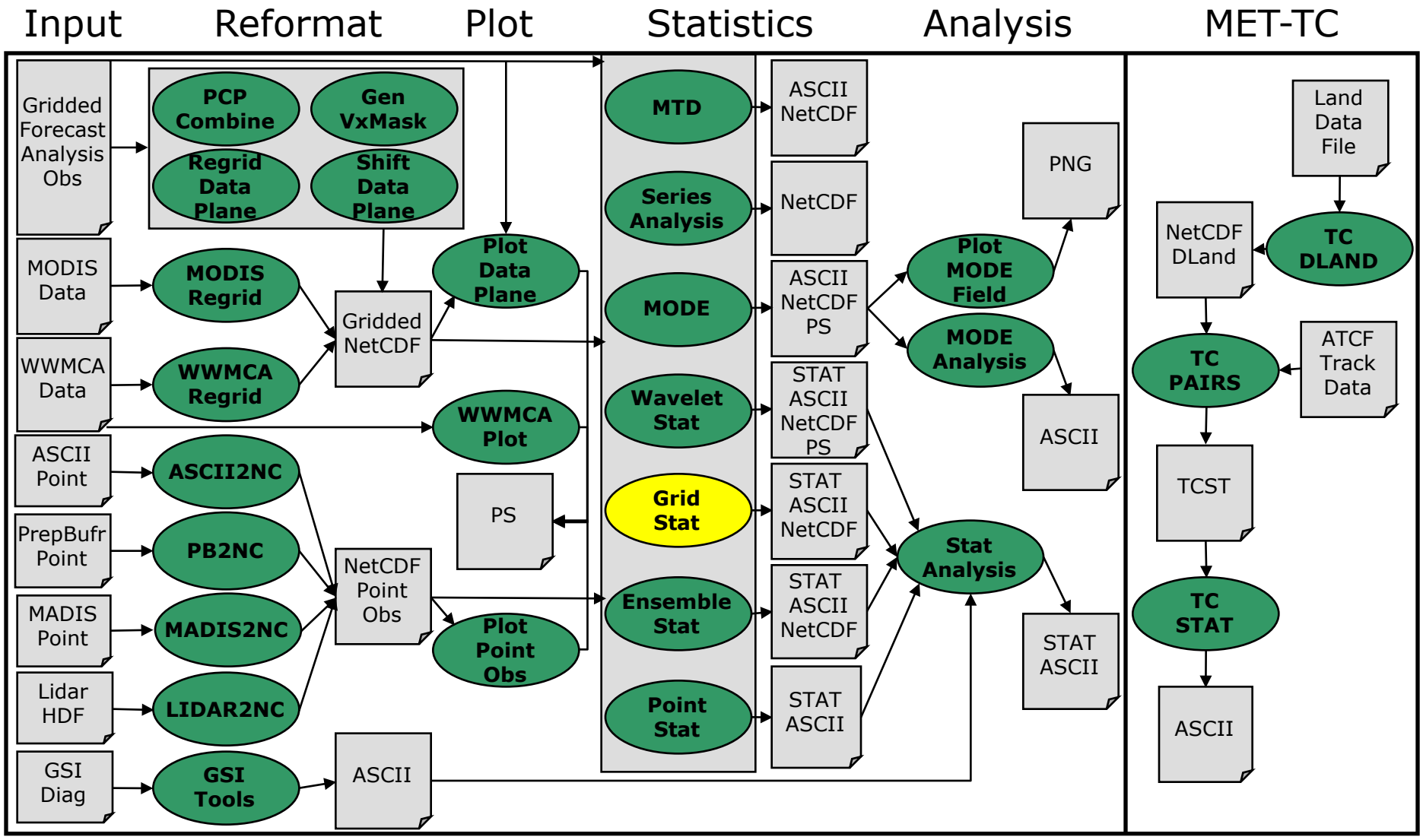
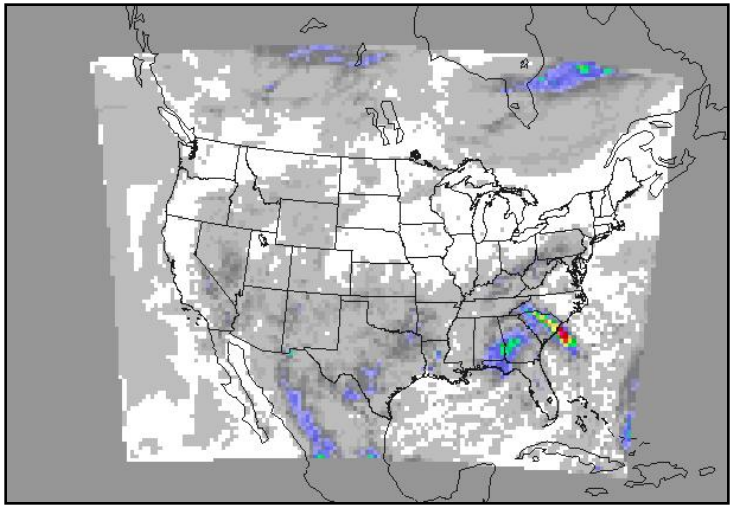
Grid-Stat Tool

Grid-Stat Tool

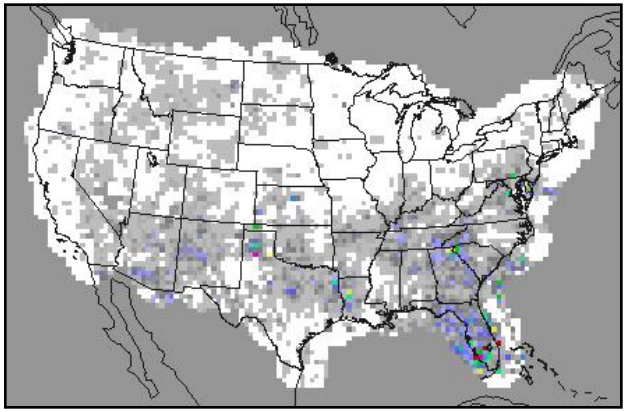


Grid-Stat: Overview



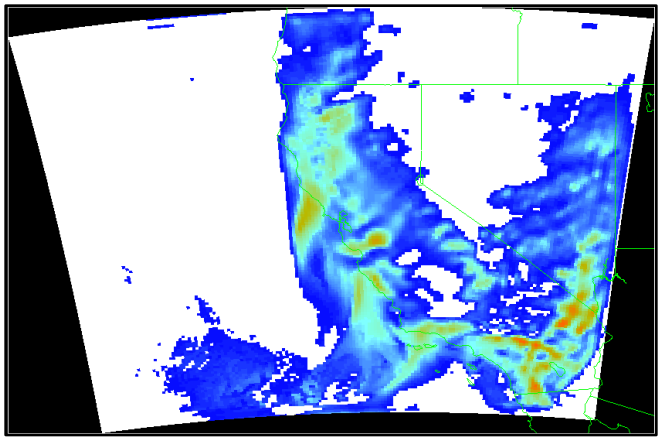
- Compare **gridded forecasts** to **gridded observations** on the same grid.
- Accumulate matched pairs over a defined area at a single point in time.
- Verify one or more variables/levels.
- Analysis tool provided to aggregate through time.

- Verification methods:
  - Continuous statistics for raw fields.
  - Single and Multi-Category counts and statistics for thresholded fields.
  - Parametric and non-parametric confidence intervals for statistics.
  - Compute partial sums for raw fields.
  - Methods for probabilistic forecasts.
  - Continuous statistics and categorical counts/statistics using neighborhood verification method.

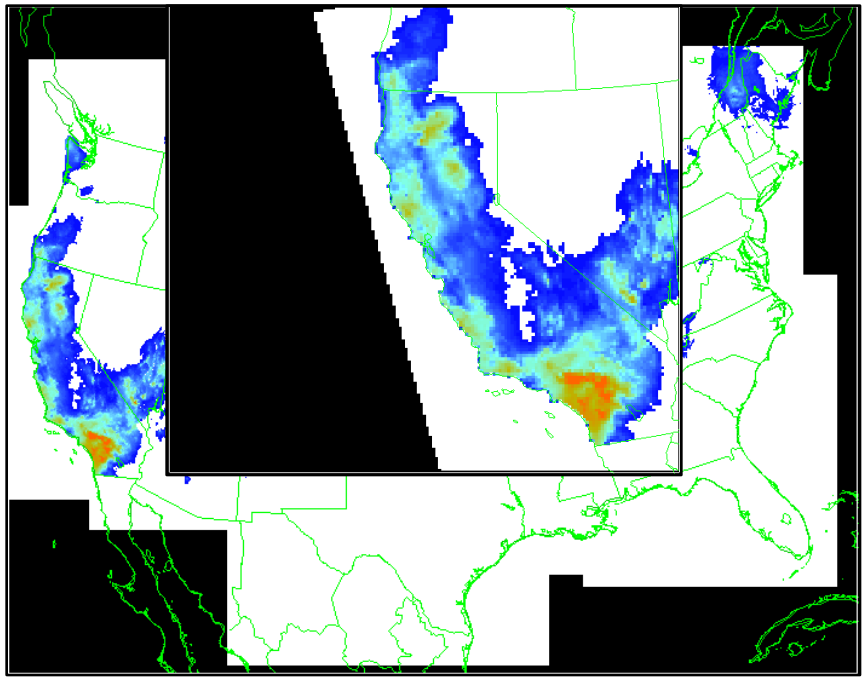


Grid-Stat: Common Grid

Model Forecast



StageIV Analysis



- Regrid the StageIV Analysis (GRIB) to the model domain:

```
copygb –xg"255 5 169 154 31357 -129770 8 -120500 10395 10395 0 64" \
        ST4.2010122212.06h ST4.2010122212.06h_regrid
```

- Automated regridding in configuration file.
- Practice running copygb in the practical session.

Grid-Stat: Input/Output

- Input Files
  - Gridded forecast and observation files
    - GRIB1 output of Unified Post-Processor (or other)
    - GRIB2 from NCEP (or other)
    - NetCDF from PCP-Combine, wrf_interp, or CF-compliant
  - ASCII configuration file
- Output Files
  - ASCII statistics file with all output lines (end with ".stat")
  - Optional ASCII files sorted by line type with a header row (ends with "_TYPE.txt")
  - Optional NetCDF matched pairs file

Grid-Stat: Usage

Usage: grid_stat

fcst_file

obs_file

config_file

[-outdir path]

[-log file]

[-v level]

fcst_file	Gridded forecast file
obs_file	Gridded observation file
config_file	ASCII configuration file
-outdir	Output directory to be used
-log	Optional log file
-v	Level of logging

Grid-Stat: Configuration

- Many configurable parameters – only set a few:
  - Precipitation accumulated over 24 hours.
    - GRIB1 forecast
    - NetCDF observation
  - Threshold any rain and moderate rain (mm).
  - Accumulate stats over all the points in the domain and just the eastern United States.
  - Compute neighborhood statistics with two sizes.
  - Generate all possible output types, except probabilistic.

```
fcst = {
 field = [
  {
   name       = "APCP";
   level      = [ "A24" ];
   cat_thresh = [ >0.0, >20.0 ];
  }
 ];
};
```

```
obs = {
 field = [
  {
   name       = "APCP_24";
   level      = [ "(*,*)" ];
   cat_thresh = [ >0.0, >20.0
  }
 ];
};
```

```
mask = {
 grid = [ "FULL" ];
 poly = [ "EAST.poly" ];
};
```

```
nbrhd = {
 vld_thresh = 1.0;
 width      = [ 3, 5 ];
 cov_thresh = [ >=0.5 ];
}
```

```
output_flag = {
 fho    = BOTH;
 ctc    = BOTH;
 cts    = BOTH;
 mctc   = BOTH;
 mcts   = BOTH;
 cnt    = BOTH;
 sl1l2  = BOTH;
 sal1l2 = BOTH;
 vl1l2  = BOTH;
 val1l2 = BOTH;
 vcnt   = NONE;
 pct    = NONE;
 pstd   = NONE;
 pjc    = NONE;
 prc    = NONE;
 eclv   = NONE;
 nbrctc = BOTH;
 nbrcts = BOTH;
 nbrcnt = BOTH;
 grad   = NONE;
};
```

