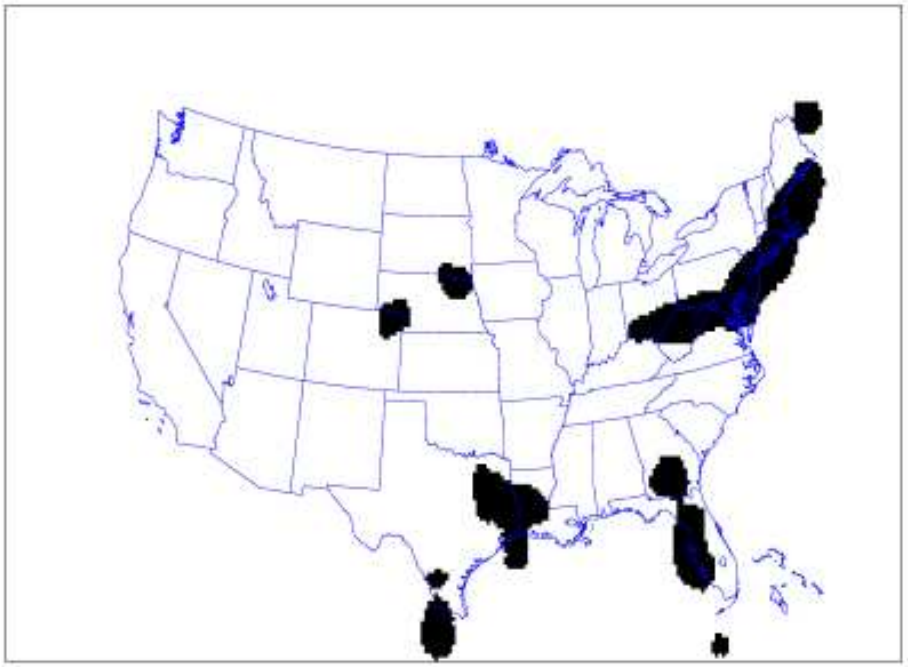


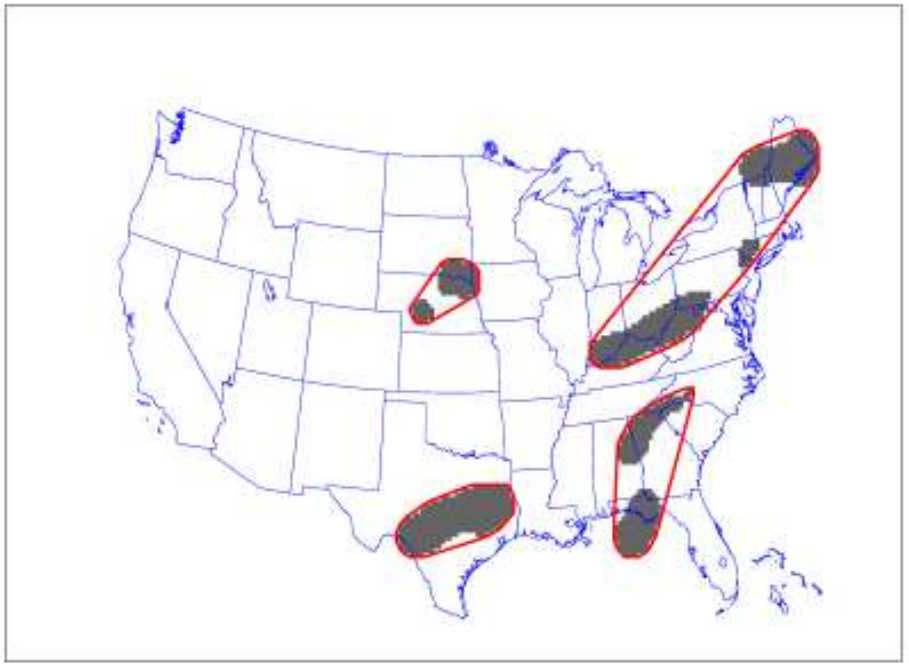
Symmetric Difference

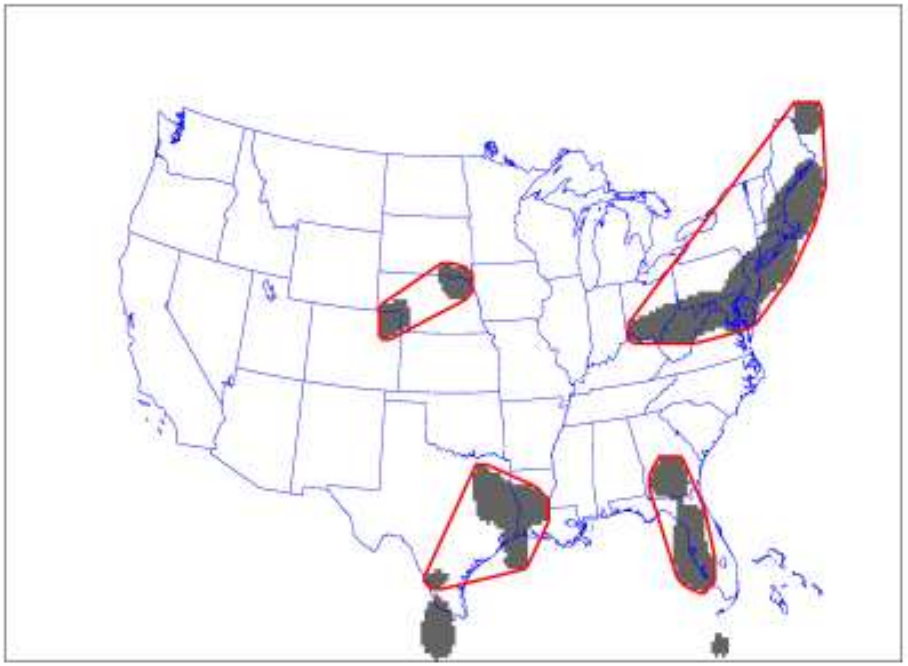
Forecast

Observed

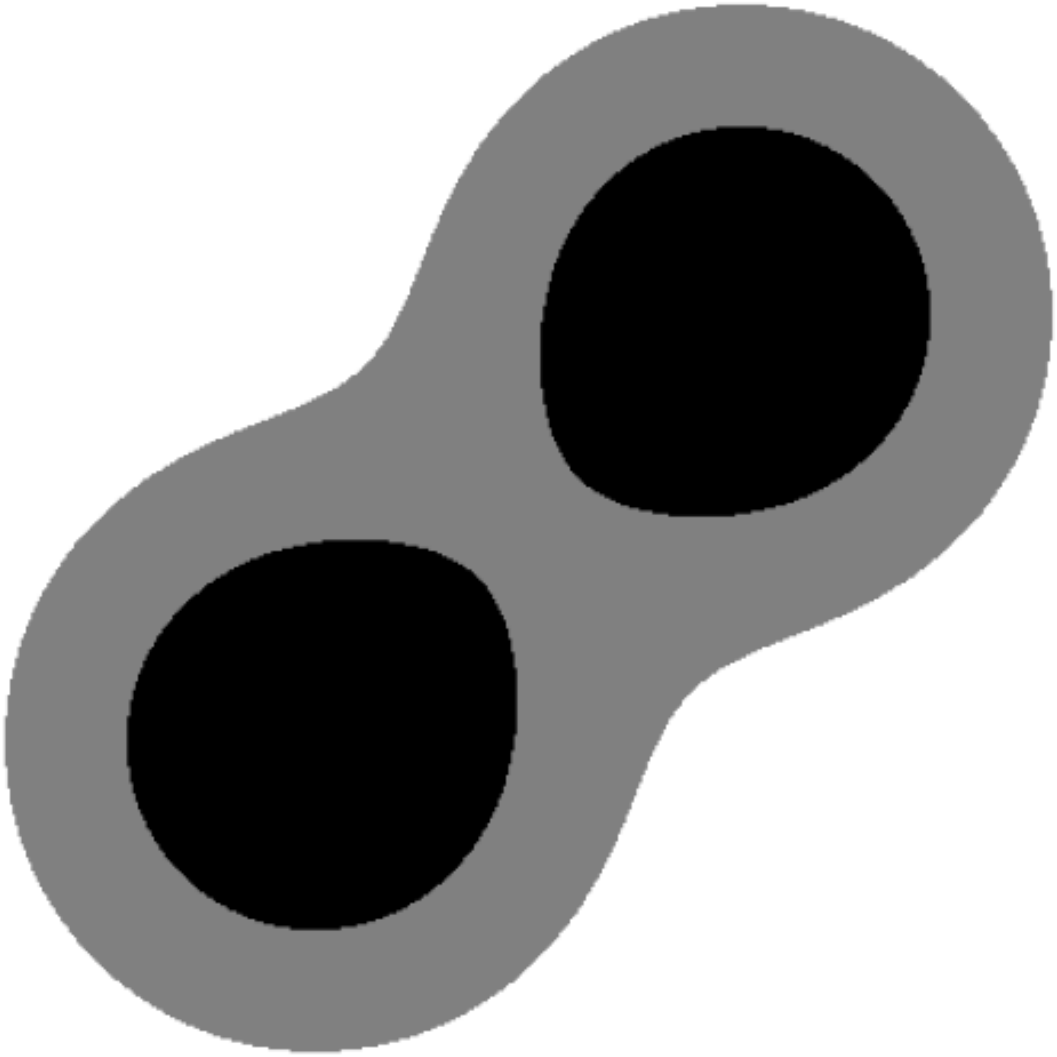






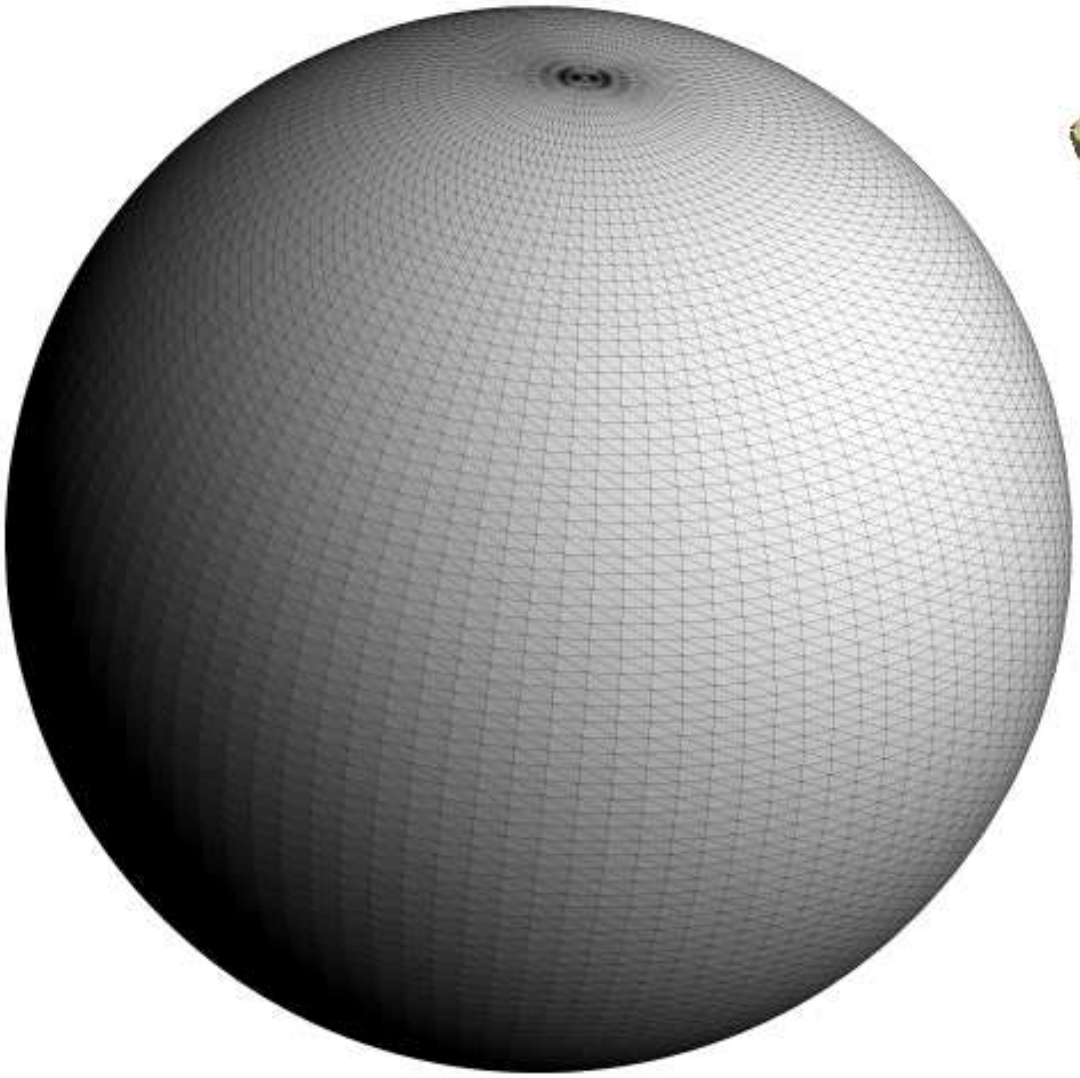


Alternative Merging Method



Double Thresholding

Space ... the Final Frontier



3D Objects

Practical: MODE Tool

- Functionality:
  - Identifies objects in two fields, computes single object attributes and pair-wise differences, matches/merges objects, writes object attributes and differences.
  - One time step at a time.
  - Data must reside on a common grid.
    - Recommend COPYGB for regridding GRIB data.
  - Configuration file specifies:
    - Forecast and observation specified separately
      - Field/level, raw filter value, mask region (grid, polyline)
      - Object definition parameters (convolution radius and threshold)
      - Object filtering parameters (area, intensity)
    - Flags for matching/merging logic
    - Fuzzy engine weights, interest functions, and confidence maps
    - Total interest threshold
    - PostScript plotting options
- Data formats
  - Reads gridded GRIB and NetCDF output of PCP-Combine.
  - Writes ASCII statistics, NetCDF object fields, PostScript summary plot.

