DART Tutorial Section 16:
Diagnostic Output

NCAR | National Center for
UCAR | Atmospheric Research

This tutorial section has not yet been updated for the Manhattan release of DART.

A new version is expected in April 2017.

DART Diagnostic Output Categories:

- State-Space:
  Values of model's state vector and inflation.
  Output using netCDF format.

- Observation-Space:
  Values of the observations.
  DART-specific *obs_sequence* format for now.

- Regression confidence factor:
  Values for state vector / observation pairs.
  Output as flat ASCII (soon to be netCDF).

- Program diagnostic output:
  Identification for source code version and namelist values.
  Error, warning, message output from modules.

State-Space Diagnostic Files:

Available in netCDF (a common data format)
    http://www.unidata.ucar.edu/software/netcdf

1. Prior state   (Prior_Diag.nc)   : state before assimilation.
2. Posterior state   (Posterior_Diag.nc) : state after assimilation.
3. Truth   (True_State.nc)   : truth for OSSEs.

Contents of prior and posterior controlled by *filter_nml*:
1. output_state_ens_mean = .true.   (include ensemble mean);
2. output_state_ens_spread = .true.   (include ensemble spread);
3. num_output_state_members = *##*   (include this many of the individual
                                              ensemble members)
4. output_interval = *N*   (only output every N^{th} assimilation time)

Note: output_interval for True_State.nc is in the *perfect_model_obs_nml* namelist.

DART State-Space Diagnostic functions

See the DART website section titled: "Configuring Matlab to work with DART"
www.image.ucar.edu/DAReS/DART/DART_Documentation.php#configure_matlab

ALL the DART Matlab state-space diagnostic functions are in *<dart>*/matlab
This **must** be in your *matlabpath*.

Only focus on the functions/scripts that start with *plot_*

- plot_bins.m
- plot_correl.m
- plot_ens_err_spread.m
- plot_ens_mean_time_series.m
- plot_ens_time_series.m
- plot_phase_space.m
- plot_reg_factor.m   Some, but not all, described here.
- plot_sawtooth.m   All functions have a 'help' section
- plot_smoother_err.m   available in the standard Matlab way.
- plot_total_err.m
- plot_var_var_correl.m
- …

1. Standard DART matlab diagnostics:
   a. plot_bins: rank histograms,
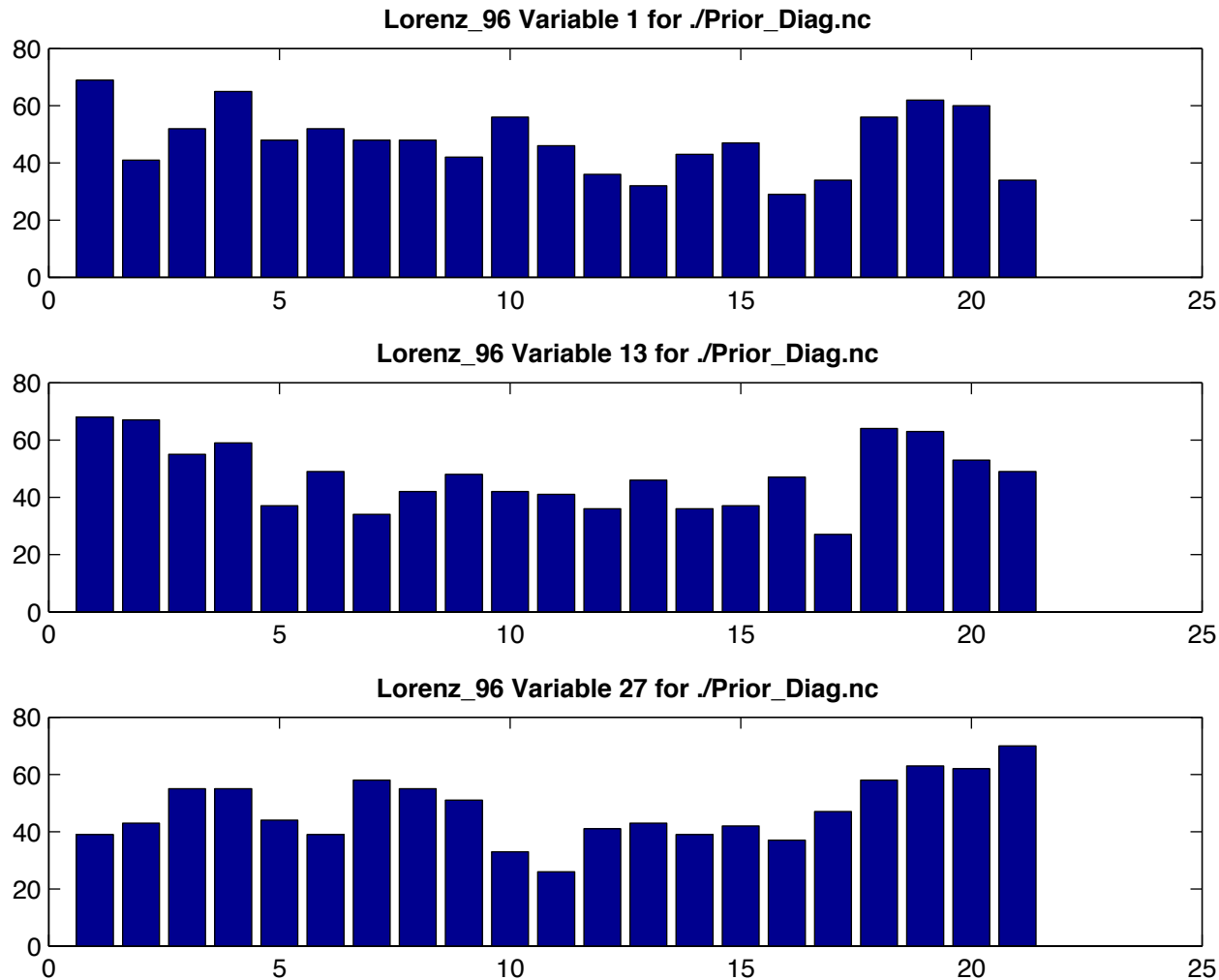
1. Standard DART matlab diagnostics:
   b. plot_correl: correlation x(t) with all other state vars at all times,