Grid-Stat: Field name and level

- GRIB1 and GRIB2 files
  - **name = "GRIB Abbreviation";**
    - http://www.nco.ncep.noaa.gov/pmb/docs/on388/table2.html
    - TMP for Temperature, APCP for accumulated precipitation.
  - **level = ["string"];** Multiple values expand to multiple vx tasks
    - Level indicator followed by level value.
      - **A** for accumulation interval in HH[MMSS] format *(A06)*.
      - **P** for pressure level (*P500*) or layer (*P500-600*).
      - **Z** for vertical level (*Z2* or *Z10*).
      - **L** for generic level type (*L100*).
      - **R** for a specific GRIB record number (*R225*).
- Gridded NetCDF files
  - **name = "string";** Defines NetCDF variable name.
  - **level = ["string"];** Defines index into dimensions.
  - For APCP_06(lat,lon) from PCP-Combine output
    - *name = "APCP_06"; level = ["(*,*)"];*
  - For TT(Time, num_metgrid_levels, south_north, west_east) from p_interp
    - *name = "TT"; level = ["(0,0,*,*)", "(0,1,*,*)", "(0,2,*,*)"];*

Grid-Stat: Config File Defaults

- MET Statistics tools parse up to 4 configuration files:
  1. **MET_BASE/config/ConfigConstants** defines configuration file constants (e.g. NONE, STAT, BOTH) and should not be modified.
  2. **MET_BASE/config/ConfigMapData** defines default map data for all plots (map data files, line colors, widths, and types for Plot-Point-Obs, Plot-Data-Plane, Wavlet-Stat, and MODE).
  3. **MET_BASE/config/GridStatConfig_default** defines default settings for the specific tool.
  4. **User-specific configuration** file passed on the command line override default settings.

NOTE: When running a shared installation of MET, override default settings in the **user-specific configuration** file rather than modifying the system-wide defaults.

Grid-Stat: Run

- met-8.0/bin/grid_stat \
  sample_fcst.grb sample_obs.nc \
  GridStatConfig_APCP24 -outdir out -v 2

```
DEBUG 1: Default Config File: met-8.0/share/met/data/config/GridStatConfig_default
DEBUG 1: User Config File: GridStatConfig_APCP24
DEBUG 1: Forecast File: sample_fcst.grb
DEBUG 1: Observation File: sample_obs.nc
DEBUG 2: ----------------------------------------------------------------------
DEBUG 2: Processing APCP/A24 versus APCP_A24, for interpolation method UW_MEAN(1), over region FULL, using 6412 pairs
DEBUG 2: Computing Categorical Statistics.
DEBUG 2: Computing Multi-Category Statistics.
DEBUG 2: Computing Continuous Statistics.
DEBUG 2: Processing APCP/A24 versus APCPA24, for interpolation method UW_MEAN(1), over region EAST, using 2582 pairs.
DEBUG 2: Processing APCP/A24 versus APCPA24, for interpolation method NBRHD(9), raw thresholds of >0.000 and >0.000,
over region EAST, using 5829 pairs.
DEBUG 2: Computing Neighborhood Categorical Statistics.
DEBUG 2: Computing Neighborhood Continuous Statistics.
… MORE NEIGHBORHOOD VERIFICATION TASKS LISTED …
DEBUG 2: ----------------------------------------------------------------------
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V.stat
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_fho.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_ctc.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_cts.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_mctc.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_mcts.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_cnt.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_sl1l2.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_vl1l2.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_nbrctc.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_nbrcts.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_nbrcnt.txt
DEBUG 1: Output file: out/grid_stat_240000L_20050808_000000V_pairs.nc
```

Grid-Stat: ASCII Output Types

- Statistics line types: 20 possible
  - Same as Point-Stat
    - FHO, CTC, CTS, MCTC, MCTS, CNT
    - SL1L2, SAL1L2, VL1L2 , VAL1L2, VCNT
    - PCT, PSTD, PJC, PRC, ECLV
  - Gradient statistics (ECLV)
  - Neighborhood – apply threshold, define neighborhood
    - Neighborhood continuous statistics (NBRCNT)
    - Neighborhood contingency table counts (NBRCTC)
    - Neighborhood contingency table statistics (NBRCTS)
- 22 header columns common to all line types
- Remaining columns specific to each line type

Grid-Stat: Sample Output

1. STAT file output for sample run:
   - 2 lines each for CNT, MCTC, MCTS, and SL1L2
     - = 2 verification regions (FULL and EAST)
   - 4 lines each for FHO, CTC, and CTS
     - = 2 regions * 2 thresholds
   - 8 lines each for NBRCNT, NBRCTC, NBRCTS
     - = 2 regions * 2 thresholds * 2 neighborhood sizes
2. Additional TXT files for each line type
3. NetCDF file containing matched pairs

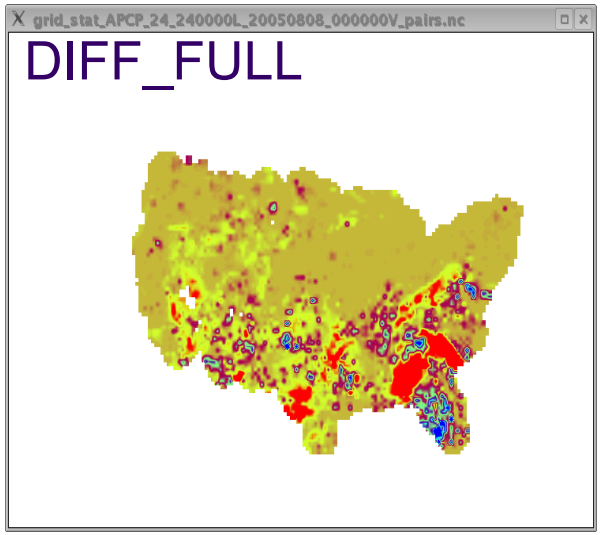
Grid-Stat: CTC Output Line

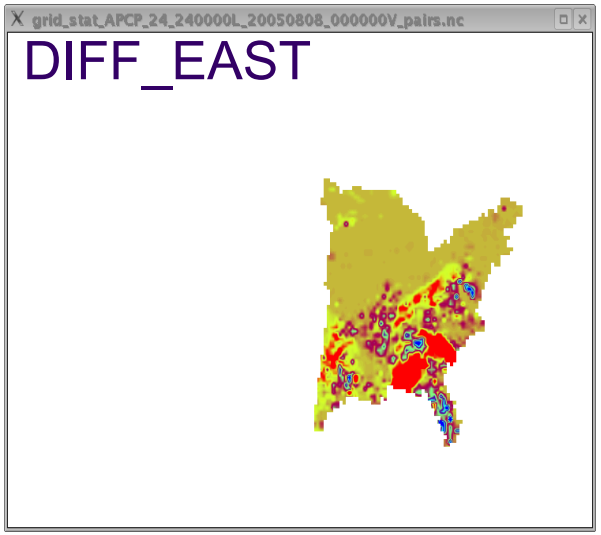
VERSION	V6.0
MODEL	WRF
DESC	NA
FCST_LEAD	240000
FCST_VALID_BEG	20050808_000000
FCST_VALID_END	20050808_000000
OBS_LEAD	000000
OBS_VALID_BEG	20050808_000000
OBS_VALID_END	20050808_000000
FCST_VAR	APCP_24
FCST_LEV	A24
OBS_VAR	APCP_24
OBS_LEV	A24
OBTYPE	MC_PCP

VX_MASK	EAST
INTERP_MTHD	UW_MEAN
INTERP_PNTS	1
FCST_THRESH	>20.000
OBS_THRESH	>20.000
COV_THRESH	NA
ALPHA	NA
LINE_TYPE	CTC
TOTAL	2582
FY_OY (hits)	5
FY_ON (f.a.)	104
FN_OY (miss)	70
FN_ON (c.n.)	2403

Grid-Stat: NetCDF Matched Pairs

- Forecast, observation, and difference fields for each combination of…
  - Variable, level, masking region, and interpolation method (smoothing)
- Sample output contains 6 fields:
  - FCST, OBS, and DIFF for FULL and EAST






```
//
// NetCDF matched
// pairs output file
//
nc_pairs_flag = {
   latlon      = TRUE;
   raw         = TRUE;
   diff        = TRUE;
   climo       = TRUE;
   weight      = FALSE;
   nbrhd       = FALSE;
   apply_mask  = TRUE;
}
```

Comparing Different Fields

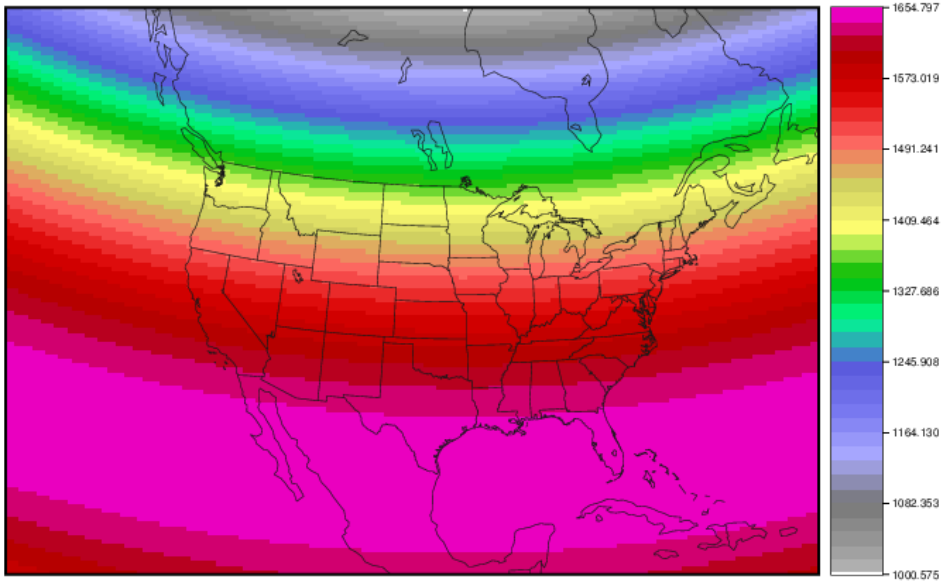
- Grid-Stat and Point-Stat may be used to compare two different variables.
  - User must interpret results.
  - Example: Convective Precip vs. Total Precip
  - Configuration file settings:
    - Selecting variable/levels

```
fcst = {
 field = [
  {
   name        = "ACPCP";
   level       = [ "A24" ];
   cat_thresh = [ >0.0 ];
  }
 ];
};
```

```
obs = {
 field = [
  {
   name        = "APCP";
   level       = [ "A24" ];
   cat_thresh = [ >0.0 ];
  }
 ];
};
```