Practical: MODE Usage

- Command Line
  - Required: fcst_file, obs_file, config_file
  - Optional: -config_merge, -outdir, -v
  - Disable output: -plot, -obj_plot, -obj_stat, ct_stat

Shell - Konsole <2>

Session Edit View Bookmarks Settings Help

```
[johnhg@billiken]% ./mode

Usage: mode
        fcst_file
        obs_file
        config_file
        [-config_merge merge_config_file]
        [-outdir path]
        [-plot]
        [-obj_plot]
        [-obj_stat]
        [-ct_stat]
        [-v level]

        where   "fcst_file" is a forecast file in either Grib or netCDF format (output of pcp_combine) containing the field to be verified (required).
                "obs_file" is an observation file in either Grib or netCDF format (output of pcp_combine) containing the verifying field (required).
                "config_file" is a WrfModeConfig file containing the desired configuration settings (required).
                "-config_merge merge_config_file" overrides the default fuzzy engine settings for merging within the fcst/obs fields (optional).
                "-outdir path" overrides the default output directory (/d1/johnhg/MET/MET_releases/METv1.1beta6/out/mode) (optional).
                "-plot" disables plotting (optional).
                "-obj_plot" disables the output of the object split and composite fields to a NetCDF file (optional).
                "-obj_stat" disables the output of the object statistics file (optional).
                "-ct_stat" disables the output of the contingency table standard statistics file (optional).
                "-v level" overrides the default level of logging (1) (optional).

        NOTE: The forecast and observation fields must be on the same grid.

[johnhg@billiken]%
```

Shell

Practical: MODE Output

- Output files:
  - ASCII object statistics file (File ends with "_obj.txt").
  - ASCII Contingency table statistics file (File ends with "_cts.txt").
  - NetCDF object file (File ends with "_obj.nc").
  - PostScript summary plot (File ends with ".ps").
- Naming conventions:
  - mode_FFIELD_FLVL_vs_OFIELD_OLVL_HHMMSSL_YYYYMMDD_HHMMSSV_HHMMSSA [.obj.txt | _cts.txt | _obj.nc | .ps]
  - Ex: mode_APCP_12_SFC_vs_APCP_12_SFC_120000L_20050807_120000V_120000A_obj.txt

Practical: MODE Object Stats

- Four object statistics line types (contents of OBJECT_ID column):
  - Simple forecast and observation objects (FNNN and ONNN)
  - Pairs of simple objects (FNNN_ONNN)
  - Composite forecast and observation objects (CFNNN and CONNN)
  - Pairs of composite object (CFNNN_CONNN)
- Same number of columns for each line type (50 in total):
  - 18 header columns.
  - 20 columns applicable to SINGLE simple and composite line types.
  - 12 columns applicable to PAIRS of simple or composite line types.
  - Columns which do not apply to a given line type contain fill data (-9999)
- May be disabled using the –obj_stat command line argument.

Practical: MODE CTStats

- Contains traditional contingency table counts and corresponding statistics computed in three ways:
  - Scoring the RAW fields by applying the convolution thresholds.
  - Scoring the FILTERed fields by first applying any raw filters and then applying the convolution thresholds.
  - Scoring point-wise using the resolved OBJECT fields.
- Meant simply as a point of reference for the MODE method.
- Differs from the VSDB CTS line type.
  - Does not include confidence intervals.
- May be disabled using the –ct_stat command line argument.

Practical: MODE NetCDF

- NetCDF output file contains 4 fields:
  - Indices for the simple forecast objects.
  - Indices for the simple observation objects.
  - Indices for the composite forecast objects.
  - Indices for the composite observation objects.
- May be disabled using the –obj_plot command line argument.

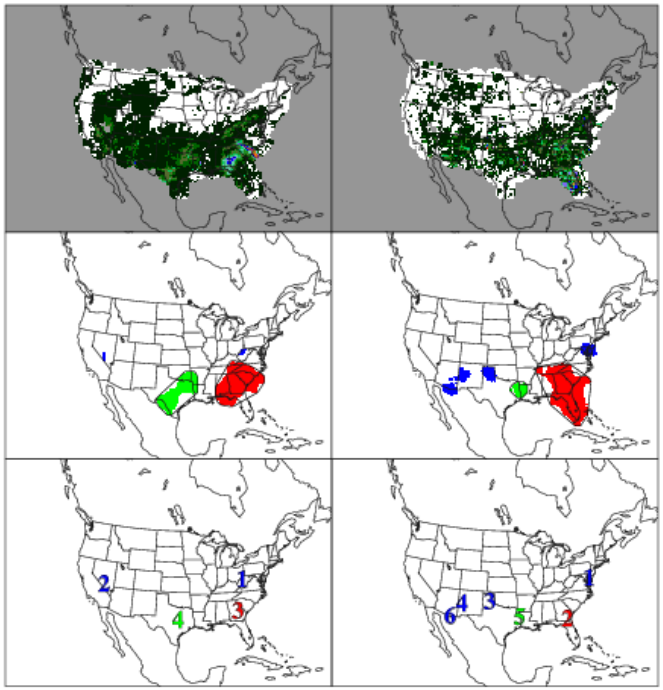
Practical: MODE PostScript

- MET does not generally provide plotting tools.
- Exception for MODE to illustrate the method.
- Configuration file plotting options:
  - Specify colortables to be used for plotting raw and object fields.
    - 61 colortables provided in data/colortables.
    - Specify how to rescale an existing colortable.
    - Or explicitly define you own.
  - Option for how colorbar is plotted.
  - Draw lines as great circle arcs or straight lines in the grid.
  - Zoom plot up to only the valid region of data.
- Number of pages of PostScript output based on configuration file selections:
  - At least 4 depicting object definition and matching.
  - Additional pages for:
    - Merging using the double-threshold technique.
    - Merging using the fuzzy engine technique.

