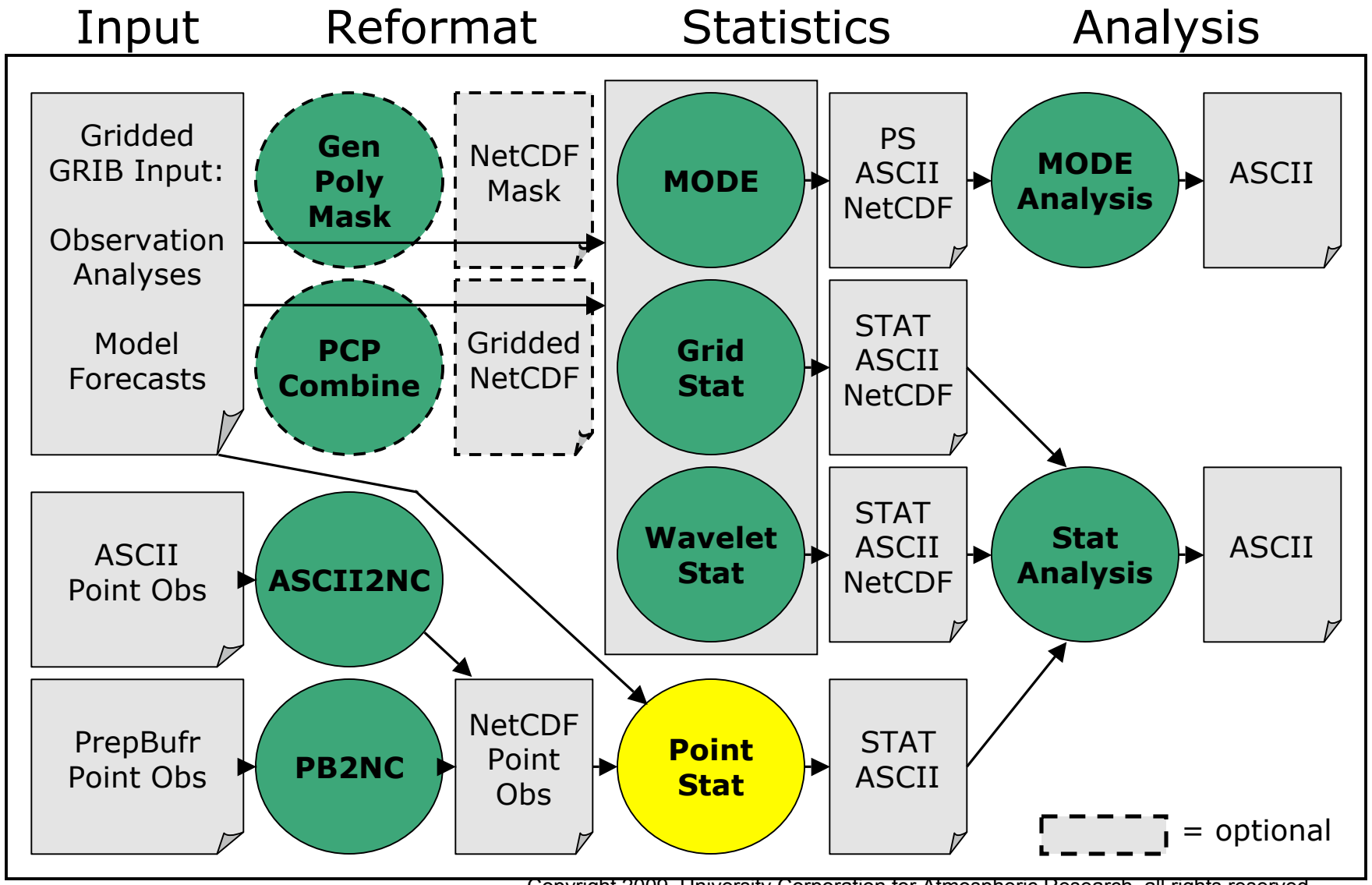
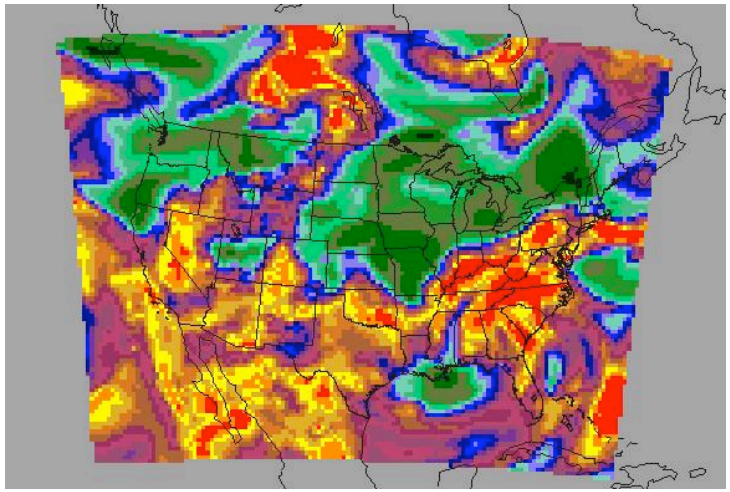
Using Point and Grid Stat