Grid Point Weighting

```
// The "grid_weight_flag" specifies how grid weighting should be applied…
//    - "NONE" to disable grid weighting using a constant weight (default).
//    - "COS_LAT" to define the weight as the cosine of the grid point latitude.
//      This an approximation for grid box area used by NCEP and WMO.
//    - "AREA" to define the weight as the true area of the grid box (km^2).
```



true_area_weight.nc

Climatologies

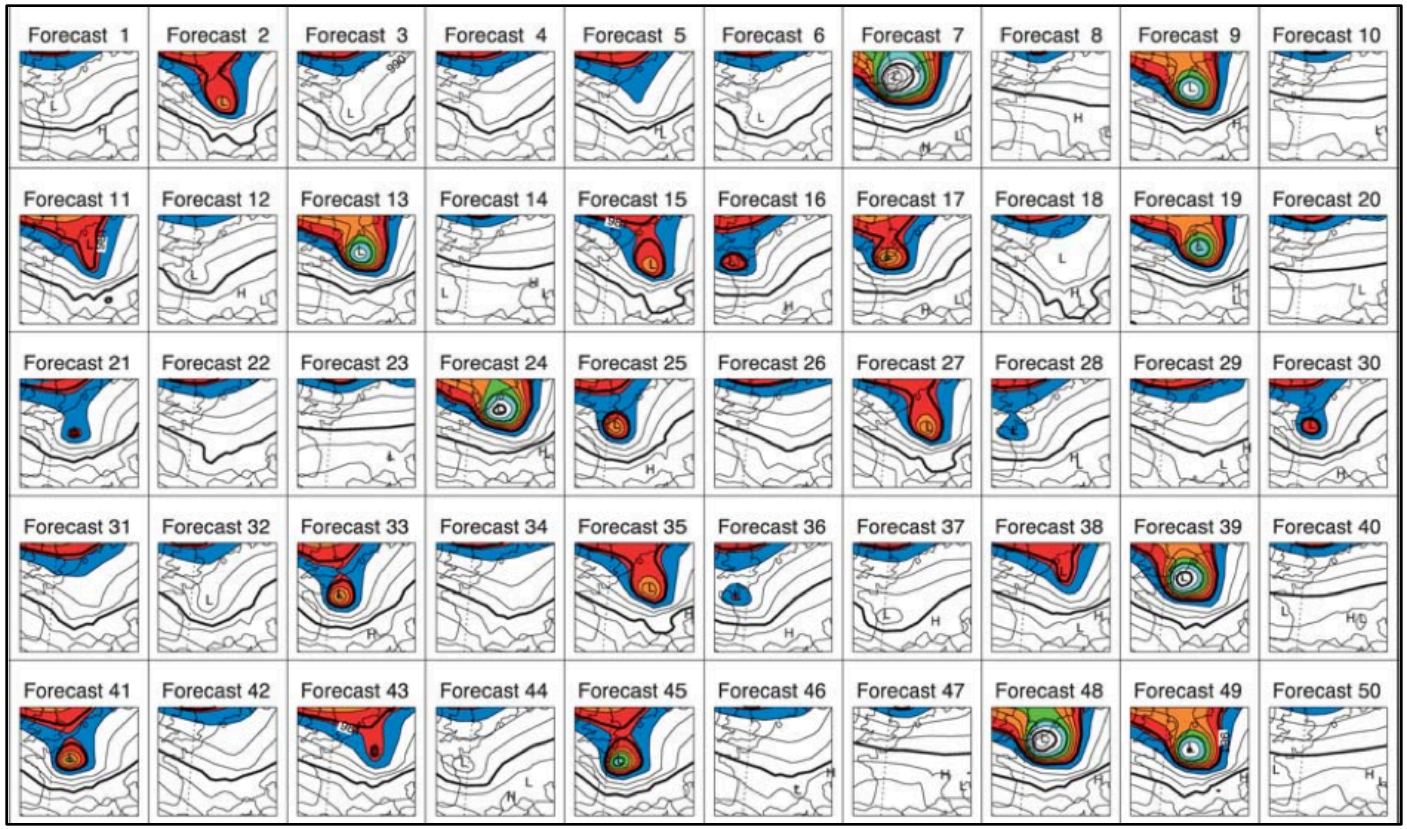
- Required for anomaly correlation (ANOM_CORR)
- Monthly 2.5 degree
  - *match_day = FALSE*
- Experimental daily 1.0 degree
  - *match_day = TRUE*
- Any other reference forecast.

```
//
// Climatology mean data
//
climo_mean   = {

   file_name = [
      // List of file names
   ];
   field     = [
      // Same length as fcst.field
   ];
   regrid    = {
      method     = NEAREST;
      width      = 1;
      vld_thresh = 0.5;
   }
   time_interp_method = DW_MEAN;
   match_month        = TRUE;
   match_day          = FALSE;
   time_step          = 21600;
}
```

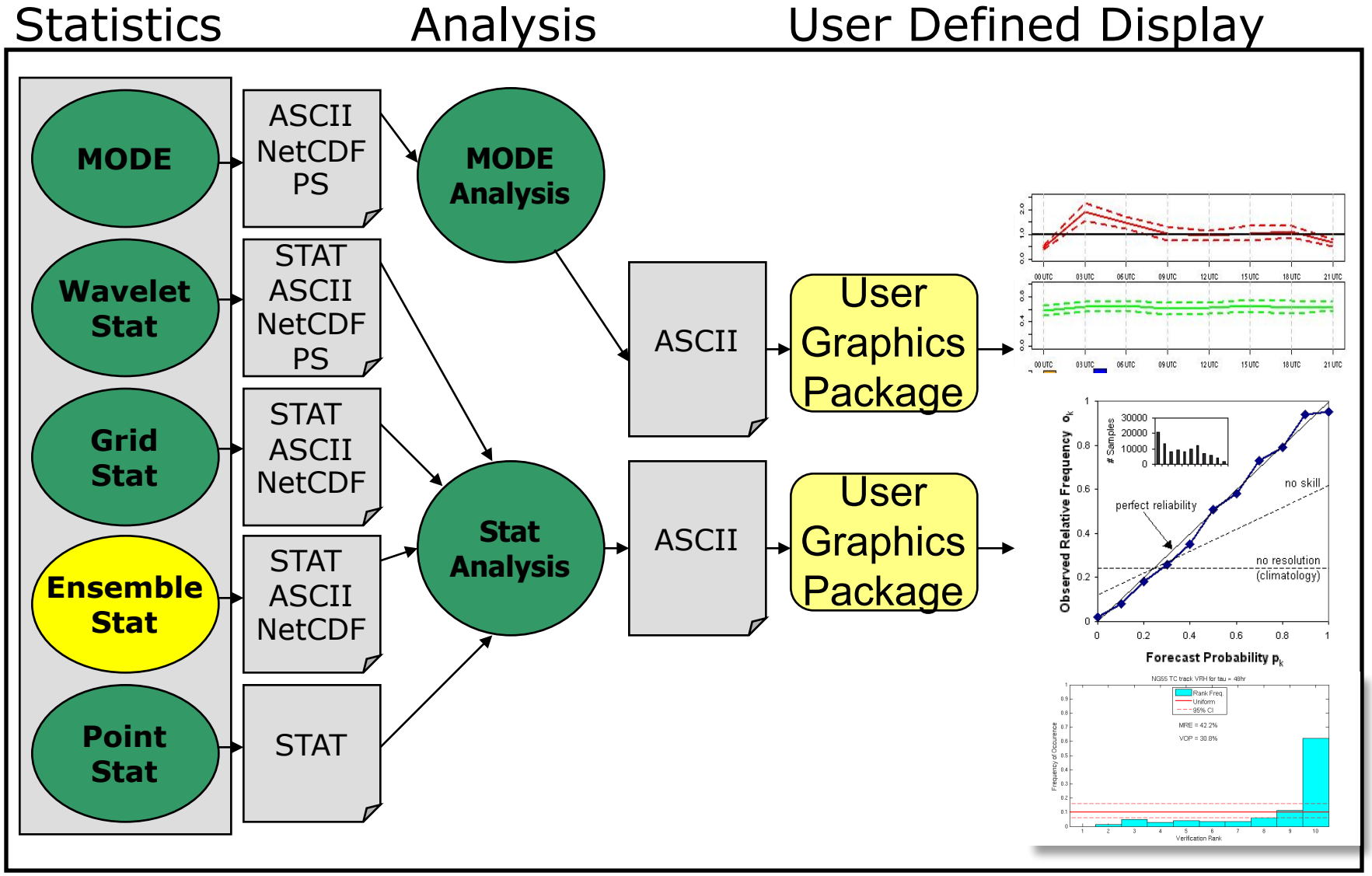
Verifying Ensembles & Probability Fcsts with MET

- Point-Stat and Grid-Stat Tool (probability)
  - Brier Score + Decomposition
  - Reliability Diagrams
  - Receiver Operating Characteristic Diagram + Area Under the Curve
  - Joint/Conditional factorization table
- **Ensemble-Stat Tool**
  - **Ensemble Mean Fields**
  - **Probability Fields**
  - **Rank Histograms**
  - **Spread-Skill Calculation**

Ensemble_Stat



Ensemble-Stat Tool



Ensemble-Stat Capabilities

Reads:

- Gridded ensemble member files
- ***Gridded AND point*** observations files

Calculates:

- Ensemble Mean, Standard Deviations, Mean ± 1 SD fields
- Ensemble Min, Max, and Range fields
- Ensemble Valid Data Count field
- Ensemble Relative Frequency by threshold fields (i.e. probability)
- Rank and PIT Histograms (if Obs Field Provided)
- Ensemble Spread-Skill (if Obs Field Provided)

Writes:

- Stat file with Rank Histogram, PIT Histogram, Spread-Skill partial sums, and Point Observation Ranks
- Gridded field of Observation Ranks to a NetCDF file

Ensemble Stat Tool: Usage

Usage: ensemble_stat

n_ens ens_file_1 \ ... ens_file_n | ens_file_list	Number of Ensemble members followed by list of ensemble member names OR ens_file_list (the name of an ASCII file with names of members)
config_file	Config file name
[-grid_obs file] **[-point_obs file]**	Name of gridded or point observed file – Required if Rank Histograms desired (optional)
[-ssvar_mean file]	Specify an ensemble mean model data file for use in calculating ensemble spread-skill (optional)
[-obs_valid_beg time] **[-obs_valid_end time]**	YYYYMMDD[_HH[MMSS]] format to set the beginning and end of the matching observation time window (optional)
[-outdir path]	Set output directory (optional)
[-log file]	Outputs log messages to the specified file (optional)
[-v level]	Set level of verbosity (optional)

Ensemble-Stat: Configuration

- Many configurable parameters
  - ens = fields to summarize
  - ens_thresh - All members must be available
  - vld_thresh – all data in grid must be valid
  - 24hr Accumulated Precip (APCP)
  - Composite Reflectivity (REFC)
  - N-S component of Wind (UGRD)
  - Thresholds used for Ensemble Relative Freq (i.e. probability)
  - GRIB1_ptv = 129; Use GRIB Table 129 instead of Table 2

```
//
// Ensemble product fields to be processed
// (i.e. mean, min, max, stdev fields)
//
ens = {
   ens_thresh = 1.0;
   vld_thresh = 1.0;

   field = [
      {
         name       = "APCP";
         level      = [ "A24" ];
         cat_thresh = [ >0.0, >=10.0 ];
      },
      {
         name       = "REFC";
         level      = [ "L0" ];
         cat_thresh = [ >=35.0 ];
         GRIB1_ptv  = 129;
      },
      {
         name       = "UGRD";
         level      = [ "Z10" ];
         cat_thresh = [ >=5.0 ];
      },
   ];
}
```

Ensemble-Stat: Configuration

- Many configurable parameters – only set a few:
  - Fcst specifies fields to be verified
  - ADPSFC message type for point obs
  - 24hr precip for gridded obs field
  - Bin size for spread-skill calculation is 0.1 mm
  - Bin size for probability integral transform statistics is 0.05 mm

```
// Forecast and observation fields to be
// verified (i.e. RHIST, PHIST, SSVAR)
//
fcst = {
   field = [
      {
         name                = "APCP";
         level               = [ "A24" ];
      }
   ];
}
obs = fcst;
```

```
// Point observation filtering options
// May be set separately in each "obs.field" entry
//
message_type   = [ "ADPSFC" ];
sid_exc        = [];
obs_thresh     = [ NA ];
obs_quality    = [];
duplicate_flag = NONE;
obs_summary    = NONE;
obs_perc_value = 50;
skip_const     = FALSE;

//
// Ensemble bin sizes
// May be set separately in each "obs.field" entry
//
ens_ssvar_bin_size = 0.1;
ens_phist_bin_size = 0.05;
```

