Stat-Analysis Tool

- Filtering
- Summarizing
- Aggregating
- Ramp

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

Tina Kalb

What can Stat Analysis do?

Questions to MET Help - Can I get…

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

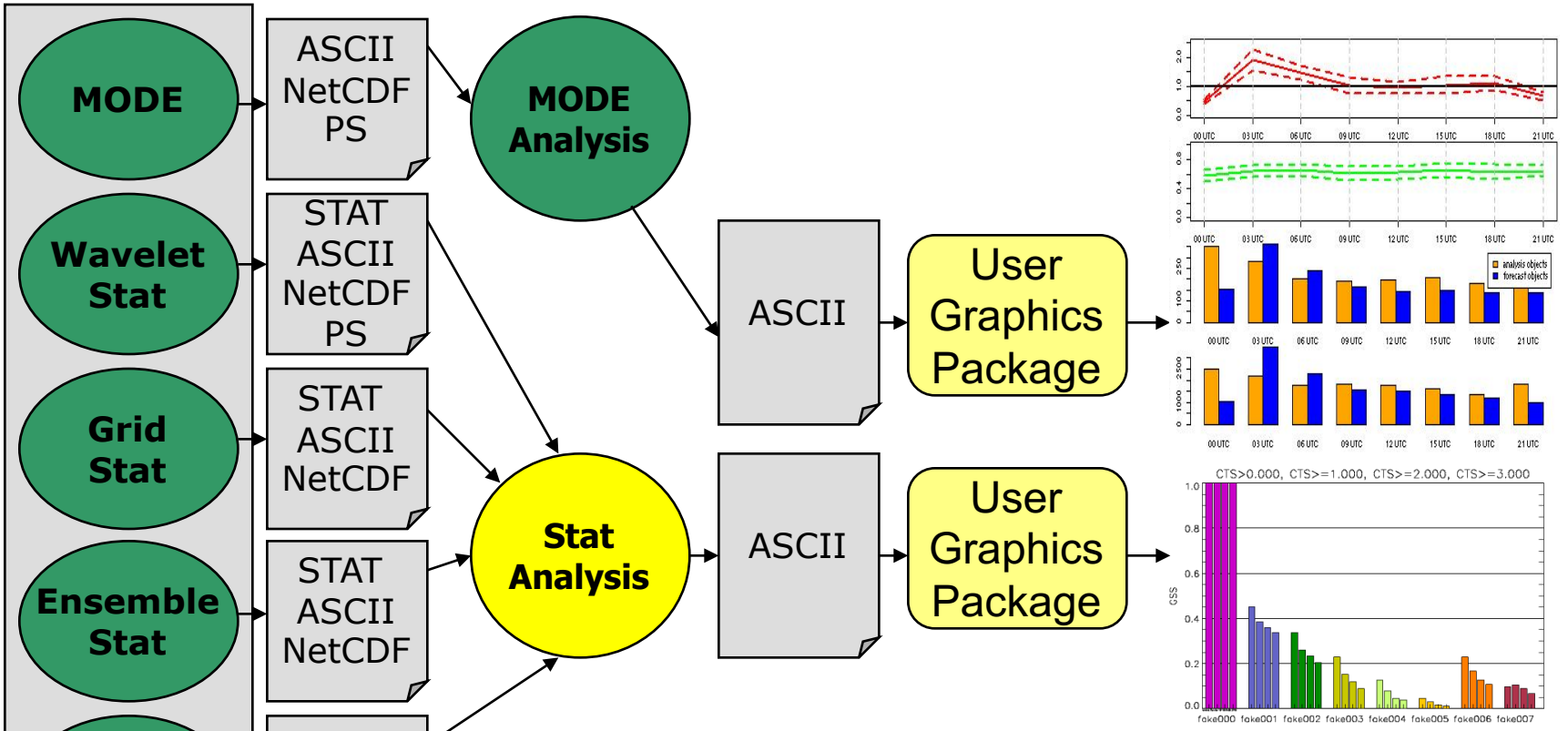
Q: Overall statistics for gridded observations compared to forecasts, hours 0 - 24?
A: Using Stat Analysis Tool on Grid-Stat output

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

Statistics Analysis User Defined Display



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 (10th, 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
ORANK	yields	RHIST, PHIST, RELP, SSVAR
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]

[ramp]

-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
ramp	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: Filtering