John Halley Gotway

Point-Stat Tool

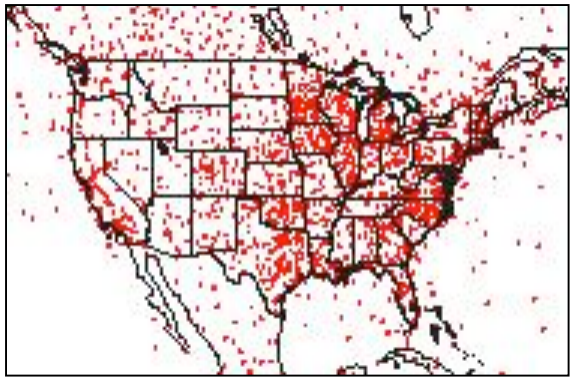


Point-Stat: Overview



- Compare **gridded forecasts** to **point observations**.
- 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.
  - Categorical counts and statistics for thresholded fields.
  - Parametric and non-parametric confidence intervals for statistics.
  - Compute partial sums for raw fields and/or the raw matched pair values.
  - Probabilistic methods.



Point-Stat: Input/Output

- Input Files
  - Gridded forecast file
    - GRIB output of WRF Post-Processor (or other)
    - NetCDF output of PCP-Combine
  - Point observation file
    - NetCDF output of PB2NC
    - NetCDF output of ASCII2NC
  - 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")

Point-Stat: Usage

Usage: point_stat

fcst_file

obs_file

config_file

[-climo climo_file]

[-ncfile netcdf_file]

[-valid_beg time]

[-valid_end time]

[-outdir path]

[-v level]

fcst_file	Forecast file in GRIB or NetCDF
obs_file	Point observation file in NetCDF (PB2NC or ASCII2NC)
config_file	ASCII configuration file
-climo	Climatological file for computing anomaly partial sums
-ncfile	Additional point observation files in NetCDF
-valid_beg	Beginning of valid time window for matching
-valid_end	End of valid time window for matching
-outdir	Output directory to be used
-v	Level of logging

Point-Stat: Configuration

- 25 configurable parameters – only set a few:
  - Temperature at the surface (2-meter).
    - **fcst_field[] = ["TMP/Z2“];**
  - Temperature below freezing.
    - **fcst_thresh[] = [“lt273“];**
  - Match to observations at the surface.
    - **message_type[] = ["ADPSFC"];**
  - Look at all the points in my domain.
    - **mask_grid[] = ["FULL"];**
  - Match observation to the nearest forecast value.
    - **interp_wdth[] = [1];**
  - Generate all possible statistic types.
    - **output_flag[] = [2, 2, 2, 2, 2, 2, 2, 2, 2];**

Point-Stat: Run

- METv2.0/bin/point_stat \
  sample_fcst.grb sample_pb.nc \
  PointStatConfig_TMPZ2 -outdir out -v 2

```
Forecast File: sample_fcst.grb
Climatology File: none
Configuration File: PointStatConfig_TMPZ2
Observation File: sample_pb.nc
----------------------------------------
Reading records for TMP/Z2.
For TMP/Z2 found 1 forecast levels and 0 climatology levels.
----------------------------------------
Searching 179772 observations from 44076 PrepBufr messages.
----------------------------------------
Processing TMP/Z2 versus TMP/Z2, for observation type ADPSFC, over region FULL, for interpolation method
UW_MEAN(1), using 11370 pairs.
Computing Categorical Statistics.
Computing Continuous Statistics.
Computing Scalar Partial Sums.
----------------------------------------
Output file: out/point_stat_360000L_20070331_120000V.stat
Output file: out/point_stat_360000L_20070331_120000V_fho.txt
Output file: out/point_stat_360000L_20070331_120000V_ctc.txt
Output file: out/point_stat_360000L_20070331_120000V_cts.txt
Output file: out/point_stat_360000L_20070331_120000V_cnt.txt
Output file: out/point_stat_360000L_20070331_120000V_sl1l2.txt
Output file: out/point_stat_360000L_20070331_120000V_sal1l2.txt
Output file: out/point_stat_360000L_20070331_120000V_vl1l2.txt
Output file: out/point_stat_360000L_20070331_120000V_val1l2.txt
Output file: out/point_stat_360000L_20070331_120000V_mpr.txt
```