MODE PS Pages 1 and 2

MODE: APCP_24 at SFC vs APCP_24 at SFC

Forecast | Observation

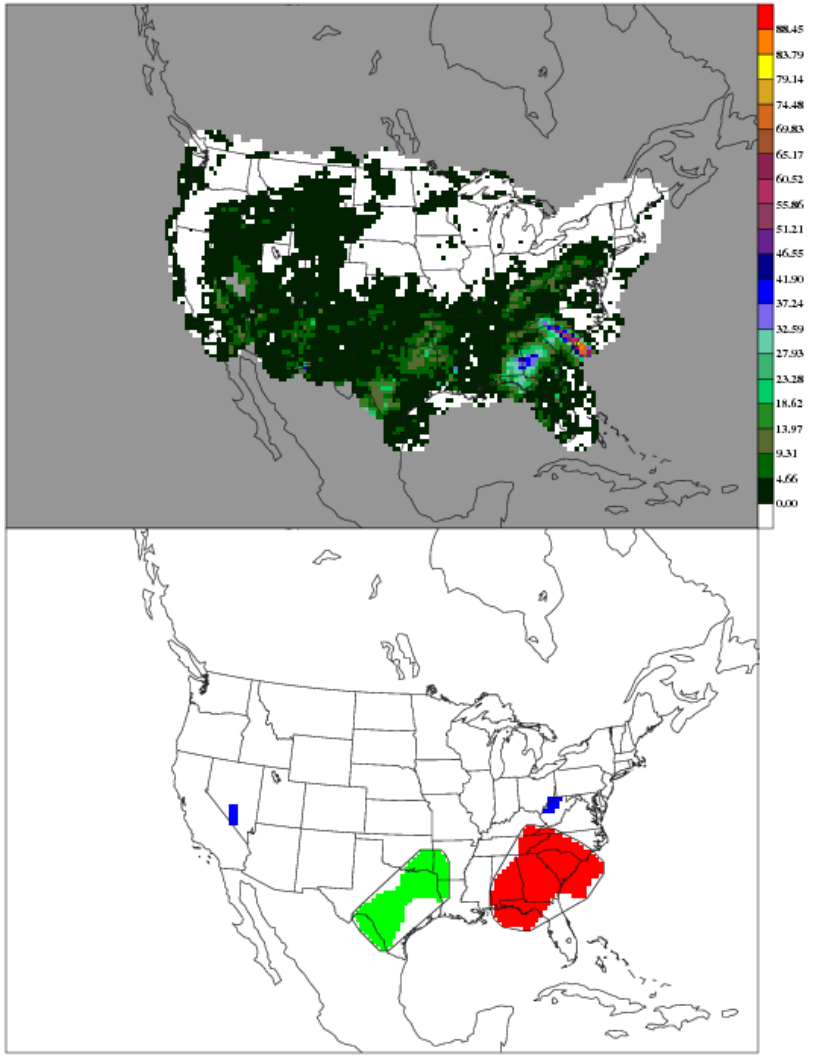


Fcst	Obs	Interest
3	2	0.9410
4	5	0.8883
1	1	0.6204
4	4	0.6024
3	5	0.5903
3	1	0.5875
4	3	0.5842
2	4	0.5701
2	6	0.5600
4	2	0.5574
1	2	0.5548
4	6	0.5508
1	5	0.4282
2	3	0.4261
3	3	0.4130
4	1	0.3446
3	4	0.2962
1	3	0.2765
2	5	0.1920
2	1	N/A
1	6	N/A
2	2	N/A
3	6	N/A
1	4	N/A

			Forecast	Observation
Model Name:	WRF	Mask Missing:	on	on
Init Time:	Aug 7, 2005 00:00:00	Mask Gr/Ply:	off/off	off/off
Valid Time:	Aug 8, 2005 00:00:00	Raw Thresh:	>=0.000 kg/m^2	>=0.000 kg/m^2
Lead Time:	24:00:00	Conv Radius:	5 gs	5 gs
Accum Time:	24:00:00	Conv Thresh:	>=5.000 kg/m^2	>=5.000 kg/m^2
Centroid Distance:	2.00	Area Thresh:	>=10.000 gs	>=10.000 gs
Boundary Distance:	4.00	Inten Perc:	100 th	100 th
Convex Hull Distance:	0.00	Inten Thresh:	>=0.000 kg/m^2	>=0.000 kg/m^2
Angle Difference:	1.00	Merge Thresh:	>=1.250 kg/m^2	>=1.250 kg/m^2
Area Ratio:	1.00	Merging:	none	none
Intersection/Area:	2.00	Matching:	match/merge	
Complexity Ratio:	0.00	Simple(M/U):	4 (2/2)	6 (2/4)
Intensity Ratio:	0.00	Composites:	2	2
Total Interest Thresh:	0.70			