Lorenz_96 Correlation of variable state index 20, T = 50 of ./Prior_Diag.nc against all variables, all times, all 20 ensemble members
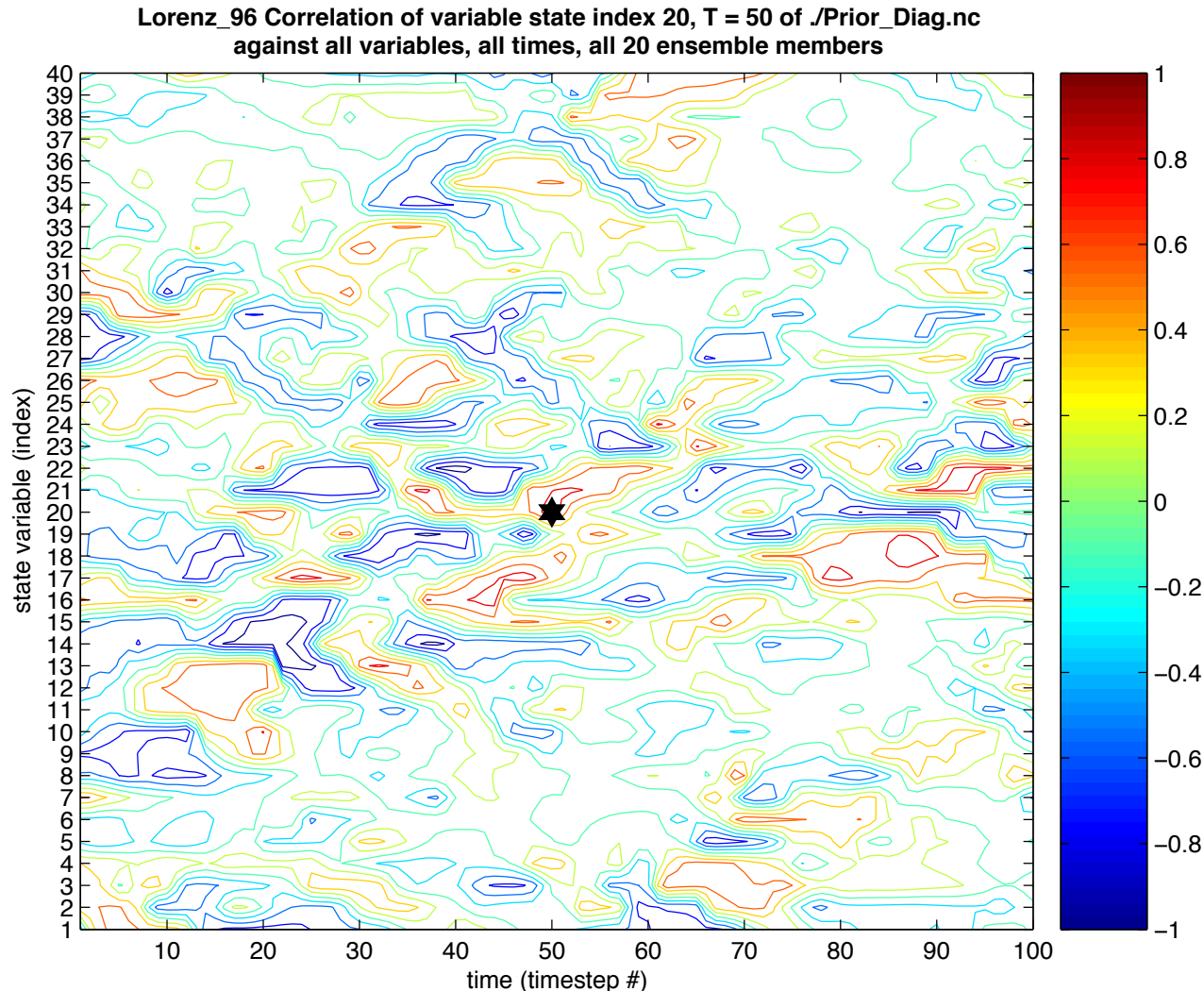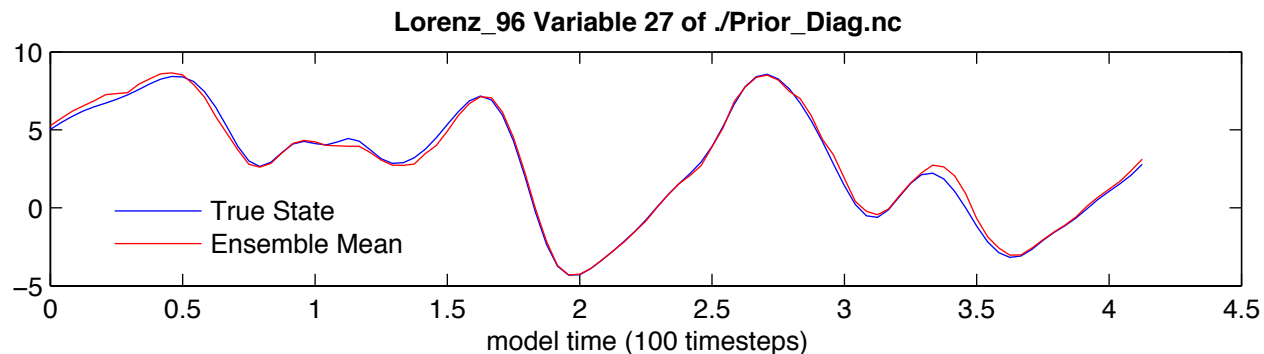
1. Standard DART matlab diagnostics:
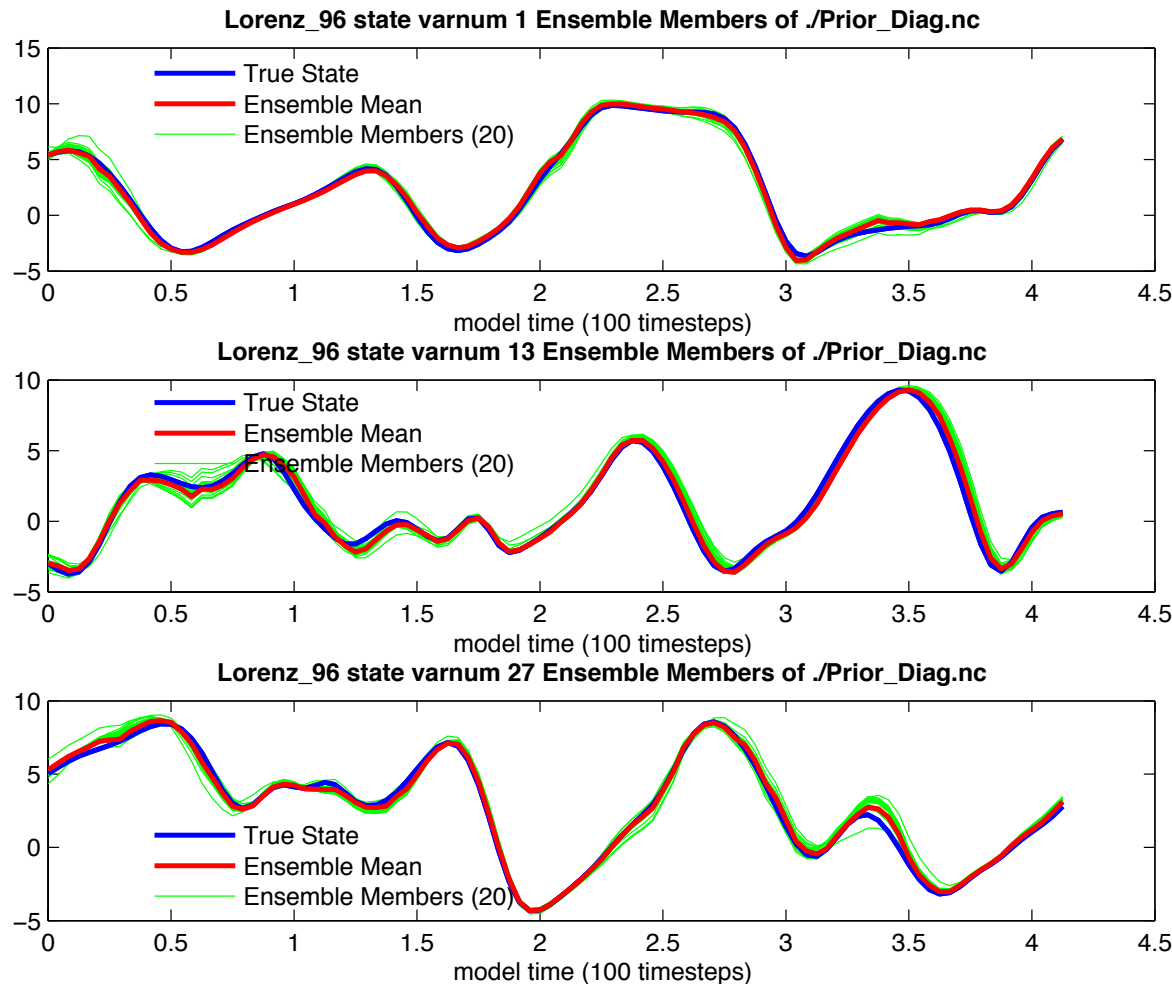   c. plot_ens_err_spread: rms error and spread,
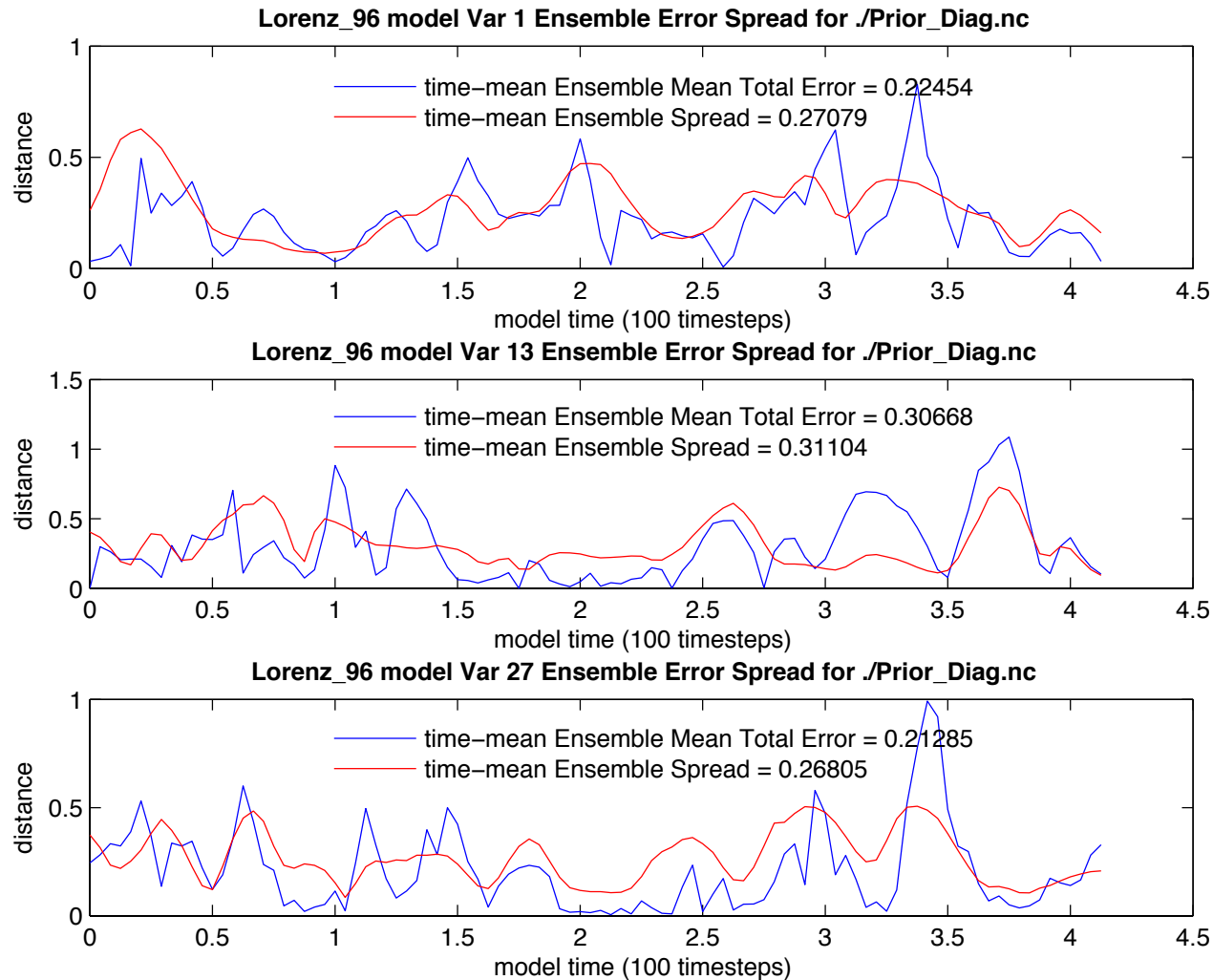
1. Standard DART matlab diagnostics:
    d. plot_ens_mean_time_series: just like the name says,