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

"-job aggregate -line_type CTC -fcst_var UGRD -vx_mask DTC165 -vx_mask DTC166
 -interp_mthd DW_MEAN"
```

- Can be a job itself, or as user specified options
- -dump_row gives an output file
  - Check intended subset
- Filtering options
  - Top of configuration file
  - In “jobs” section

Stat Analysis Tool: Run –job aggregate

```
-line_type required
"-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

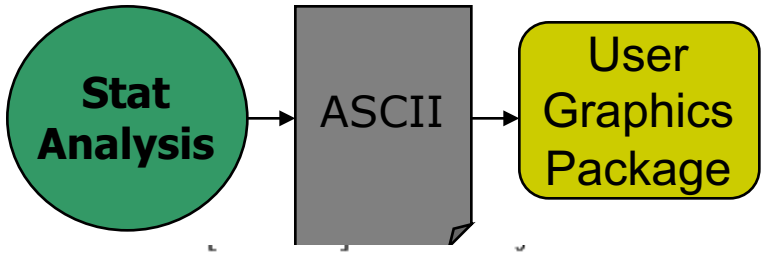
```
"-job summary -fcst_var UGRD -interp_mthd DW_MEAN -line_type CTS -column GSS
-dump_row out/job_summary.stat"
```

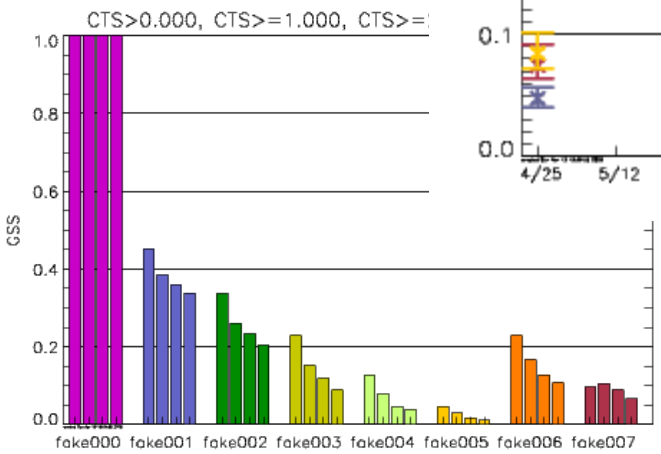
#	Description	
1	Column Name	Summary
2	Total	3
3-7	Mean* ***Includes normal and bootstrap upper and lower confidence limits***	*0.109*
8-10	Standard deviation** ***Includes bootstrap upper and lower confidence limits***	*0.150*
11	Minimum value	*-0.036*
12	10th percentile	-0.008
13	25th percentile	0.032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

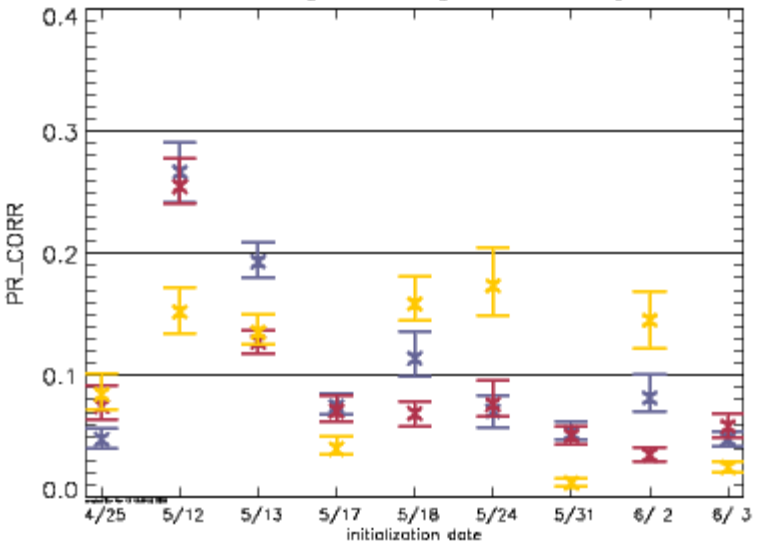
Summary Output *(stat_analysis.out cont.)*