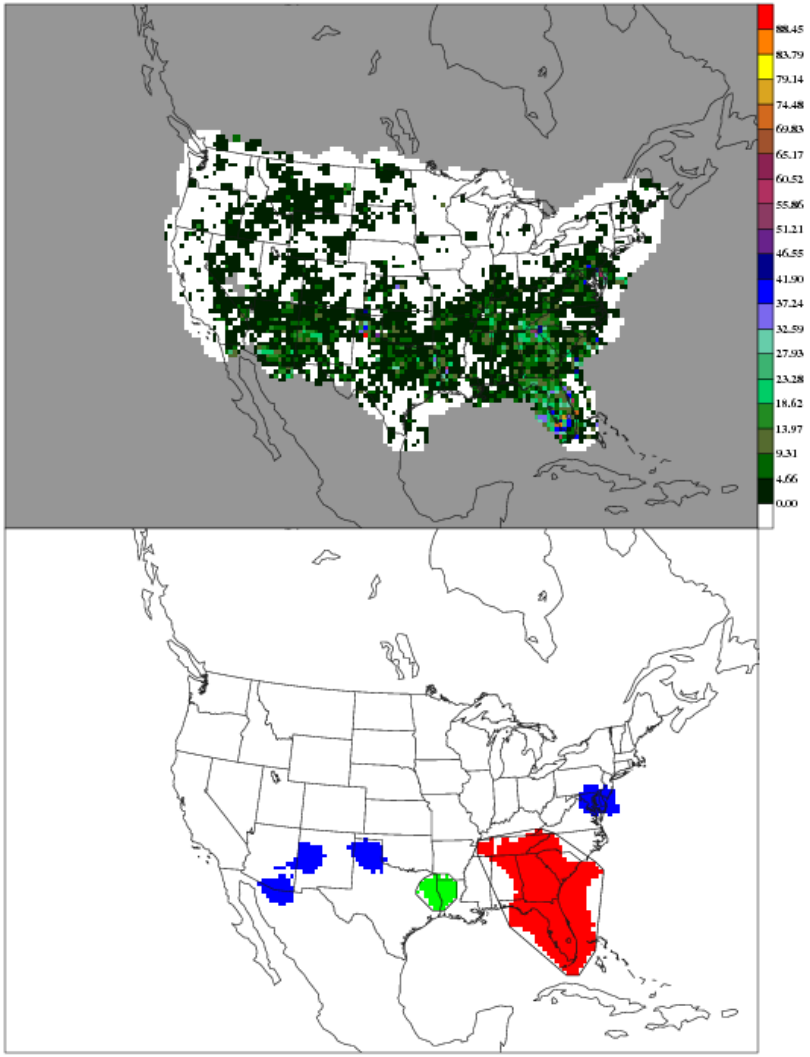
MODE: APCP_24 at SFC vs APCP_24 at SFC

Forecast

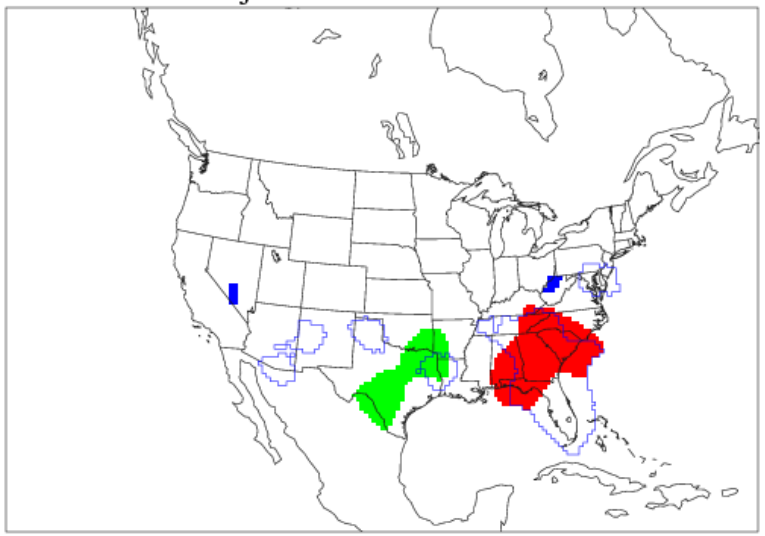


MODE PS Pages 3 and 4

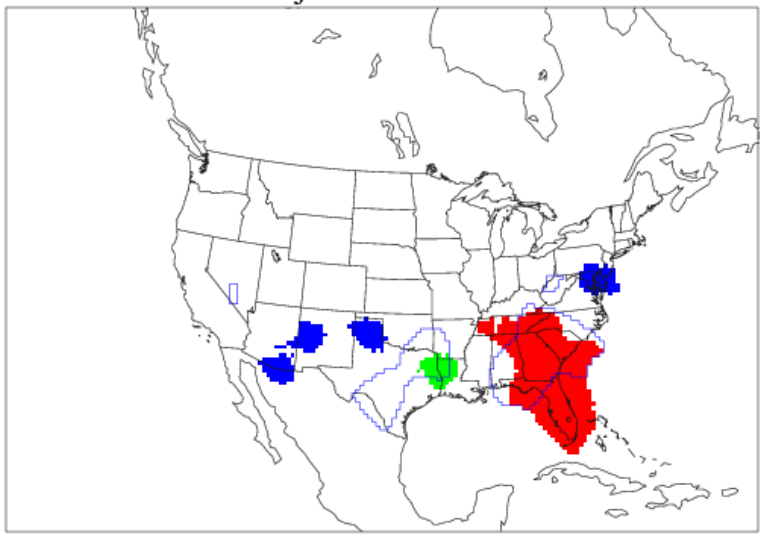
MODE: APCP_24 at SFC vs APCP_24 at SFC
Observation