Ensemble-Stat Tool: Run

```
ensemble_stat \
  6 sample_fcst/2009123112/*gep*/d01_2009123112_02400.grib \
  config/EnsembleStatConfig \
  -grid_obs sample_obs/ST4/ST4.2010010112.24h \
  -point_obs out/ascii2nc/precip24_2010010112.nc \
  -outdir out/ensemble_stat -v 2
```

NOTE:
You can pass in a path with wildcards to pull out the files you would like to process or you can pass in an ASCII filename that contains a list of ensemble members

Gridded and Obs fields are included for use in calculating Rank Histogram, PIT Histogram, and Spread-Skill

Ensemble Stat Tool: Run

```
*** Running Ensemble-Stat on APCP using GRIB forecasts, point observations, and gridded observations ***
DEBUG 1: Default Config File: /d3/projects/MET/MET_releases/met-8.0/data/config/EnsembleStatConfig_default
DEBUG 1: User Config File: config/EnsembleStatConfig
GSL_RNG_TYPE=mt19937
GSL_RNG_SEED=1
DEBUG 1: Ensemble Files[6]:
DEBUG 1:    ../data/sample_fcst/2009123112/arw-fer-gep1/d01_2009123112_02400.grib
DEBUG 1:    ../data/sample_fcst/2009123112/arw-fer-gep5/d01_2009123112_02400.grib
DEBUG 1:    ../data/sample_fcst/2009123112/arw-sch-gep2/d01_2009123112_02400.grib
DEBUG 1:    ../data/sample_fcst/2009123112/arw-sch-gep6/d01_2009123112_02400.grib
DEBUG 1:    ../data/sample_fcst/2009123112/arw-tom-gep3/d01_2009123112_02400.grib
DEBUG 1:    ../data/sample_fcst/2009123112/arw-tom-gep7/d01_2009123112_02400.grib
DEBUG 1: Gridded Observation Files[1]:
DEBUG 1:    ../data/sample_obs/ST4/ST4.2010010112.24h
DEBUG 1: Point Observation Files[1]:
DEBUG 1:    ../out/ascii2nc/precip24_2010010112.nc
DEBUG 2:
DEBUG 2: ----------------------------------------------------------------------
DEBUG 2:
DEBUG 2: Processing ensemble field: APCP/A24
DEBUG 2:
DEBUG 2: ----------------------------------------------------------------------
...
Processing gridded verification APCP_24/A24 versus APCP_24/A24, for observation type MC_PCP, over region FULL, for
interpolation method UW_MEAN(1), using 15480 pairs.
---------------------------------------------------------------------------
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V.stat
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_rhist.txt
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_phist.txt
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_orank.txt
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_ssvar.txt
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_ens.nc
DEBUG 1: Output file: out/ensemble_stat/ensemble_stat_20100101_120000V_orank.nc
```

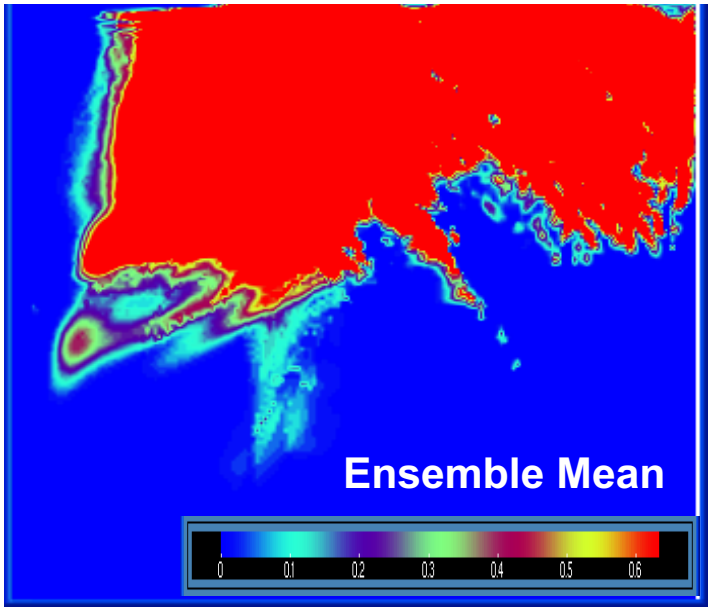
Ensemble-Stat: Output Files

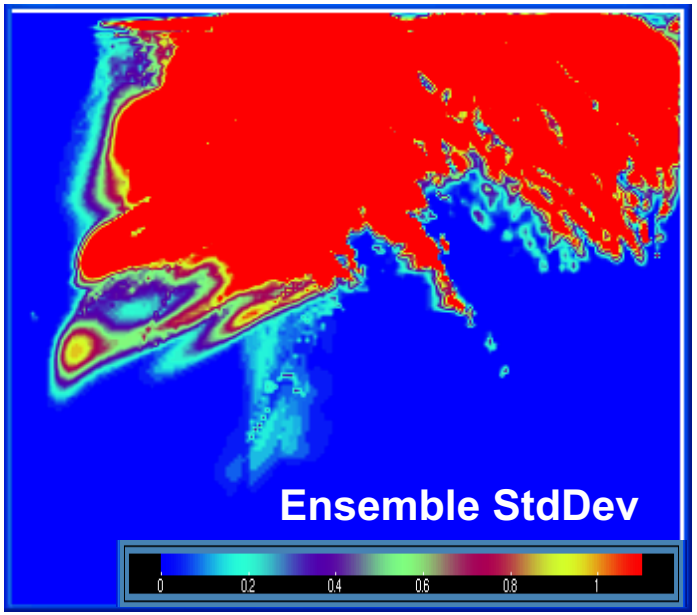
- **ASCII**
  - Up to 4 txt files and stat file
  - Ranked histogram (CPSS, IGN)
  - Probability integral transform histogram
  - Skill/spread variance
    - e.g. FBAR, OBAR, MSE, RMSE, PR_CORR
  - Relative position
- **netCDF**
  - Gridded ensemble mean, standard deviation, min, max, range, frequency
  - "orank" file (gridded obs rank)

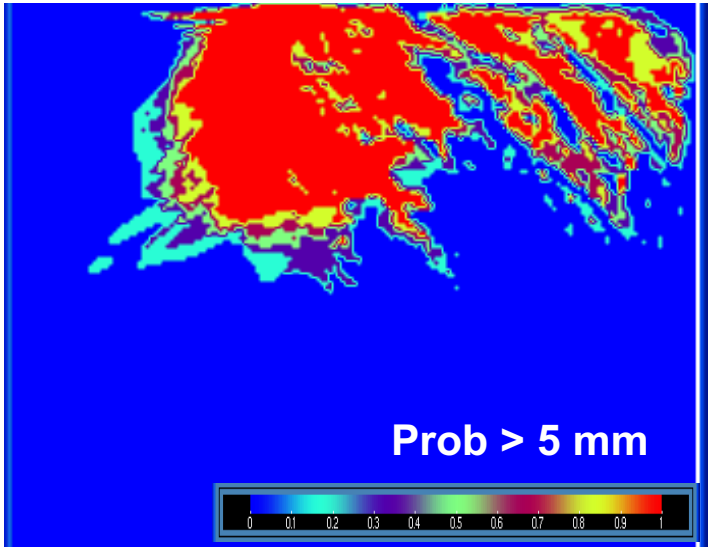
```
ensemble_flag = {
   mean       = TRUE;
   stdev      = TRUE;
   minus      = TRUE;
   plus       = TRUE;
   min        = TRUE;
   max        = TRUE;
   range      = TRUE;
   vld_count  = TRUE;
   frequency  = TRUE;
   rank       = TRUE;
   weight     = FALSE;
};
```

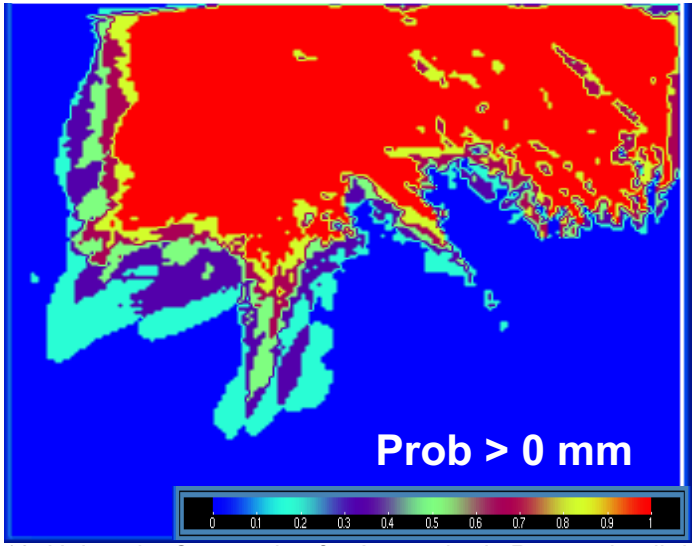
```
   output_flag = {
      ecnt  = BOTH;
      rhist = BOTH;
      phist = BOTH;
      orank = BOTH;
      ssvar = BOTH;
      relp  = BOTH;
   };
```

Ensemble Stat Tool: nc Output









Ensemble Stat Tool: txt Output

Output from *_rhist.txt

VERSION MODEL FCST_LEAD FCST_VALID_BEG FCST_VALID_END OBS_LEAD OBS_VALID_BEG OBS_VALID_END FCST_VAR FCST_LEV OBS_VAR OBS_LEV OBTYPE VX_MASK INTERP_MTHD INTERP_PNTS FCST_THRESH OBS_THRESH COV_THRESH ALPHA LINE_TYPE TOTAL CRPS IGN N_RANK RANK_1 RANK_2 RANK_3 RANK_4 RANK_5 RANK_6 RANK_7

V6.0 WRF 240000 20100101_120000 20100101_120000 000000 20100101_103000 20100101_133000 APCP_24 A24 APCP_24 A24 ADPSFC FULL UW_MEAN 1 NA NA NA NA RHIST 1125 8.21904 6.53721 7 261 160 138 141 149 111 165