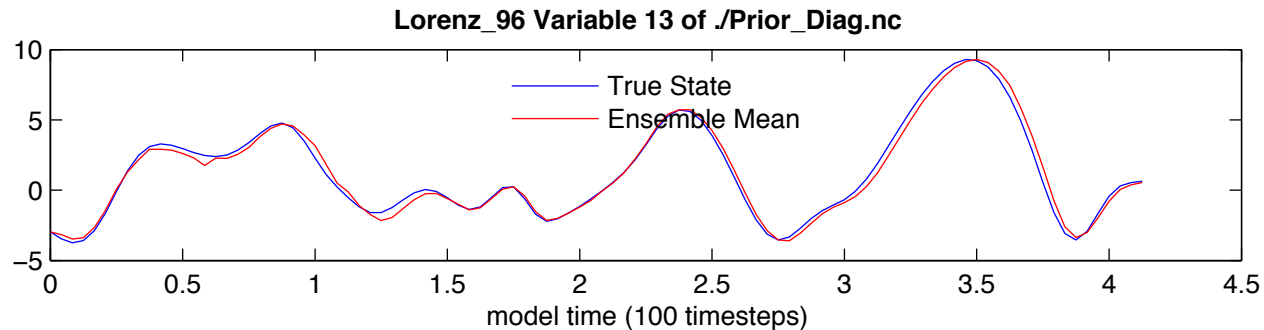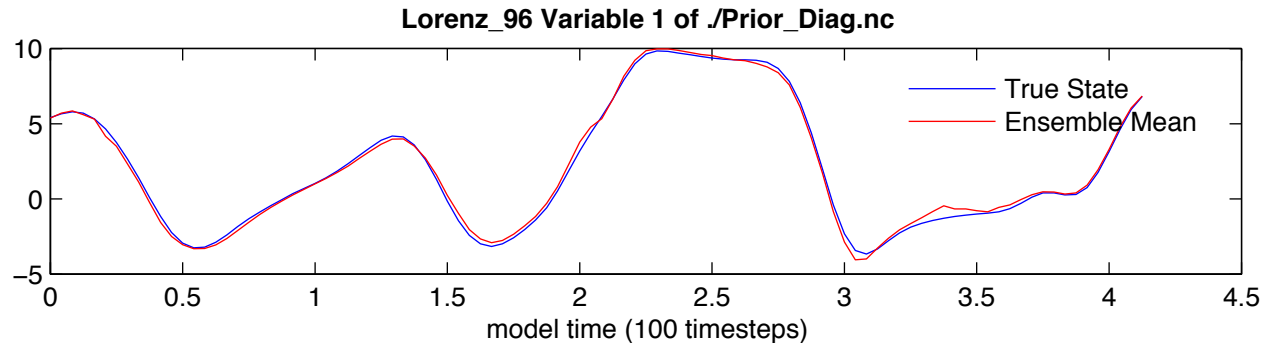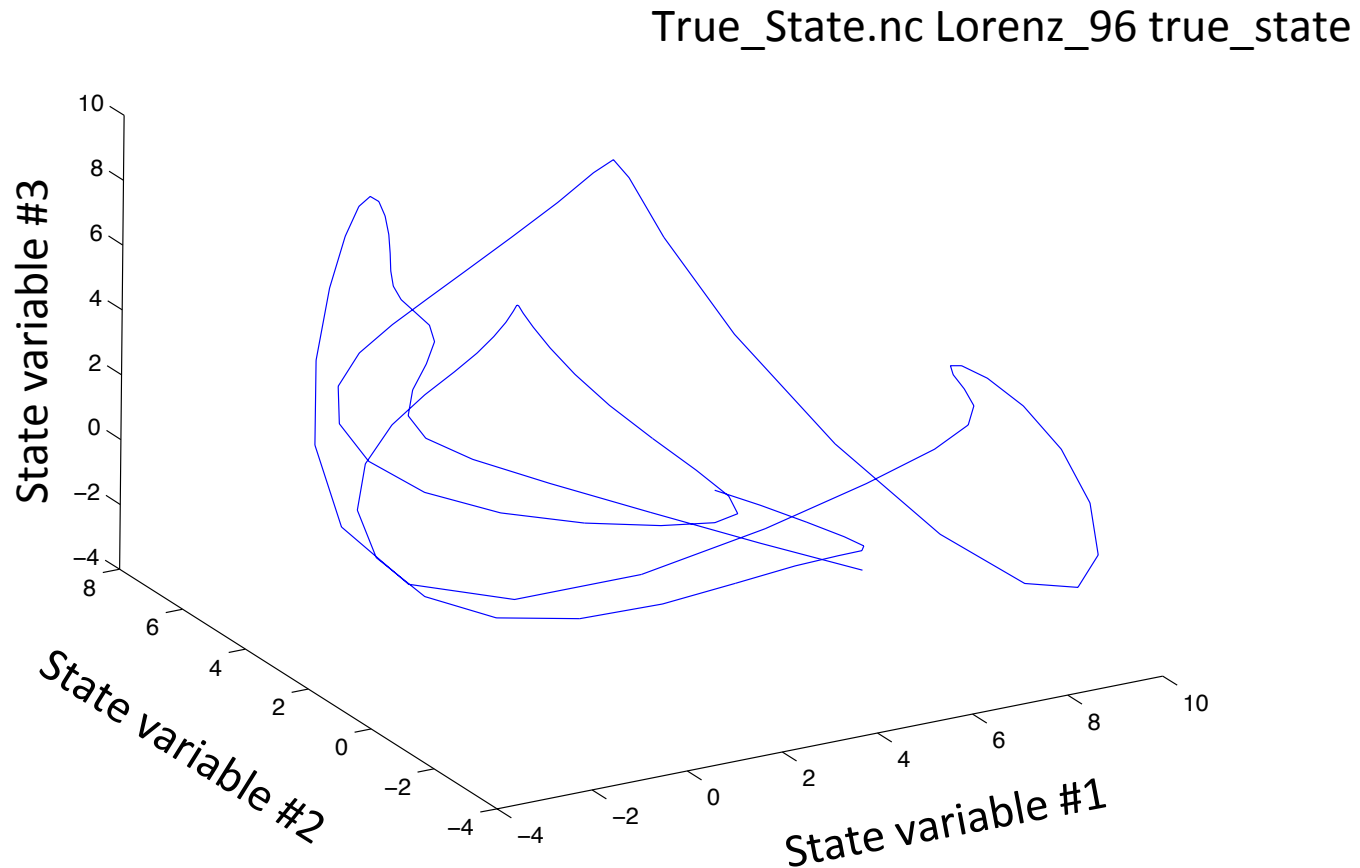
1. Standard DART matlab diagnostics:
   e. plot_ens_time_series: plots the ensemble
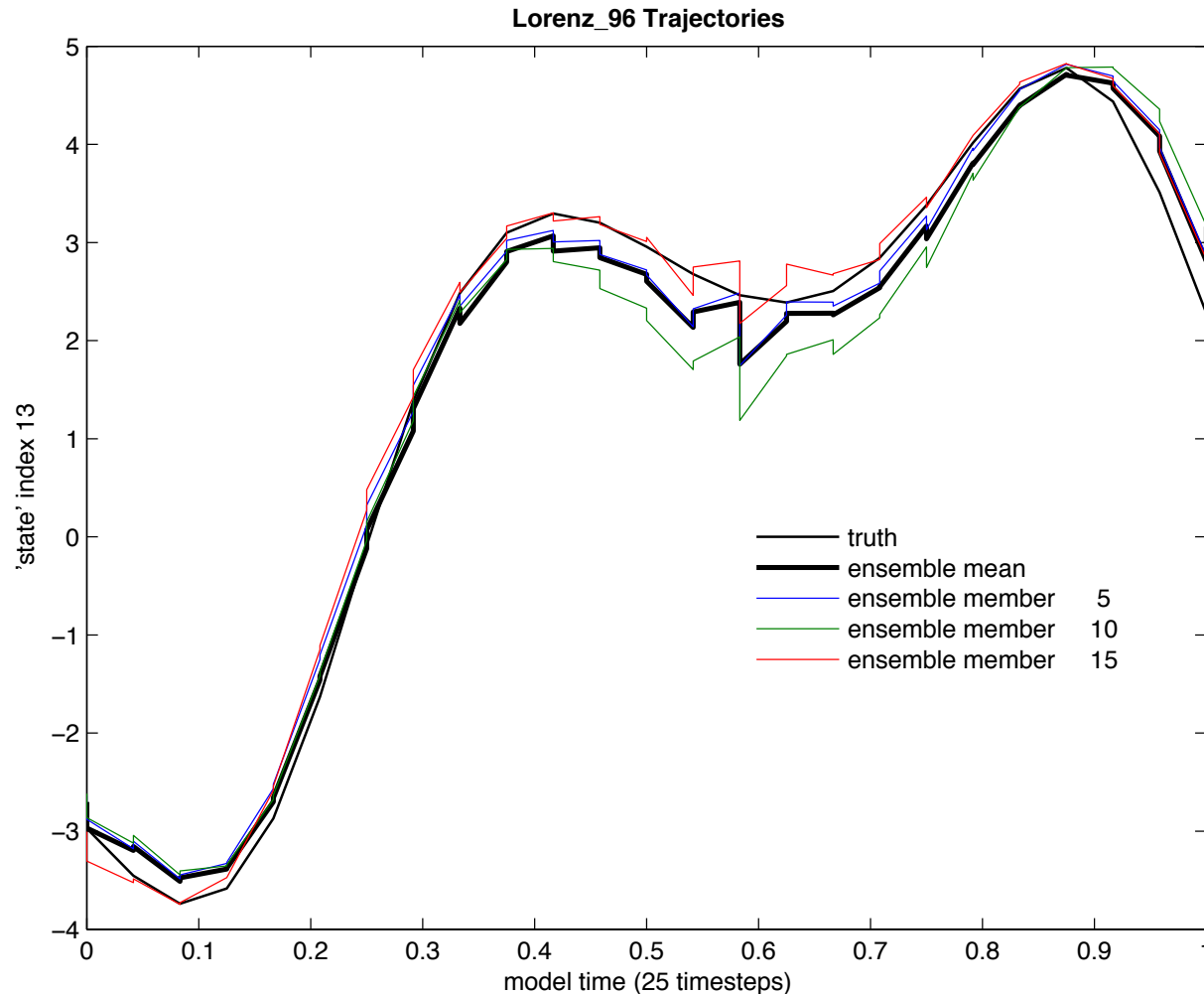   (as available from *num_output_state_members*) ,