Forecast Objects with Observation Outlines

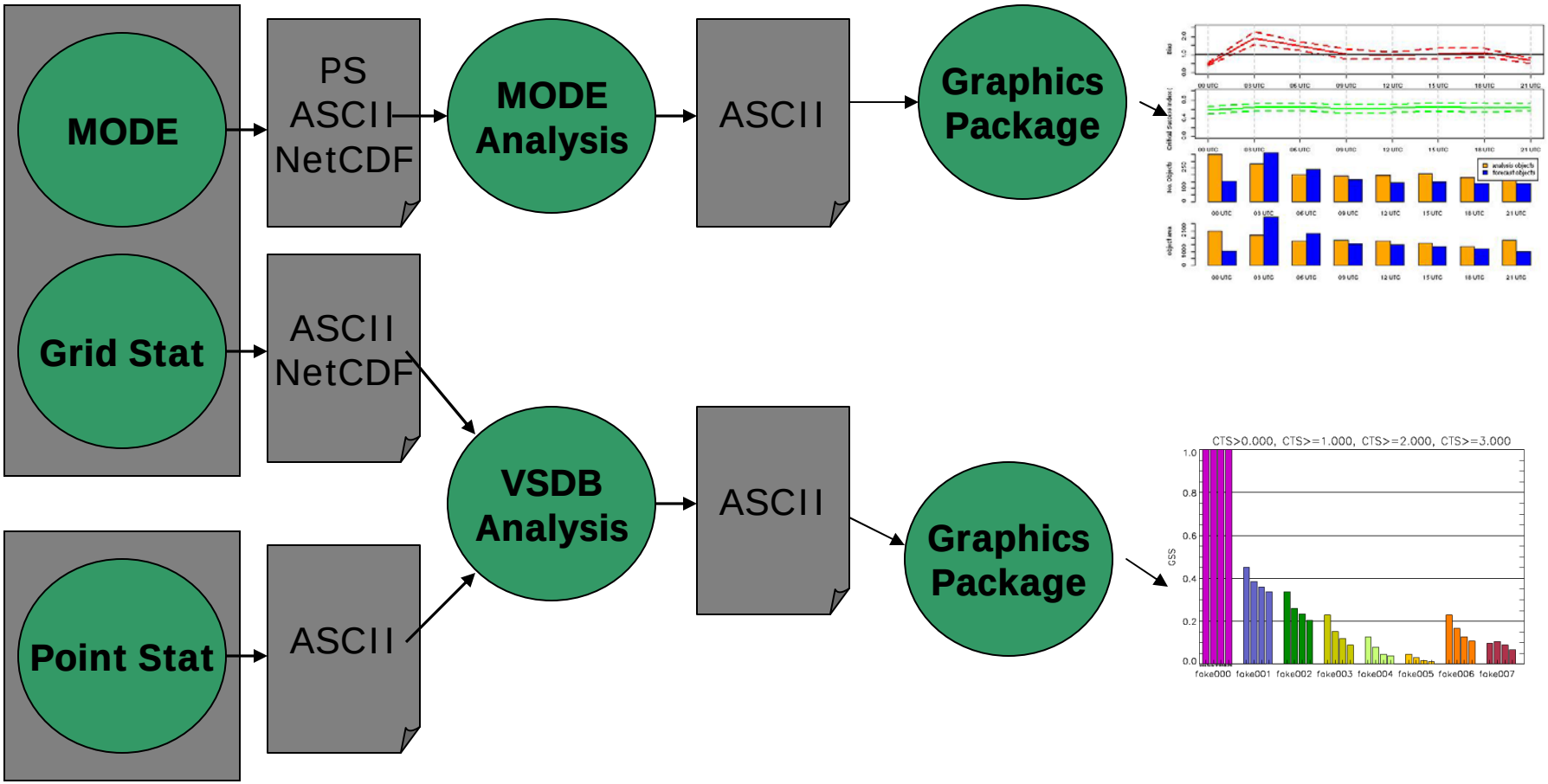


Observation Objects with Forecast Outlines

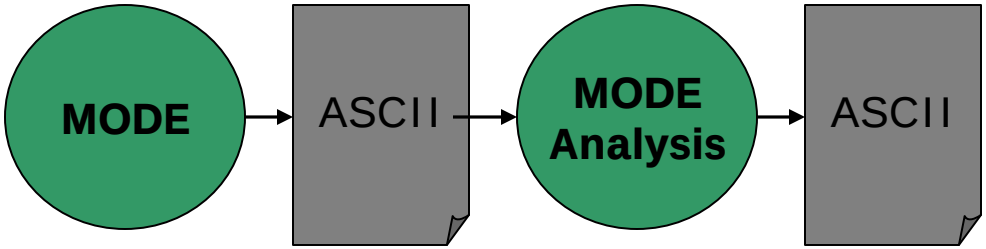


MET Analysis tools and examples

→STATS→OUTPUT→AGGREGATE→ User-defined display



MODE Analysis - input



- Usage:

mode_analysis -lookin path -summary | -bycase [-column name] [-dump_row filename] [-out filename] [-help] [MODE FILE LIST] [-config config_file] | [MODE LINE OPTIONS]

- example:

mode_analysis -lookin ./mode_output -summary -column area -column intensity_90 -column centroid_lat -column axis_ang -single -simple -obs_thr ge0.3 -area_min 1000

MODE Analysis - output

output of previous command:

```
Total mode lines read =  460,802
Total mode lines kept =       61
       Field    N       Min       Max      Mean    StdDev       P10       P25       P50       P75       P90       Sum
------------   --   -------   -------   -------   ------   -------   -------   -------   -------   -------   --------
        area   61   1002.00   3835.00   1613.34   836.77   1017.00   1081.00   1356.00   1624.00   3362.00   98414.00
intensity_90   61      8.89     54.97     20.10    11.45     13.21     13.46     16.51     20.07     40.64    1226.11
centroid_lat   61     34.32     41.86     38.35     2.75     34.37     35.38     39.93     40.51     41.03    2339.45
    axis_ang   61    -66.30     89.51     33.72    52.33    -55.66      0.66     59.94     74.81     82.56    2057.17
```

with slightly different options: `-summary -pair -simple -model wrf4ncep`

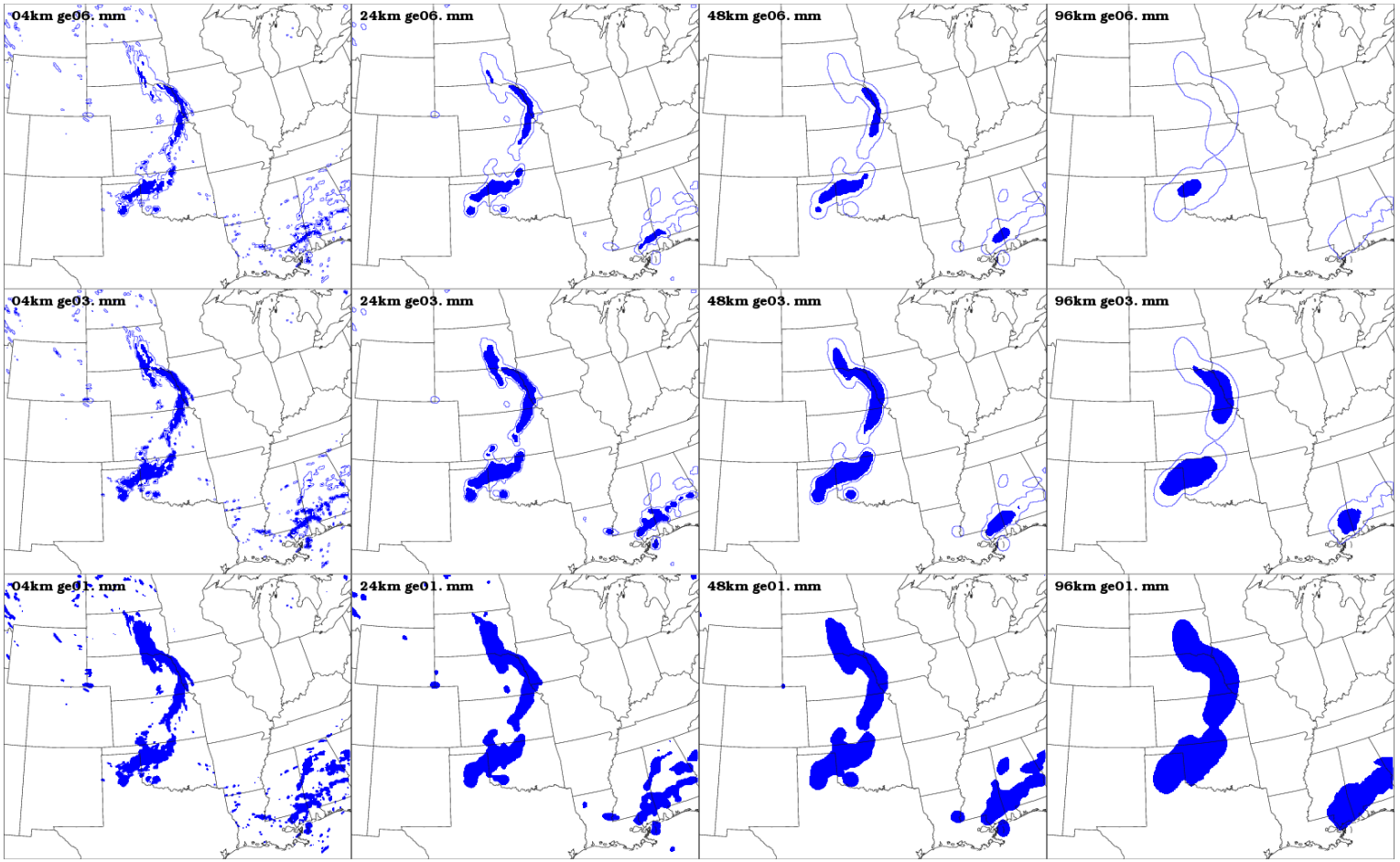
```
Total mode lines read =  460,802
Total mode lines kept =       94
     Field    N    Min     Max    Mean    StdDev    P10     P25     P50     P75     P90       Sum
----------   --   ----   -----   -----   ------   ----   -----   -----   -----   -----   -------
angle_diff   94   0.00   77.45   27.33    18.60   6.26   13.59   26.06   34.36   52.20   2569.32
  interest   94   0.90    1.00    0.92     0.02   0.91    0.91    0.92    0.93    0.96     86.95
```

with `-bycase` option:

```
Total mode lines read =  460,802
Total mode lines kept =   20,851
      Fcst Valid Time   Area Matched   Area Unmatched   # Fcst Matched   # Fcst Unmatched   # Obs Matched   # Obs Unmatched
---------------------   ------------   --------------   --------------   ----------------   -------------   ---------------
Apr 26, 2005  00:00:00          20989            10087               75                103              29               928
May 13, 2005  00:00:00          79894            40067              157                137              41              1036
May 14, 2005  00:00:00         114407            21398              186                 84             125              3291
```