CRPS IGN RANK HIST

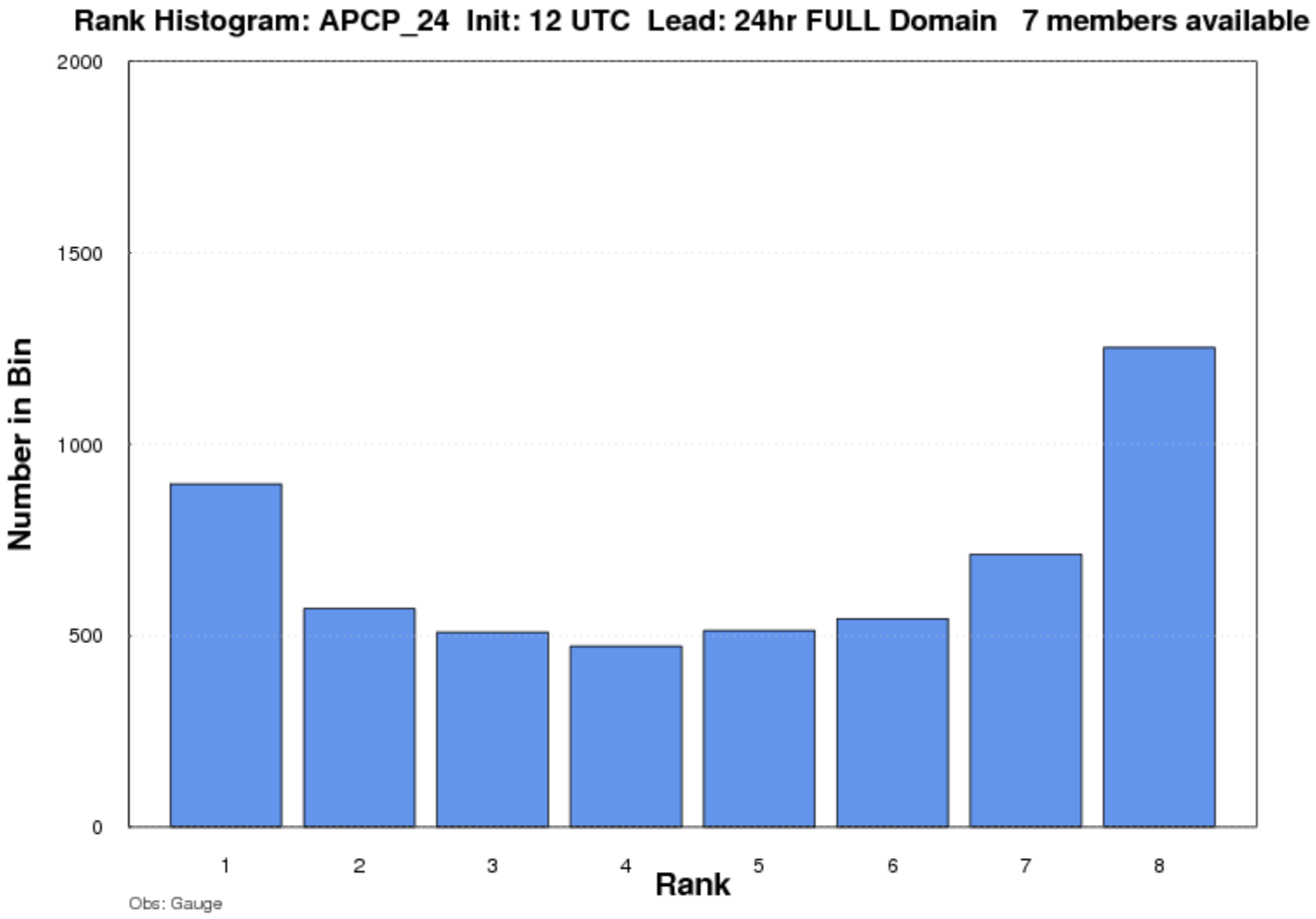
Output from *_phist.txt

VERSION MODEL FCST_LEAD FCST_VALID_BEG FCST_VALID_END OBS_LEAD OBS_VALID_BEG OBS_VALID_END FCST_VAR FCST_LEV OBS_VAR OBS_LEV OBTYPE VX_MASK INTERP_MTHD INTERP_PNTS FCST_THRESH OBS_THRESH COV_THRESH ALPHA LINE_TYPE TOTAL BIN_SIZE N_BIN BIN_1 BIN_2 BIN_3 BIN_4 BIN_5 BIN_6 BIN_7 BIN_8 BIN_9 BIN_10 BIN_11 BIN_12 BIN_13 BIN_14 BIN_15 BIN_16 BIN_17 BIN_18 BIN_19 BIN_20

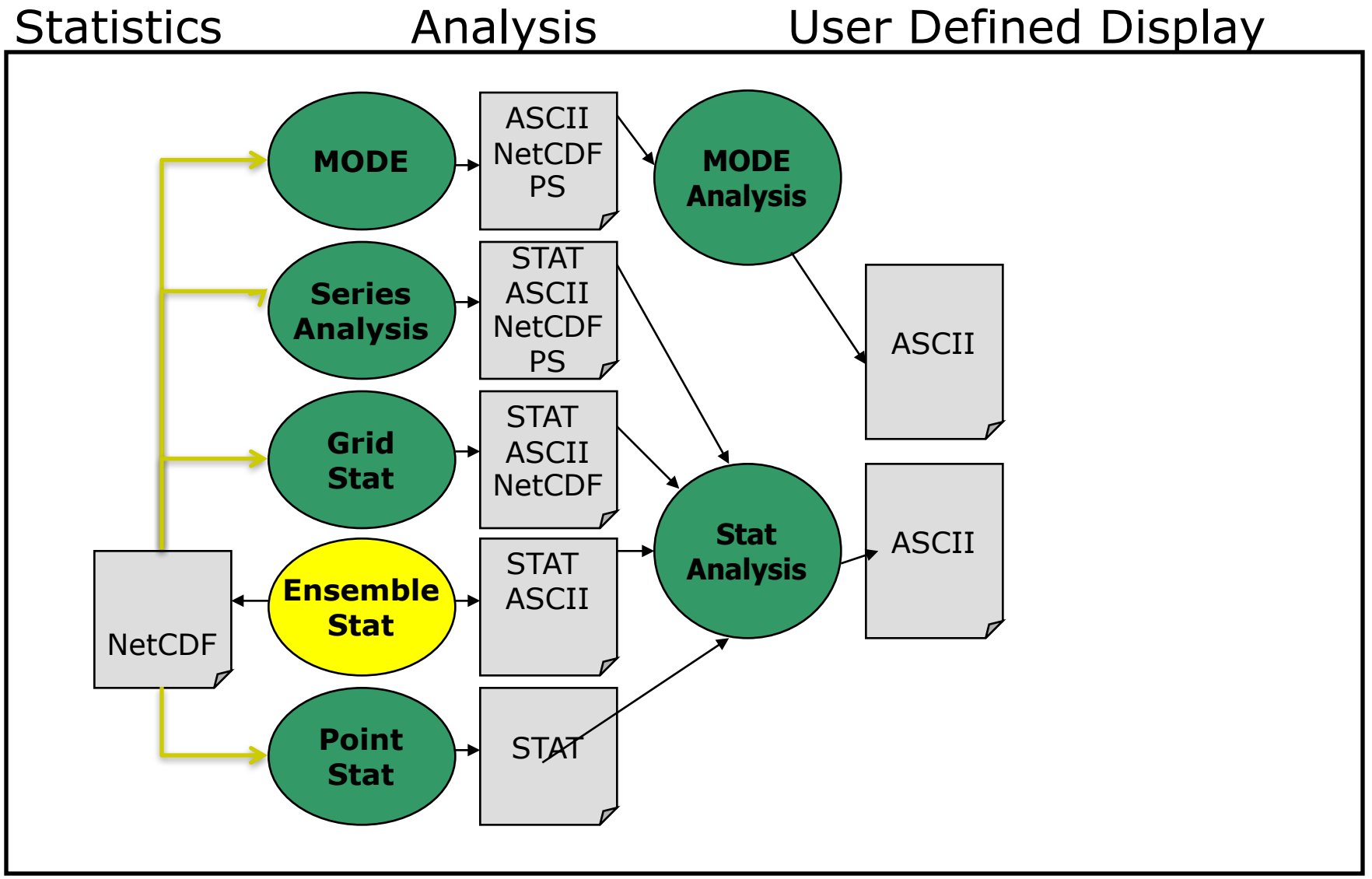
V6.0 WRF 240000 20100101_120000 20100101_120000 000000 20100101_103000 20100101_133000 APCP_24 A24 APCP_24 A24 ADPSFC FULL UW_MEAN 1 NA NA NA NA PHIST 55296 0.05 20 9261 3135 2565 2258 2237 2043 2084 2167 2059 2155 2205 2202 2198 2174 2097 2153 1992 2023 2315 5973

Probability integral transform histogram

Rank Histogram



Uses for Output from Ensemble Stat



Verifying Probabilities

- Probabilistic verification method tools:
  - Grid-Stat, Point-Stat, and Stat-Analysis
- Define Nx2 contingency table using:
  - Multiple forecast probability thresholds
  - One observation threshold

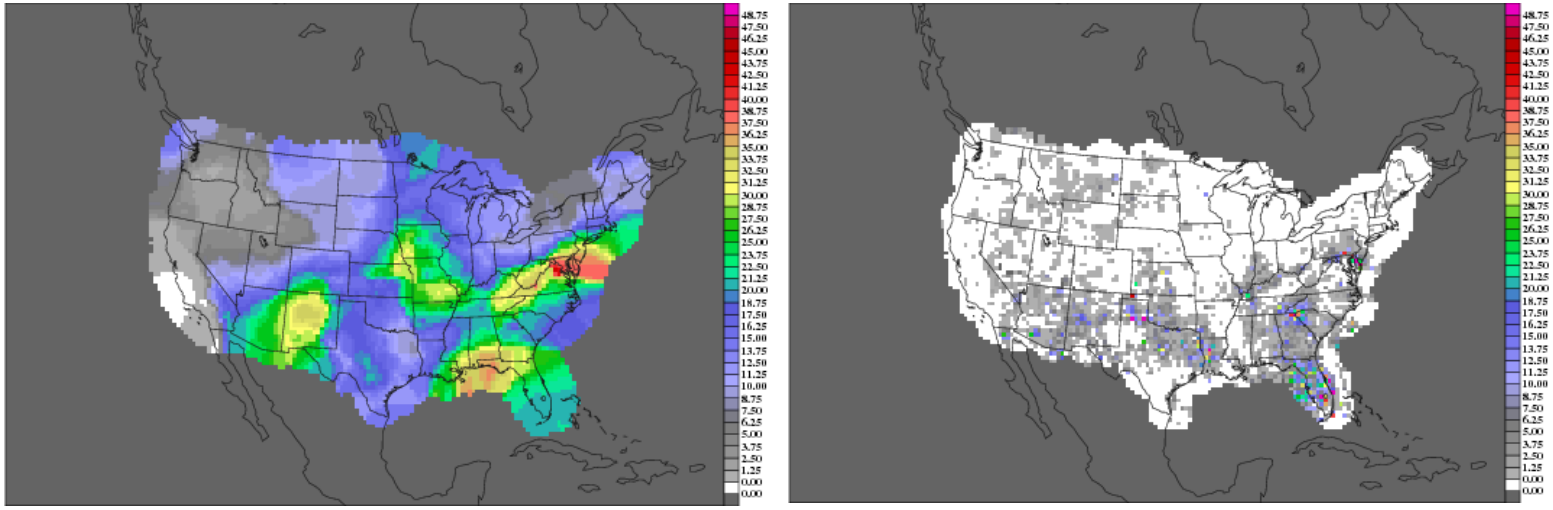
Forecast	Observation		Total
	o = 1 (e.g., "Yes")	o = 0 (e.g., "No")	
p_1 = midpoint of (0 and threshold1)	n_{11}	n_{10}	$n_{1\cdot} = n_{11} + n_{10}$
p_2 = midpoint of (threshold1 and threshold2)	n_{21}	n_{20}	$n_{2\cdot} = n_{21} + n_{20}$
⋮	⋮	⋮	⋮
p_j = midpoint of (threshold$i$ and 1)	n_{i1}	n_{i0}	$n_{j\cdot} = n_{j1} + n_{j0}$
Total	$n_{\cdot 1} = \Sigma n_{i1}$	$n_{\cdot 0} = \Sigma n_{i0}$	$T = \Sigma n_i$

Example:

- FCST: Probability of precip [0.00, 0.25, 0.50, 0.75, 1.00] ==0.25
- OBS: Accumulated precip > 0.00

Verifying Probabilities: Example

- Verify probability of precip with total precip:



- Configuration file settings:

```
fcst = {
 field = [
  {
   name          = "POP";
   level         = [ "Z0" ];
   //cat_thresh  = [ >=0.0, >=0.25, >=0.50, >=0.75, >=1.00 ];
   cat_thresh    = [ ==0.25 ];
   prob          = TRUE;
  }
 ];
};
```

```
obs = {
 field = [
  {
   name        = "APCP";
   level       = [ "A12" ];
   cat_thresh  = [ >0.0 ];
  }
 ];
};
```

Grid-Stat: Probability Config.