Point-Stat: ASCII Output Types

- Statistics line types: 9 possible
  - Categorical - apply threshold
    - Contingency table counts (**FHO**, **CTC**)
    - Contingency table statistics (**CTS**)
  - Continuous - raw fields
    - Continuous statistics (**CNT**)
    - Partial Sums (**SL1L2**, **SAL1L2**, **VL1L2**, **VAL1L2**)
  - Matched Pairs
    - Raw matched pairs – a lot of data! (**MPR**)
  - Additional line type(s) for probabilistic methods
- 21 header columns common to all line types
- Remaining columns specific to each line type

Point-Stat: Sample Output

- STAT file output for sample run:
  - 1 line each for **FHO**, **CTC**, **CTS**, **CNT**, and **SL1L2**
  - 11,370 lines for **MPR**!
- 6 additional TXT files with lines sorted by type

Point-Stat: CTC Output Line

VERSION	V2.0
MODEL	WRF
FCST_LEAD	360000
FCST_VALID_BEG	20070331_120000
FCST_VALID_END	20070331_120000
OBS_LEAD	000000
OBS_VALID_BEG	20070331_103000
OBS_VALID_END	20070331_133000
FCST_VAR	TMP
FCST_LEV	Z2
OBS_VAR	TMP
OBS_LEV	Z2
OBTYPE	ADPSFC

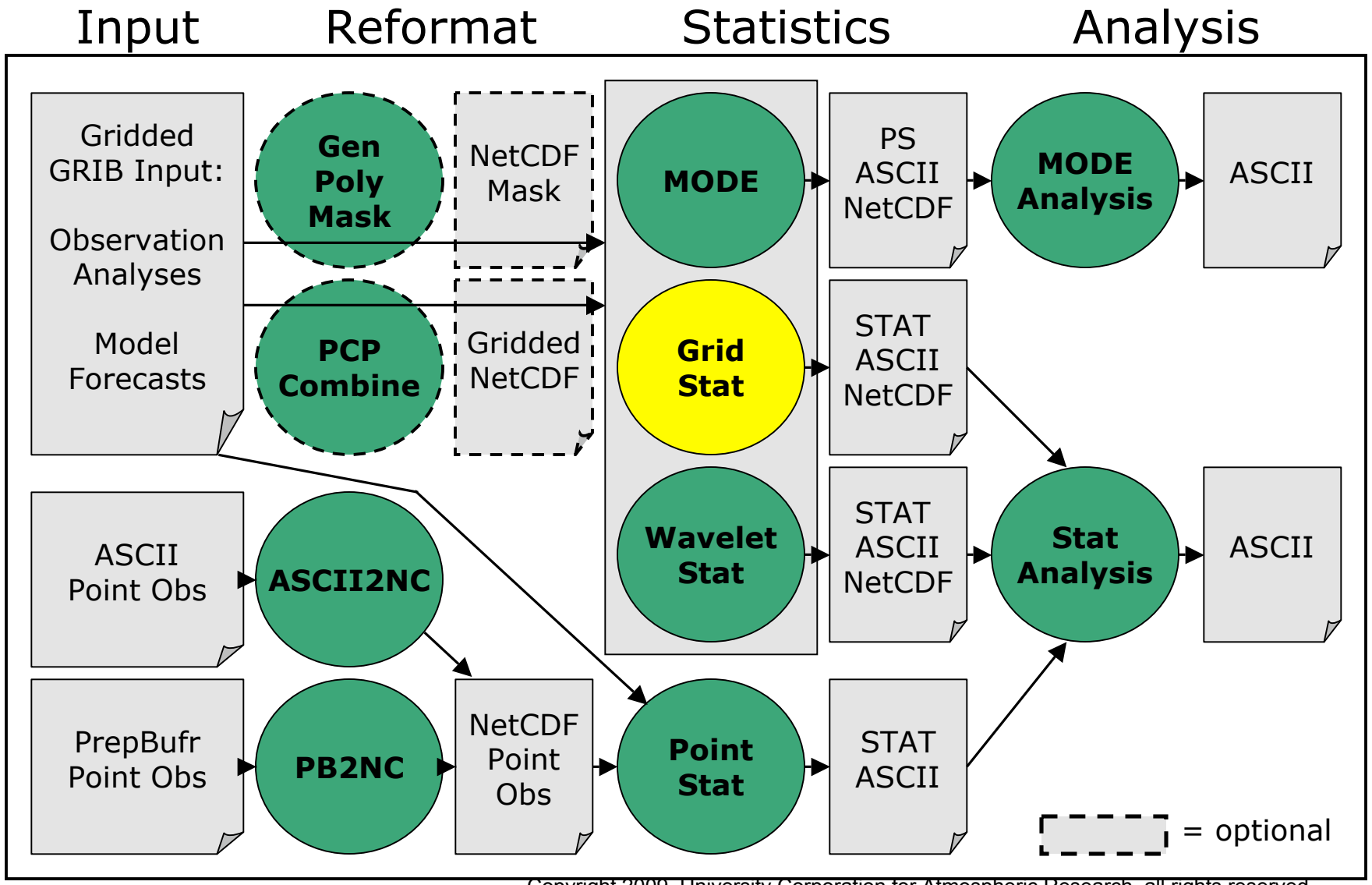
VX_MASK	FULL
INTERP_MTHD	UW_MEAN
INTERP_PNTS	1
FCST_THRESH	<273.000
OBS_THRESH	<273.000
COV_THRESH	NA
ALPHA	NA
LINE_TYPE	CTC
TOTAL	11370
FY_OY	1635
FY_ON	380
FN_OY	438
FN_ON	8917