1. Standard DART matlab diagnostics:
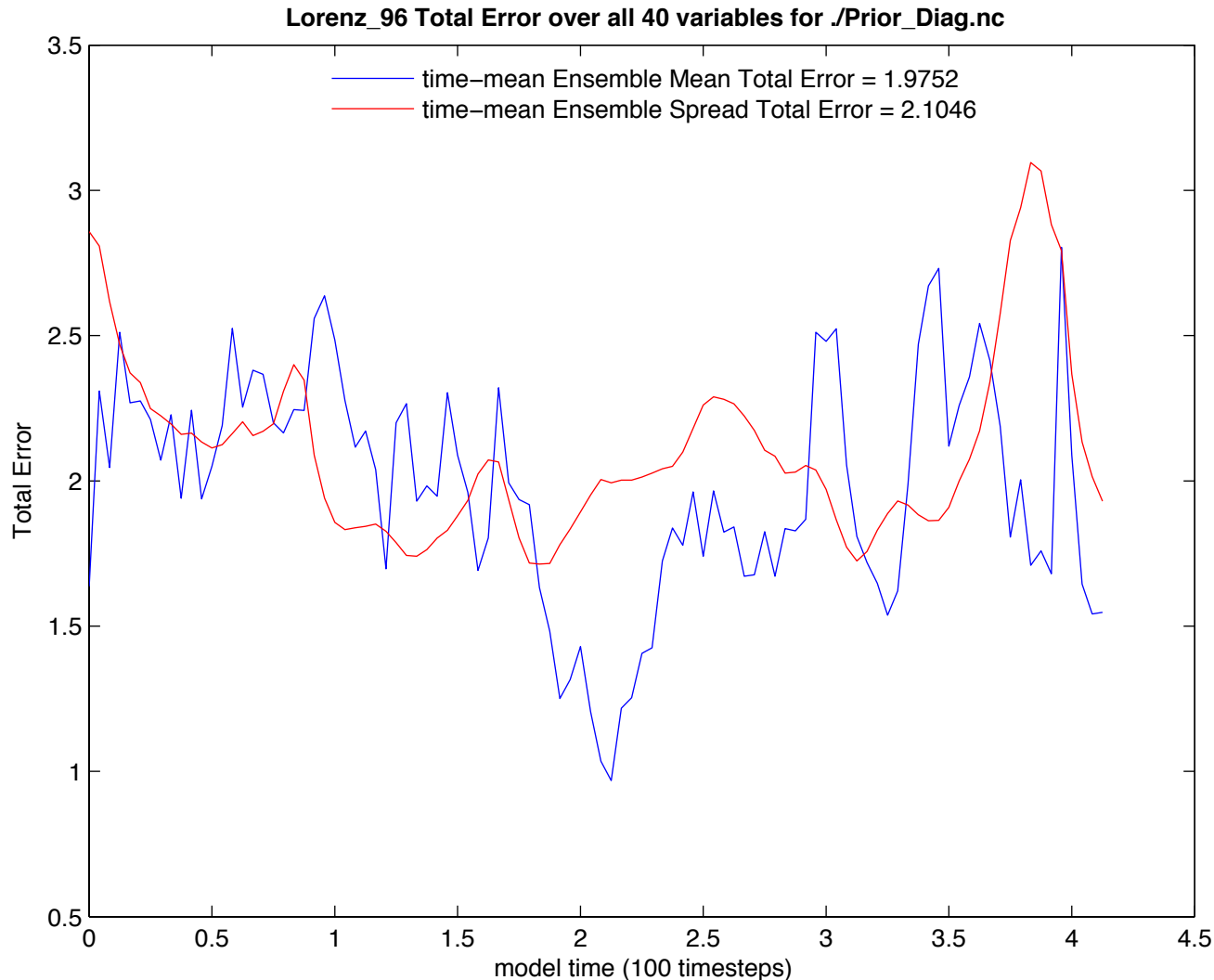   f. plot_phase_space: 3D phase space time evolution.