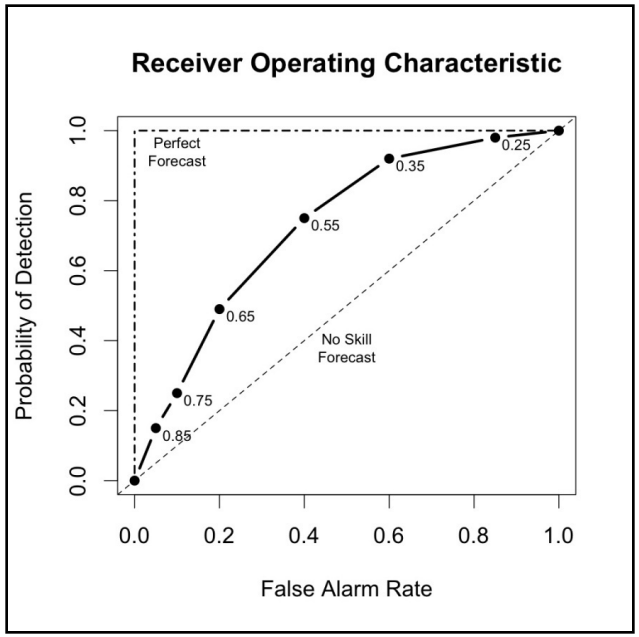
- Many configurable parameters – only set a few:
  - APCP_24… is name of ens mean in netcdf file
  - prob = True important
  - cat_thresh used for reliability and roc curves
  - Use 24hr Accumulation in grib file threshold at >10 mm
  - Generate probabilistic statistics

```
fcst = {
  wind_thresh = [ NA ];

  field = [
    {
     name       = "APCP_24_A24_ENS_FREQ_ge10.000";
     level      = [ "(*,*)" ];
     prob       = TRUE;
     cat_thresh = [ >=0.0, >=0.1, >=0.2, >=0.3,
  >=0.4, >=0.5, >=0.6, >=0.8, >=1.0 ];
     //cat_thresh = [ ==0.1 ];
     }
  ];

};
obs = {

  wind_thresh = [ NA ];

  field = [
    {
     name       = "APCP";
     level      = [ "A24" ];
     cat_thresh = [ >10.000 ];
    }
  ];

};
```

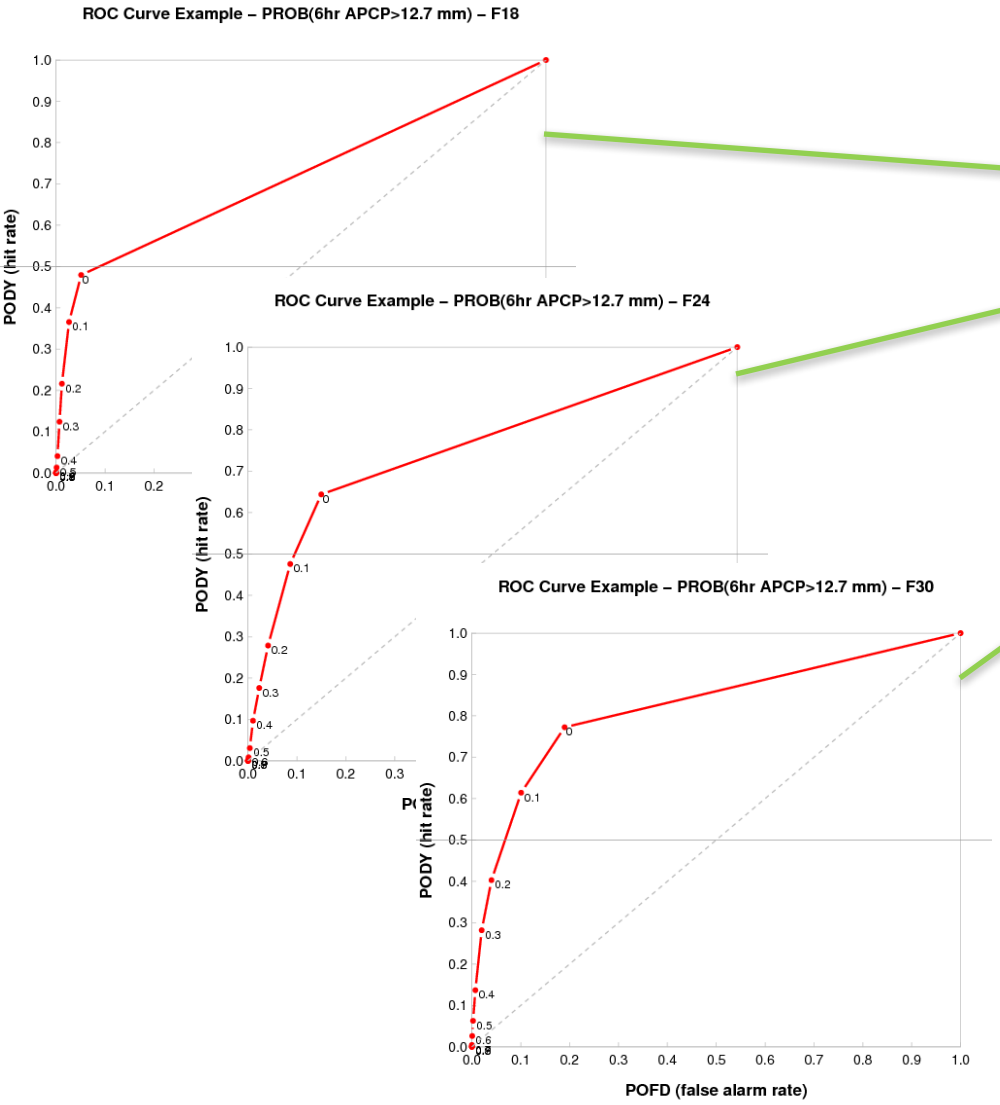
```
output_flag = {
   fho    = NONE;
   ctc    = NONE;
   cts    = NONE;
   mctc   = NONE;
   mcts   = NONE;
   cnt    = NONE;
   sl1l2  = NONE;
   vl1l2  = NONE;
   pct    = BOTH;
   pstd   = BOTH;
   pjc    = BOTH;
   prc    = BOTH;
   nbrctc = NONE;
   nbrcts = NONE;
   nbrcnt = NONE;
};
```

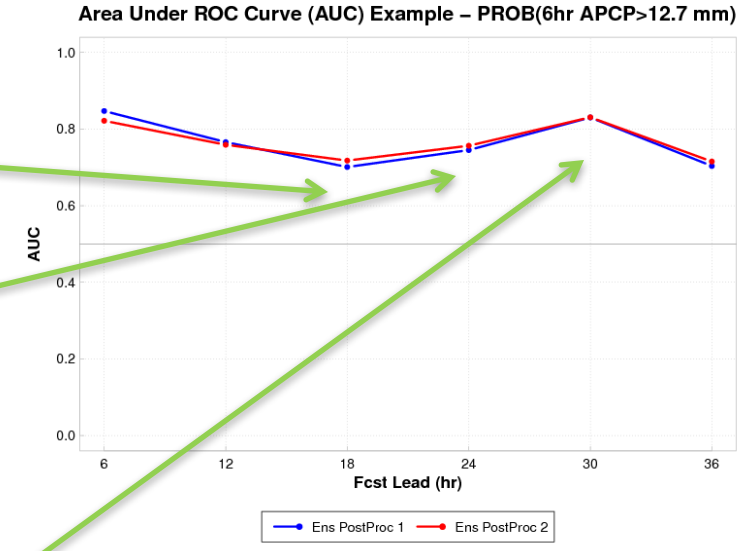
Grid Stat for Probability: Run

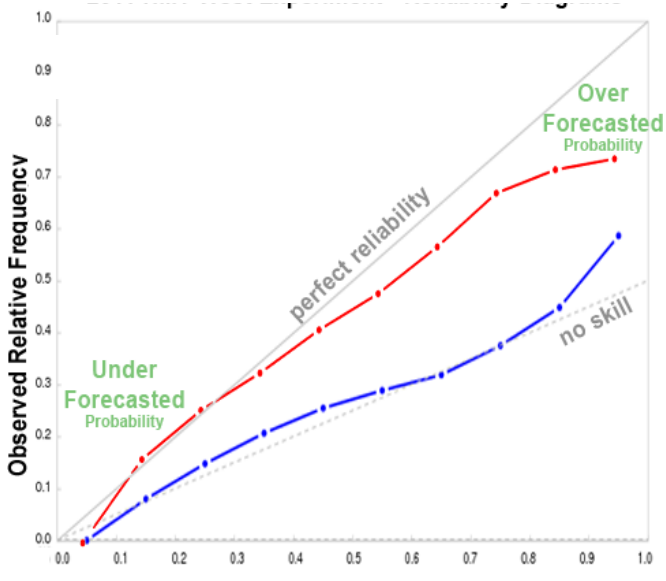
- Output written to .stat file and, if desired, to individual text files:
  - PCT – Probability Contingency Table Counts
  - PSTD – Probability Contingency Table Scores
    - Brier Score, Reliability, Resolution, Uncertainty, Area Under ROC
  - PJC – Joint/Continuous Statistics of Probabilistic Variables
    - Calibration, Refinement, Likelihood, Base Rate, Reliability points
  - PRC – ROC Curve Points for Probabilistic Variables



Grid Stat Probability: Examples







Stat-Analysis Tool

- Filtering
- Summarizing
- Aggregating

of Grid-Stat, Point-Stat,
Ensemble-Stat & Wavelet-Stat output

What can Stat Analysis do?

Questions to MET Help - Can I get…

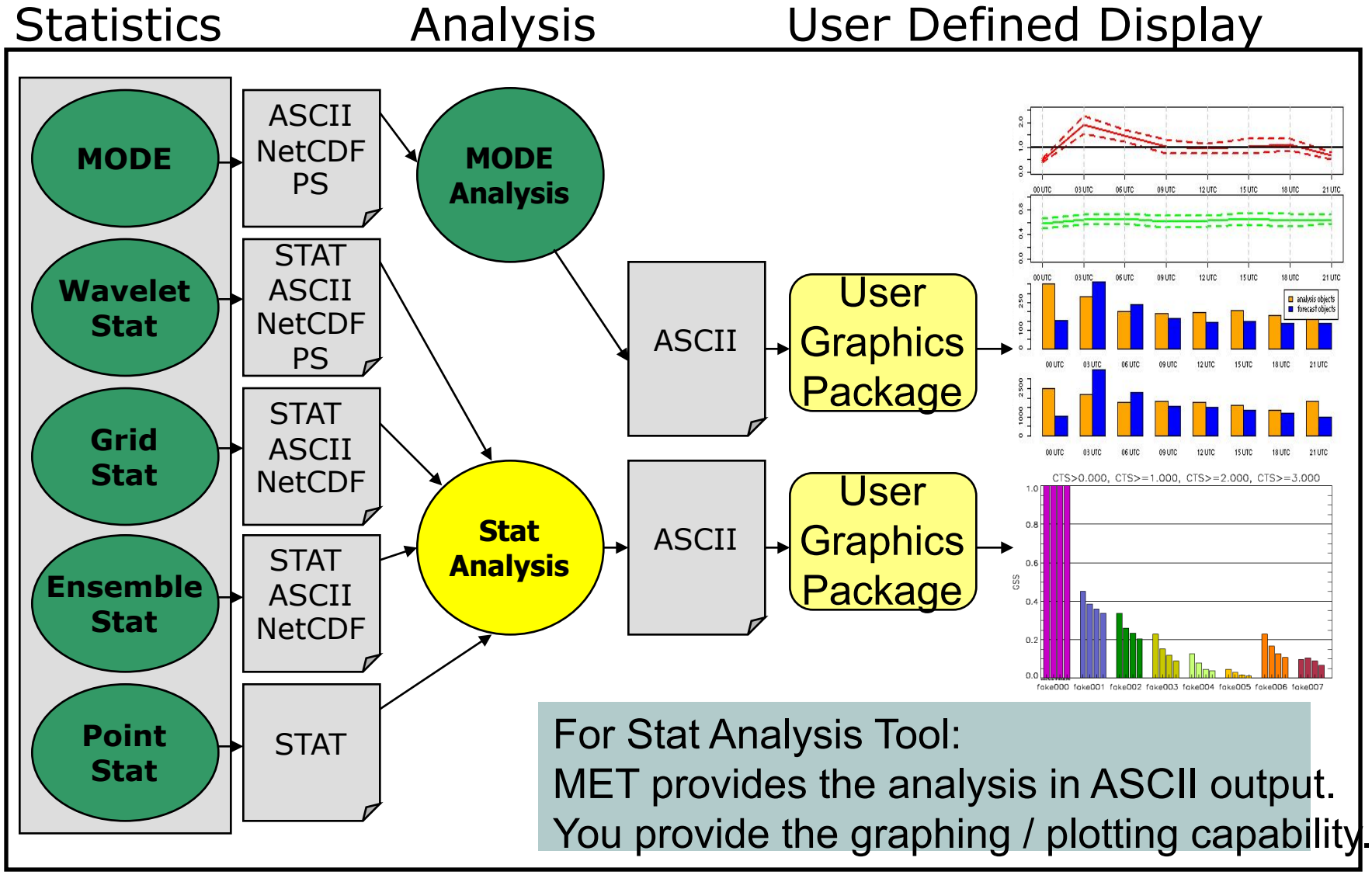
Q: Overall statistics for gridded observations compared to forecasts, hours 0 - 24?
A: Using Stat Analysis Tool on Grid-Stat output