Point-Stat: Matched Pairs

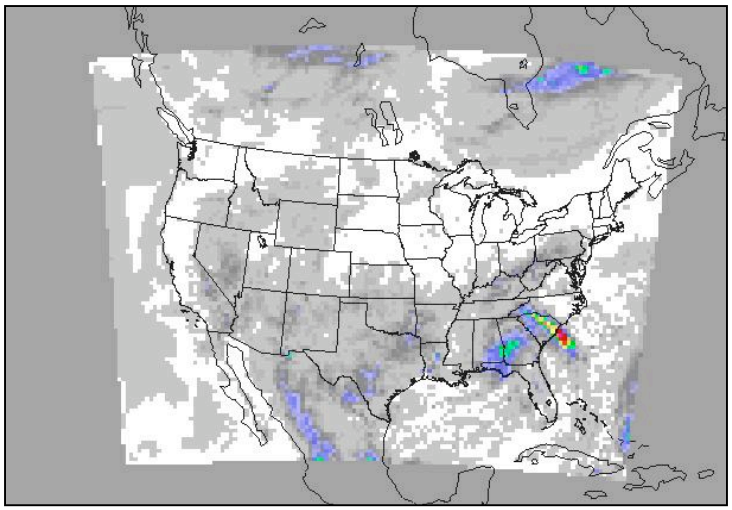
- **Matched Pair (MPR) line type contains raw matched pairs.**
  - **Data overload!**

TOTAL	INDEX	OBS_LAT	OBS_LON	OBS_LVL	OBS_ELV	FCST	OBS	CLIMO
11370	1	43.93000	-60.01000	1010.79999	4.01053	271.87788	271.54999	NA
11370	2	46.43000	-71.93000	1016.09998	102.04903	268.50255	269.45001	NA
11370	3	44.23000	-78.36000	1004.50000	191.44466	272.94013	272.35001	NA
11370	4	51.67000	-124.40000	916.50000	872.82202	263.69020	264.95001	NA
11370	5	58.61000	-117.16000	973.90002	337.50449	272.37757	270.95001	NA
11370	6	52.18000	-122.04000	906.50000	938.08594	271.37738	264.35001	NA
11370	7	50.68000	-127.36000	1020.20001	22.03931	275.44020	275.04999	NA
11370	8	50.45000	-100.59000	949.09998	562.38477	272.18978	271.75000	NA
11370	9	57.13000	-61.47000	899.70001	834.87476	258.06464	254.64999	NA
11370	10	47.56000	-59.16000	1000.90002	40.06803	272.06486	269.54999	NA
11370	11	47.41000	-72.79000	1006.90002	169.37592	266.37724	265.95001	NA
11370	12	45.76000	-62.68000	1014.00000	1.99518	268.94018	268.64999	NA
11370	13	49.24000	-65.33000	1014.90002	28.96468	264.25276	267.25000	NA
11370	14	43.29000	-79.79000	1017.79999	77.03765	273.56474	275.85001	NA
11370	15	48.78000	-123.04000	1015.70001	23.93772	278.12724	280.25000	NA