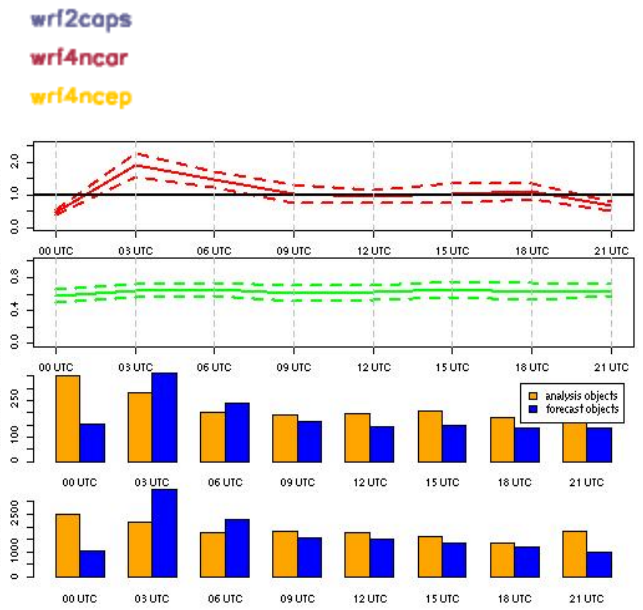
```
COL_NAME: TOTAL MEAN
MEAN_NCL MEAN_NCU MEAN_BCL
MEAN_BCU STDEV   STDEV_BCL
STDEV_BCU MIN      P10
P25     P50     P75     P90
MAX
SUMMARY: 3     0.10963
-0.26321 0.48247  -0.03613
0.26370  0.15009 0.00000
0.17311   -0.03613 -0.00864
0.03259 0.10131 0.18251
0.23122 0.26370
```

Use your favorite plotting software



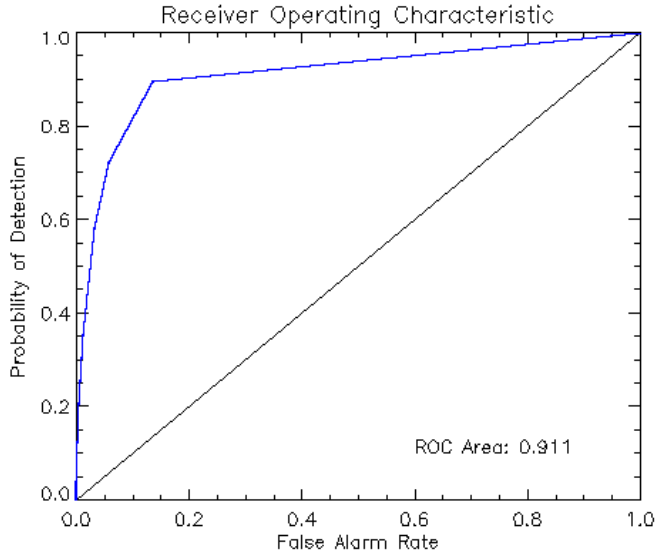


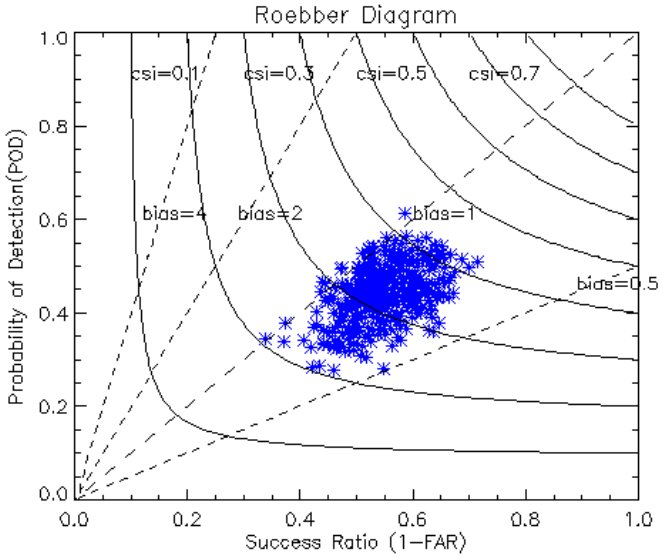


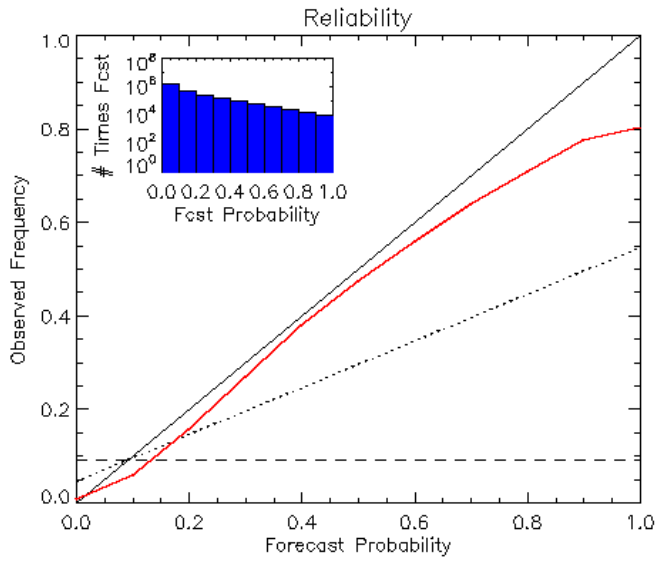


MET provides the analysis in ascii ouput.
You provide the graphing / plotting capability.