MODE Analysis – by radius,threshold

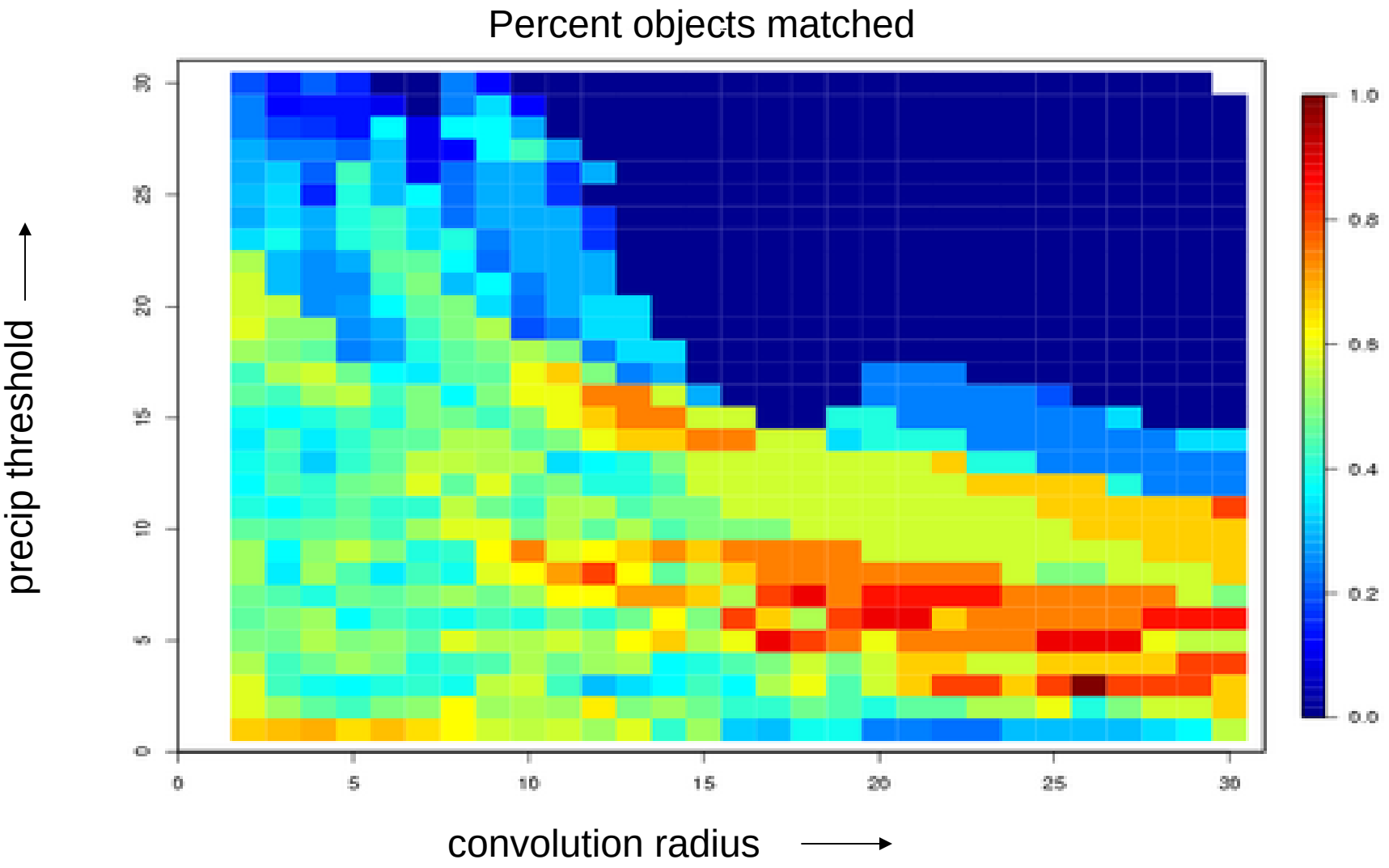
1-h precipitation forecast, 24-h lead time, valid 1 June 2005