Grid-Stat Tool

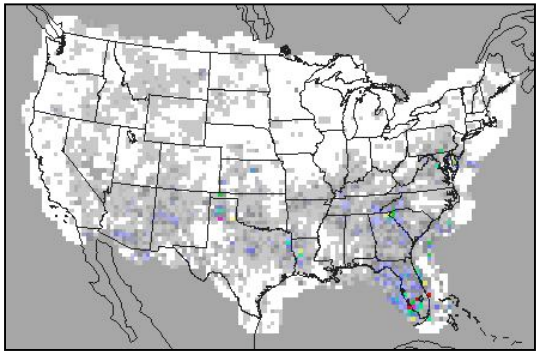


Grid-Stat: Overview



- 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.
  - Categorical counts and statistics for thresholded fields.
  - Parametric and non-parametric confidence intervals for statistics.
  - Compute partial sums for raw fields.
  - Probabilistic methods.
  - Continuous statistics and categorical counts/statistics using neighborhood verification method.



Grid-Stat: Input/Output

- Input Files
  - Gridded forecast and observation files
    - GRIB output of WRF Post-Processor (or other)
    - NetCDF output of PCP-Combine
  - 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: point_stat

fcst_file

obs_file

config_file

[-outdir path]

[-v level]

fcst_file	Forecast file in GRIB or NetCDF
obs_file	Observation file in GRIB or NetCDF
config_file	ASCII configuration file
-outdir	Output directory to be used
-v	Level of logging

Grid-Stat: Configuration

- 24 configurable parameters – only set a few:
  - Precipitation accumulated over 24 hours.
    - **fcst_field[] = [“APCP/A24“];**
  - Any rain and moderate rain.
    - **fcst_thresh[] = [“gt0.00, gt20.00“]; (mm)**
  - Look at all the points and only the eastern United States.
    - **mask_grid[] = ["FULL"];**
    - **mask_poly[] = [“EAST.poly"];**
  - Compute neighborhood statistics with two sizes.
    nbr_width[] = [3, 5];
  - Generate all possible statistic types.
    - **output_flag[] = [2, 2, 2, 2, 2, 2, 2, 2, 2];**

Grid-Stat: Run

- METv2.0/bin/grid_stat \
  sample_fcst.grb sample_obs.nc \
  GridStatConfig_APCP24 -outdir out -v 2

```
Forecast File: sample_fcst.grb
Observation File: sample_obs.nc
Configuration File: GridStatConfig_APCP24
---------------------------------------
Processing APCP/A24 versus APCP/A24, for interpolation method UW_MEAN(1), over region FULL, using 6412 pairs.
Computing Categorical Statistics.
Computing Continuous Statistics.
Processing APCP/A24 versus APCP/A24, for interpolation method UW_MEAN(1), over region EAST, using 2586 pairs.
Computing Categorical Statistics.
Computing Continuous Statistics.
Processing APCP/A24 versus APCP/A24, for interpolation method NBRHD(9), raw thresholds of >0.000 and >0.000,
over region EAST, using 5829 pairs.
... MORE NEIGHBORHOOD COMPUTATIONS ...
---------------------------------------
Output file: out/grid_stat_240000L_20050808_000000V.stat
Output file: out/grid_stat_240000L_20050808_000000V_fho.txt
Output file: out/grid_stat_240000L_20050808_000000V_ctc.txt
Output file: out/grid_stat_240000L_20050808_000000V_cts.txt
Output file: out/grid_stat_240000L_20050808_000000V_cnt.txt
Output file: out/grid_stat_240000L_20050808_000000V_sl1l2.txt
Output file: out/grid_stat_240000L_20050808_000000V_nbrctc.txt
Output file: out/grid_stat_240000L_20050808_000000V_nbrcts.txt
Output file: out/grid_stat_240000L_20050808_000000V_nbrcnt.txt
Output file: out/grid_stat_240000L_20050808_000000V_pairs.nc
```