Q: Contingency table statistics aggregated over multiple runs?
A: Using Stat Analysis Tool on any output

Q: Long-term statistics at individual sites (e.g., MAE or RMS error, daily forecasts for a month)?
A: Using Stat Analysis Tool on Point-Stat output

Q: Statistics aggregated for a large number (N) of individual stations in one simultaneous run?
A: It would be cumbersome. You would have to configure Stat Analysis Tool to run (N) number of jobs
A: OR use METViewer tool.

Stat Analysis Tool



For Stat Analysis Tool:
MET provides the analysis in ASCII output.
You provide the graphing / plotting capability.

Stat Analysis Jobs

- **Filtering (filter)**
  - filters out lines from one or more stat files
  - filters based on user-specified filtering options.
- **Summarizing (summary)**
  - Summary information from a single data column
  - Includes *mean, standard deviation, min, max, IQR, percentiles (0th, 25th, 50th, 75th, and 90th)*
- **Customized tool for AFWA (go_index)**
  - computes GO Index, performance statistic used primarily by the US Air Force
- **Ramp**
  - Computes amount of change from one time to next
  - Changes thresholded to produce contingency table

Stat Analysis Jobs

- **Aggregation**
  - **aggregate -** aggregates stat data across multiple time steps or masking regions.
    - Output line type is same as input line type (i.e. SSVAR = SSVAR)
  - **aggregate_stat –** aggregates across multiple times/regions then calculates statistics.
    - Output line is different from input line types.

Valid line type combinations include:

-line_type		-out_line_type
FHO, CTC	yields	CTS
MCTC	yields	MCTS
SL1L2, SAL1L2	yields	CNT
VL1L2, VAL1L2	yields	WDIR
PCT	yields	PSTD, PJC, PRC
NBRCTC	yields	NBRCTS
MPR	yields	FHO, CTC, CTS, MCTC, MCTS, CNT, SL1L2, SAL1L2, PCT, PSTD, PJC, PRC

Stat Analysis Tool: Usage

Usage: stat_analysis

-lookin path

[-out filename]

[-tmp_dir path]

[-v level]

-config config_file

or –job at command line options with associated arguments

[filter]

[summary]

[aggregate]

[aggregate_stat]

[go_index]

-lookin	Path to *.stat files – this can be a directory or a single file name (Use one or more times)
-out	Output name for ASCII file
-tmp_dir	Folder for temporary files
-v	Level of logging
-config	StatAnalysisConfig file
filter	See previous 2 slides
summary	See previous 2 slides
aggregate	See previous 2 slides
aggregate_stat	See previous 2 slides
go_index	See previous 2 slides

Stat-Analysis: Configuration

- Many configurable parameters – only set a few:
  - 10-m U-component of wind.
  - Aggregate stats over DTC165 and DTC166 regions
  - Accumulate only CTCs calculated using Distance-Weighted Mean interpolation
  - Dump lines included in accumulation
  - Dump aggregation to file

- OR -

can put it all in the “jobs”area…

```
fcst_var = ["UGRD"];
obs_var  = [];

fcst_lev = [];
obs_lev  = [];

obtype = [];

vx_mask = ["DTC165", "DTC166"];

interp_mthd = ["DW_MEAN"];
```

```
jobs = [
   "-job filter -line_type CTC -dump_row
 outdir/job_filter_ctc_ugrd.stat",
   "-job aggregate -line_type CTC -dump_row
 outdir/job_aggregate_ctc_ugrd.stat"
];
```

- OR -

```
jobs = [
"-job filter -line_type CTC -dump_row
 out/job_filter_ctc_ugrd.stat",
"-job aggregate -line_type CTC -fcst_var UGRD
 -vx_mask DTC165 -vx_mask DTC166
 -interp_mthd DW_MEAN -dump_row
 out/job_aggregate_ctc_ugrd.stat"
];
```

Stat Analysis Tool: Run –job aggregate

```
"-job aggregate -line_type CTC -fcst_var UGRD -vx_mask DTC165 -vx_mask DTC166
 -interp_mthd DW_MEAN -dump_row out/job_aggregate.stat"
```

Stat Analysis Filter Output in `job_aggregate.stat`

```
V4.1     WRF    360000
    20070331_120000 20070331_120000
    000000   20070331_103000
    20070331_133000 UGRD       Z10
    UGRD       Z10       ADPSFC DTC165
    DW_MEAN      9             >=5.000
    >=5.000     NA           NA       CTC
    934  32          43          32
    827
V4.1     WRF    360000
    20070331_120000 20070331_120000
    000000   20070331_103000
    20070331_133000 UGRD       Z10
    UGRD       Z10       ADPSFC DTC166
    DW_MEAN      9             >=5.000
    >=5.000     NA           NA       CTC
    2955 24          104         72
    2755
```

(NOTE: header modified to show only pertinent info)

		OBS		
FCST		Y	N	
	Y	32	43	75
	N	32	827	859
		64	870	934

		OBS		
FCST		Y	N	
	Y	24	104	128
	N	72	2755	2827
		96	2859	2955

Stat Analysis Tool: Run –job aggregate

```
"-job aggregate -line_type CTC -fcst_var UGRD -vx_mask DTC165 -vx_mask DTC166
 -interp_mthd DW_MEAN -dump_row out/job_aggregate.stat"
```

Stat Analysis Output in the file specified by –out flag ***(i.e. stat_analysis.out)***

```
JOB_LIST:      -job aggregate
   -fcst_var UGRD -vx_mask DTC165 -
   vx_mask DTC166 -interp_mthd
   DW_MEAN -line_type CTC -dump_row
   out/aggregate2.stat
COL_NAME: TOTAL FY_OY FY_ON FN_OY
   FN_ON
     CTC: 3889  56    147   104
   3582
```

		OBS		
FCST		Y	N	
	Y	56	147	251
	N	104	3582	317
		241	327	3889

Stat Analysis Tool: Run –job aggregate_stat

```
"-job aggregate_stat -line_type CTC -out_line_type CTS -fcst_var UGRD -
 vx_mask DTC165 -vx_mask DTC166 -interp_mthd DW_MEAN -dump_row
 out/job_aggregate_stat.stat"
```

Aggregate_stat Output *(stat_analysis.out continued)*

```
COL_NAME: TOTAL BASER BASER_NCL BASER_NCU
BASER_BCL BASER_BCU FMEAN FMEAN_NCL FMEAN_NCU FMEAN_BCL
FMEAN_BCU ACC ACC_NCL ACC_NCU ACC_BCL ACC_BCU FBIAS
FBIAS_BCL FBIAS_BCU PODY PODY_NCL PODY_NCU PODY_BCL
PODY_BCU PODN PODN_NCL PODN_NCU PODN_BCL PODN_BCU POFD
POFD_NCL POFD_NCU POFD_BCL POFD_BCU FAR FAR_NCL
FAR_NCU FAR_BCL FAR_BCU CSI CSI_NCL CSI_NCU CSI_BCL
CSI_BCU GSS GSS_BCL GSS_BCU HK HK_NCL HK_NCU HK_BCL
HK_BCU HSS     HSS_BCL HSS_BCU ODDS ODDS_NCL ODDS_NCU
ODDS_BCL ODDS_BCU  CTS: 3889  0.04114 0.03534
0.04785  NA  NA  0.05220 0.04564 0.05964   NA
NA  0.93546 0.92730 0.94276  NA  NA  1.26875 NA
NA  0.35000 0.33516  0.36513  NA  NA  0.96058
0.95400  0.96625  NA  NA  0.03942 0.03375
0.04600  NA  NA  0.72414 0.70987 0.73796  NA
NA  0.18241 0.17059 0.19486  NA  NA  0.15955
NA  NA  0.31058 0.23588 0.38528 NA  NA  0.27519
NA  NA  13.12088 9.11454  18.88823  NA  NA
```

		OBS		
		Y	N	
FCST	Y	56	147	203
	N	104	3582	3686
		160	3729	3889

Base Rate: 0.04
Freq Bias: 1.27
PODY: 0.35
FAR: 0.72
CSI: 0.18
GSS: 0.15

Stat Analysis Tool: Run –job summary

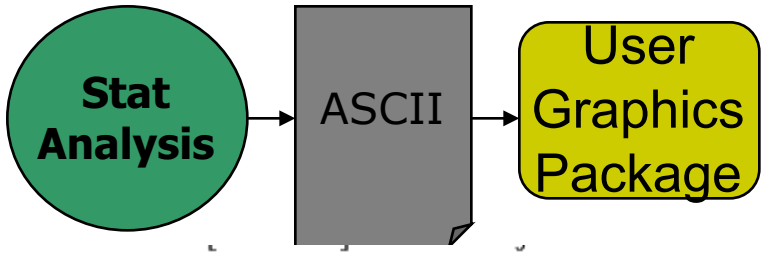
```
"-job summary -fcst_var UGRD -interp_mthd DW_MEAN -line_type CTS -column GSS
-dump_row out/job_summary.stat"
```

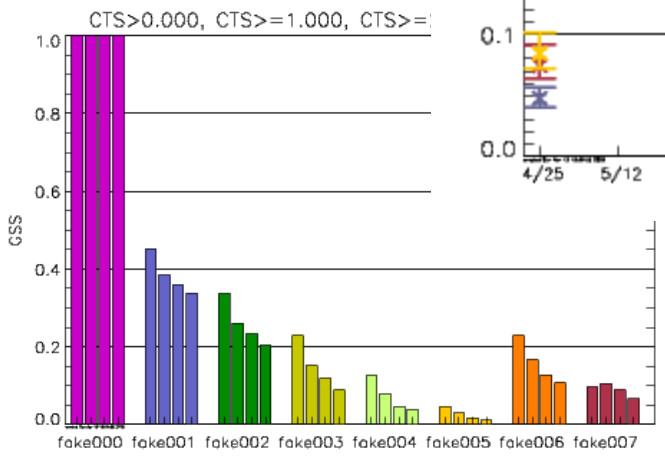
#	Description	
1	Column Name	Summary
2	Total	3
3-7	Mean* ***Includes normal and bootstrap upper and lower confidence limits***	*0.109*
8-10	Standard deviation** ***Includes bootstrap upper and lower confidence limits***	*0.150*
11	Minimum value	*-0.036*
12	10th percentile	-0.008
13	25th percentile	0.032
14	Median (50th percentile)	0.101
15	75th percentile	0.182
16	90th percentile	0.231
17	Maximum value	0.263