precip threshold ⟶

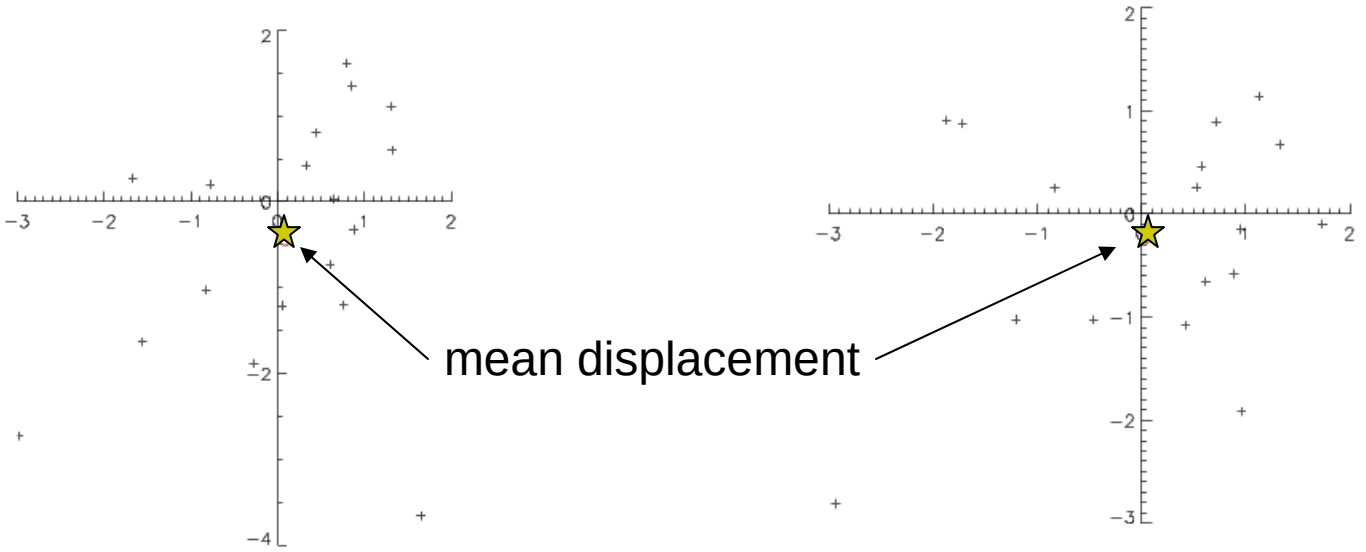
convolution radius ⟶

MODE Analysis – quilt plot



MODE object displacment analysis

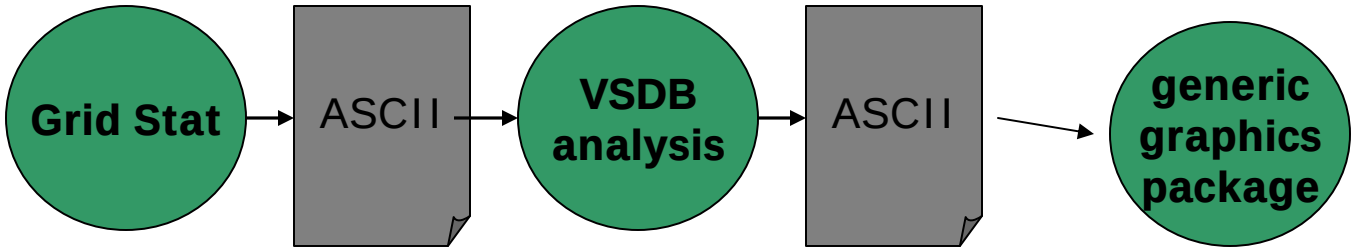
- 9 cases from 2005
  - 48 km convolution radius, 3 mm rain threshold
- displacement of composite forecast objects from matched observed objects (in degrees)

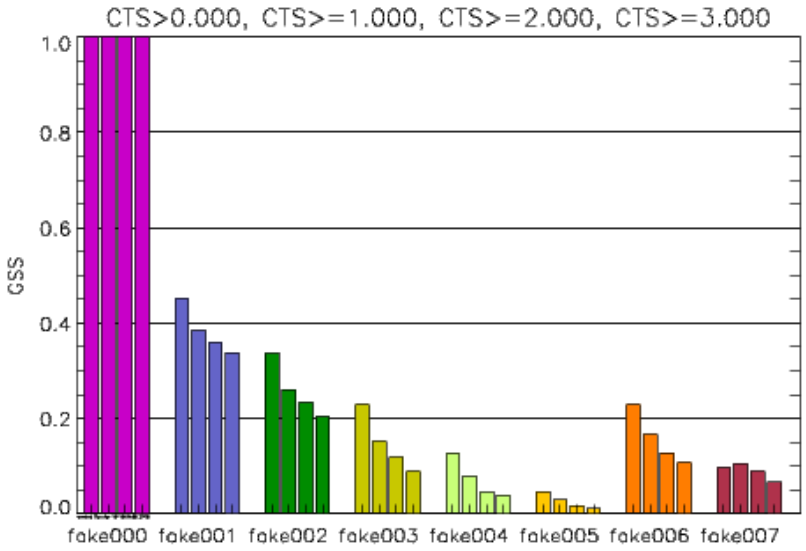


wrf2caps model

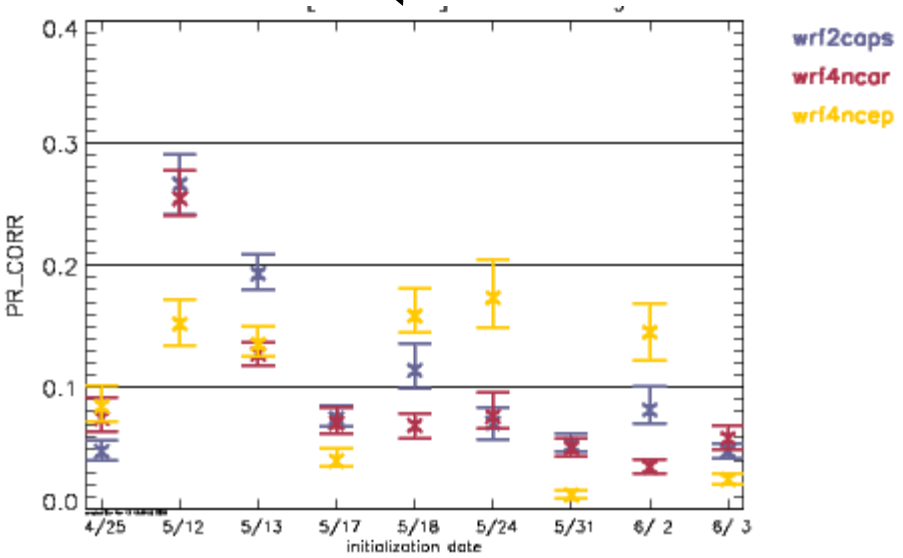
wrf4ncar model

VSDB Analysis





Gilbert skill score for 7 geometric test cases with 4 rain thresholds



Pearson correlation coefficient for 9 cases from 2005, 3 different models with 95% bootstrap confidence intervals

MET plans for the future

- New input formats
  - Grib2
  - More general NetCDF format
- New output
  - New analysis tools
  - Matched pairs
  - User-supplied graphics tools
- MET GUI
  - Automatic creation of configuration files and execution statements
- Expand verification methods
- Inclusion in a verification service developed by NOAA/GSD

MET plans for the future – verification methods

- Additional spatial approaches
  - Scale separation - Intensity-scale (Casati)
  - Contiguous rain area (Ebert/McBride)
  - Image warping (Keil and Craig; Lindstrom and Gilleland)
- Ensemble and probability forecast methods
- Methods for extremes
- Multi-model comparisons
- Cyclone track verification
- Other? Depends on the community…

What goes into MET? Community contributions

- DTC verification workshops
  - 2007, 2008 (http://www.ral.ucar.edu/~ericg/dtcworkshop.html)
  - *Goal*: Obtain guidance from verification and model evaluation experts
    - Current State-of-the-art
    - New methods, ready for implementation
- New verification methods
- Approaches for display of verification information – using your favorite graphics package

What goes into MET?

- Verification Advisory Group (VAG)
  - Establish approaches for determining if new methods are “ready” and are useful
- Verification testbed
  - Standard datasets for evaluation of methods
- Desires of MET and WRF users
  - Send feedback to bgb@ucar.edu or met_help@rap.ucar.edu

How to learn more…

- MET help: met_help@ucar.edu
- MET home page
  - http://www.dtcenter.org/met/users/
- MET on-line tutorial
  - To be released soon… Keep an eye on the MET web page http://www.dtcenter.org/met/users/
- MET tutorial
  - Included in WRF tutorial – July 2008; January 2009
- DTC verification workshop presentation
  - http://www.ral.ucar.edu/~ericg/dtcworkshop.html
- Spatial verification methods – Intercomparison Project
  - http://www.ral.ucar.edu/projects/icp/index.html
- RAL verification page
  - http://www.rap.ucar.edu/research/verification/index.php
- WMO verification page
  - http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html