Grid-Stat: ASCII Output Types

- Statistics line types: 8 possible
  - Same as Point-Stat
    - **FHO**, **CTC**, **CTS**, **CNT**, and **SL1L2**
  - Omitted for Grid-Stat
    - ***SAL1L2***, ***VL1L2***, or ***VAL1L2***
  - Neighborhood – define neighborhood, apply threshold
    - Neighborhood continuous statistics (**NBRCNT**)
    - Neighborhood contingency table counts (**NBRCTC**)
    - Neighborhood contingency table statistics (**NBRCTS**)
  - Additional line type(s) for probabilistic methods
- 21 header columns common to all line types
- Remaining columns specific to each line type

Grid-Stat: Sample Output

- STAT file output for sample run:
  - 2 lines each for **CNT** 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
- 8 additional TXT files with lines sorted by type
- NetCDF file containing matched pairs

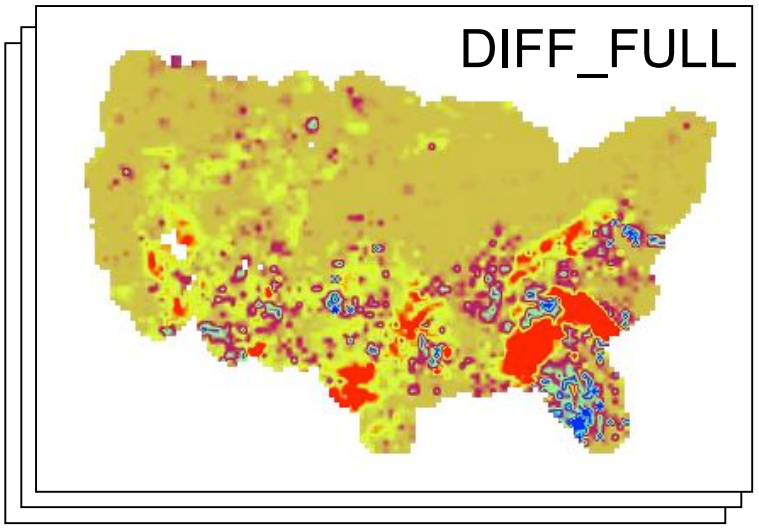
Grid-Stat: CTC Output Line

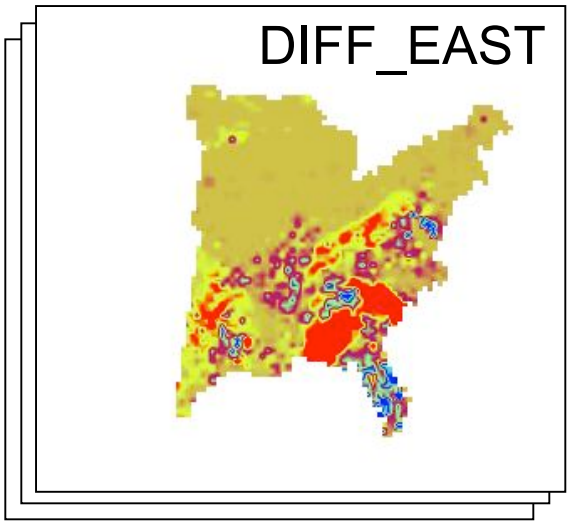
VERSION	V2.0
MODEL	WRF
FCST_LEAD	240000
FCST_VALID_BEG	20050808_000000
FCST_VALID_END	20050808_000000
OBS_LEAD	000000
OBS_VALID_BEG	20050808_000000
OBS_VALID_END	20050808_000000
FCST_VAR	APCP
FCST_LEV	A24
OBS_VAR	APCP
OBS_LEV	A24
OBTYPE	MC_PCP

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	2586
FY_OY	5
FY_ON	104
FN_OY	70
FN_ON	2407

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 variables:
  - FCST, OBS, and DIFF for **FULL** and **EAST**





Variations

- **Grid-Stat** and **Point-Stat** may be used to compare two *different* variables.
  - Not strictly verification.
  - Leave observation blank to use forecast setting.
  - Selecting variable/levels:
    - **fcst_field[] = ["61/A24"];**
    - **obs_field[] = [];**
  - Selecting thresholds:
    - **fcst_thresh[] = ["gt0.0 ge20.0"];**
    - **obs_thresh[] = [];**