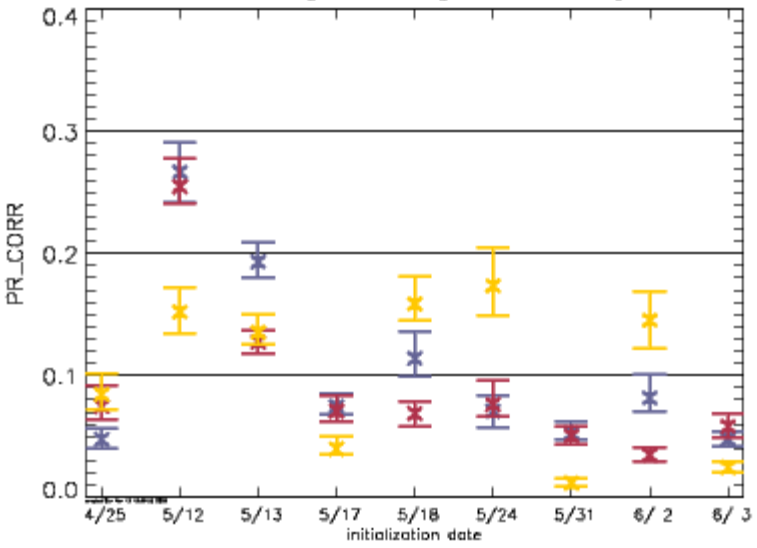
Summary Output *(stat_analysis.out cont.)*

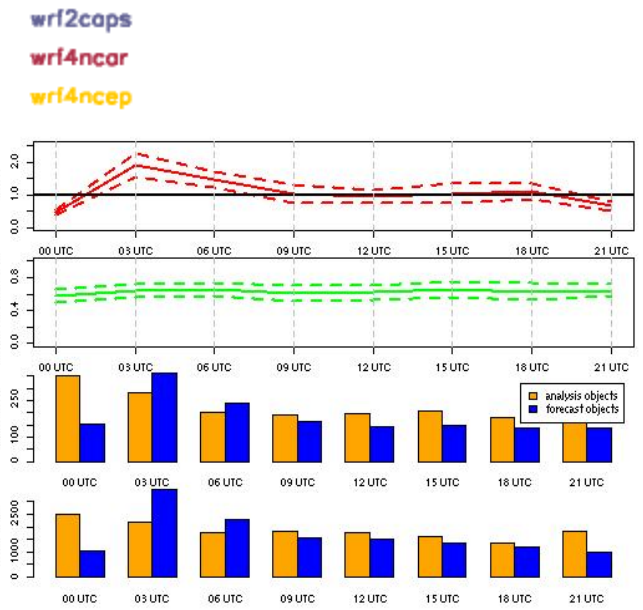
```
COL_NAME: TOTAL MEAN
MEAN_NCL MEAN_NCU MEAN_BCL
MEAN_BCU STDEV   STDEV_BCL
STDEV_BCU MIN      P10
P25     P50     P75     P90
MAX
SUMMARY: 3     0.10963
-0.26321 0.48247  -0.03613
0.26370  0.15009 0.00000
0.17311   -0.03613 -0.00864
0.03259 0.10131 0.18251
0.23122 0.26370
```

Use your favorite plotting software



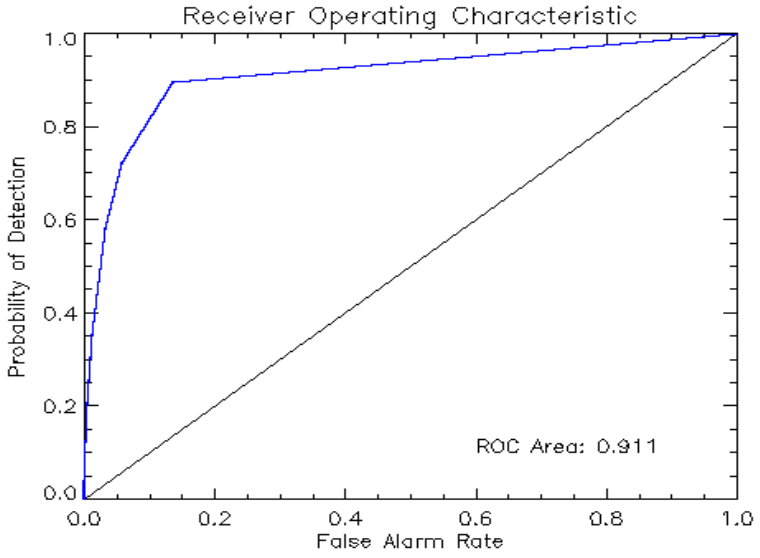


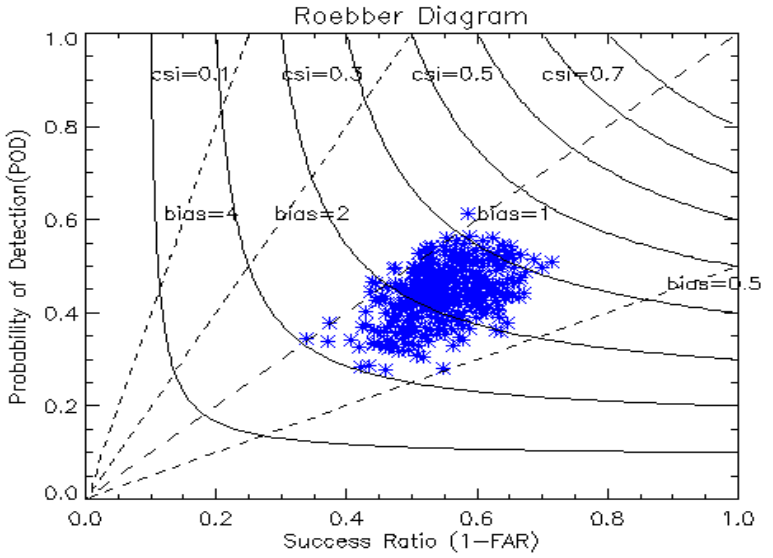


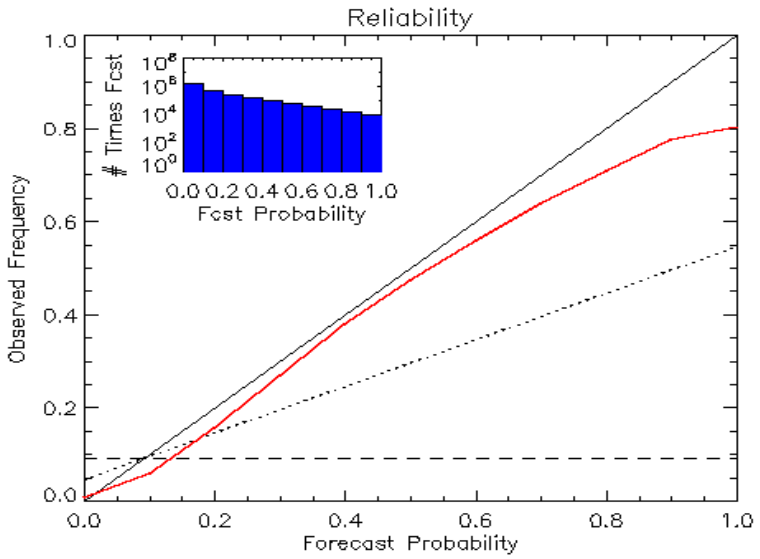


MET provides the analysis in ascii ouput.
You provide the graphing / plotting capability.

Stat_Analysis Example





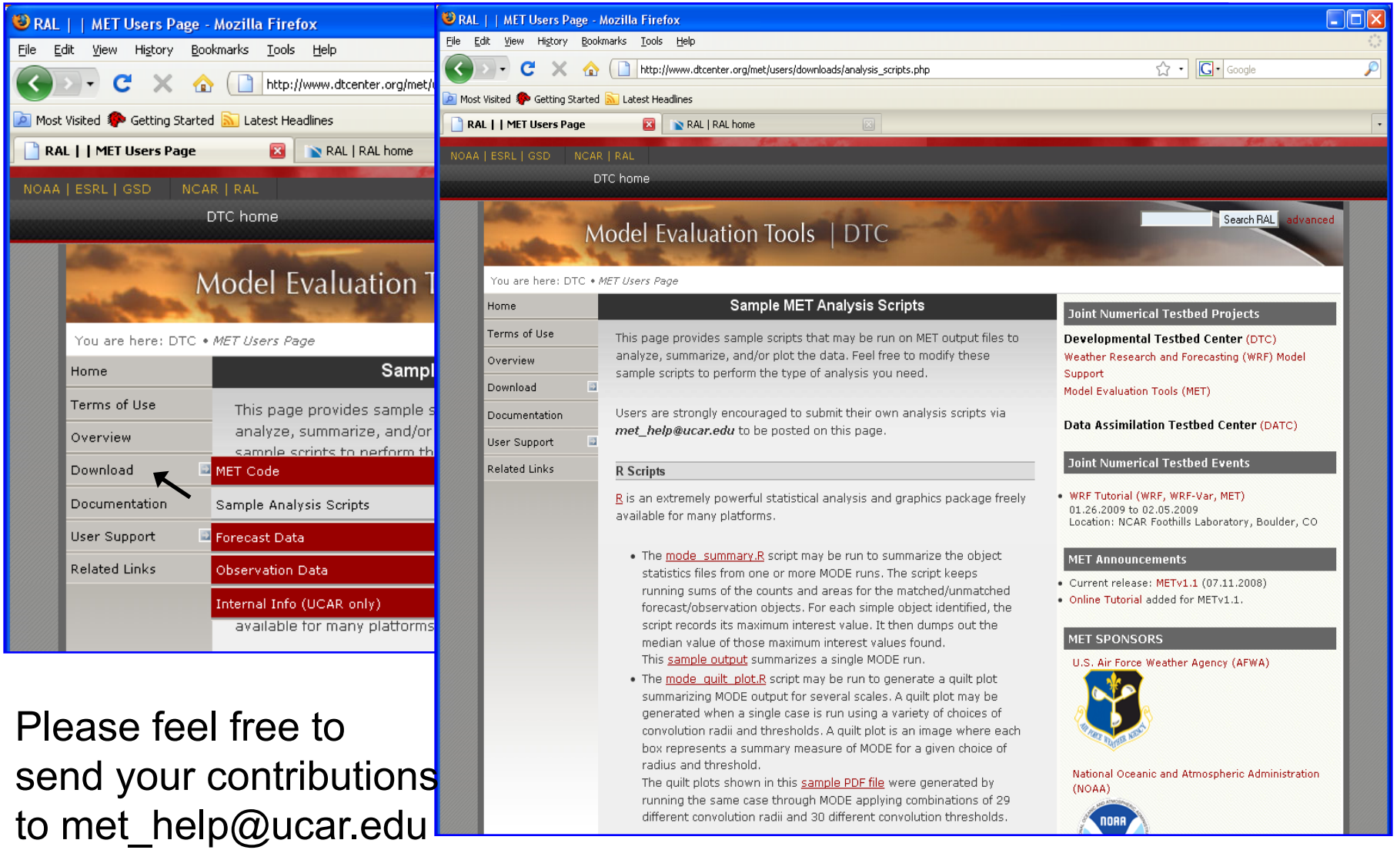


03/01/2013 - 08/30/2013

vx_mask = [“EPOCH”]

Jobs: aggregate PCT
aggregate_stat, PCT to PRC
aggregate_stat PCT to PSTD

User Contributed Plotting Scripts



Please feel free to send your contributions to met_help@ucar.edu