Stat_Analysis Example





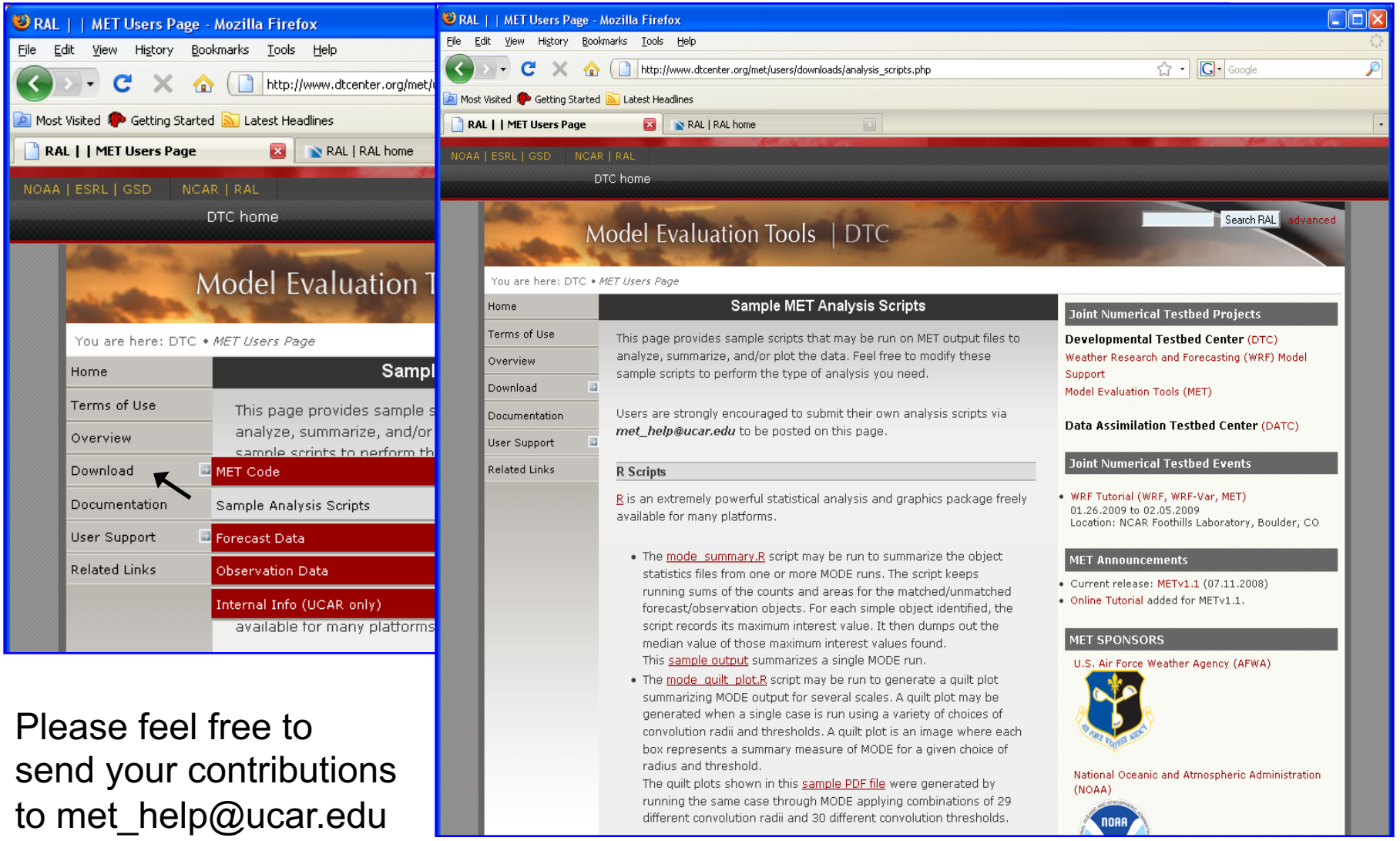


03/01/2013 - 08/30/2013
fcst_lead = “240000”
fcst_valid_hr = [“000000”,”120000”]

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