True_State.nc Lorenz_96 true_state

1. Standard DART matlab diagnostics:
   g. plot_sawtooth: truth, prior and posterior time series.

Lorenz_96 Trajectories

1. Standard DART matlab diagnostics:
   h. plot_total_err: total error for different fields,

Lorenz_96 Total Error over all 40 variables for ./Prior_Diag.nc

1. Standard DART matlab diagnostics:
   i. plot_var_var_correl: x(t) correlation to single variable, all times.

Lorenz_96 Correlation of variable state 20, T = 50, with variable state 18
20 ensemble members –– Prior_Diag.nc
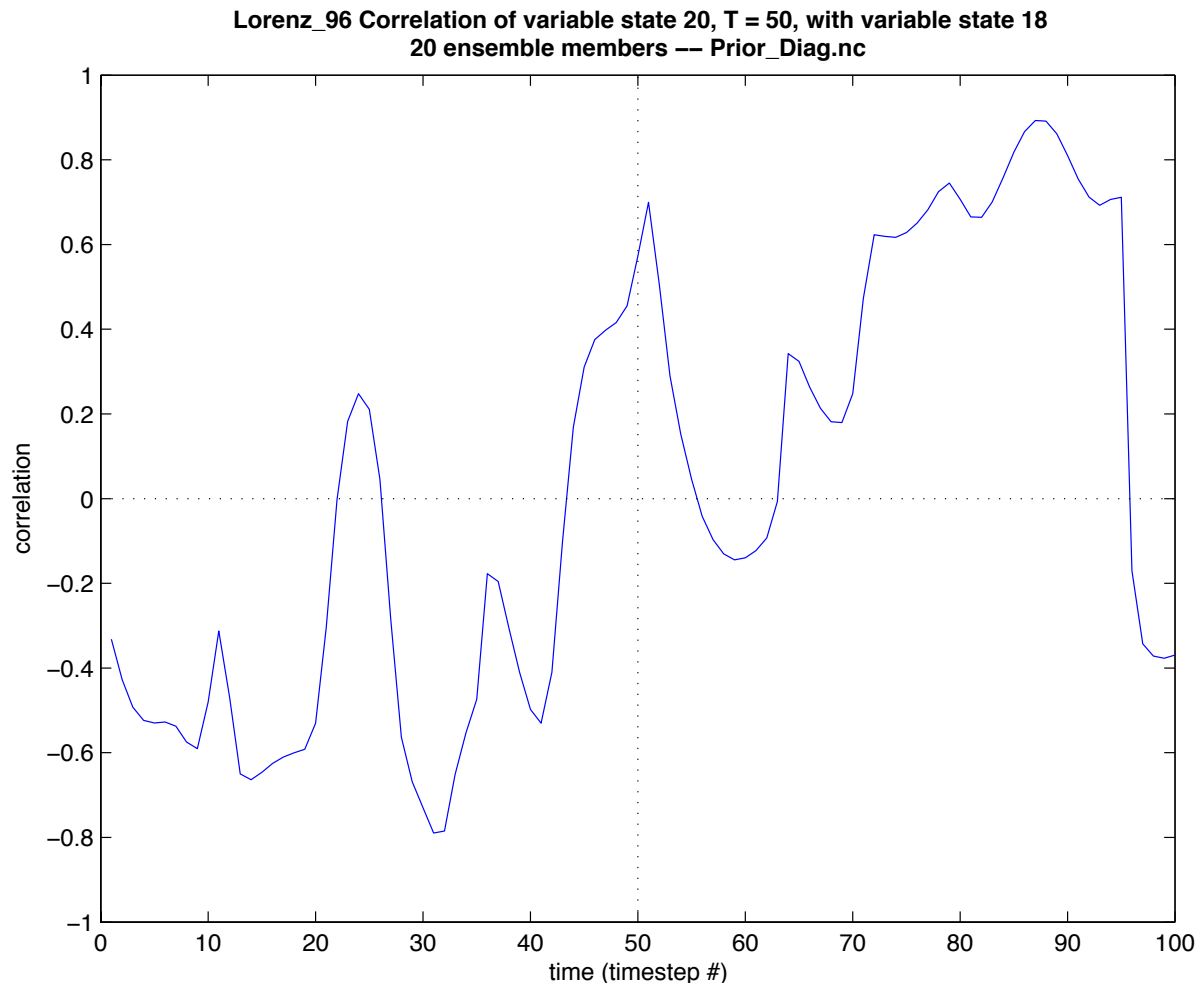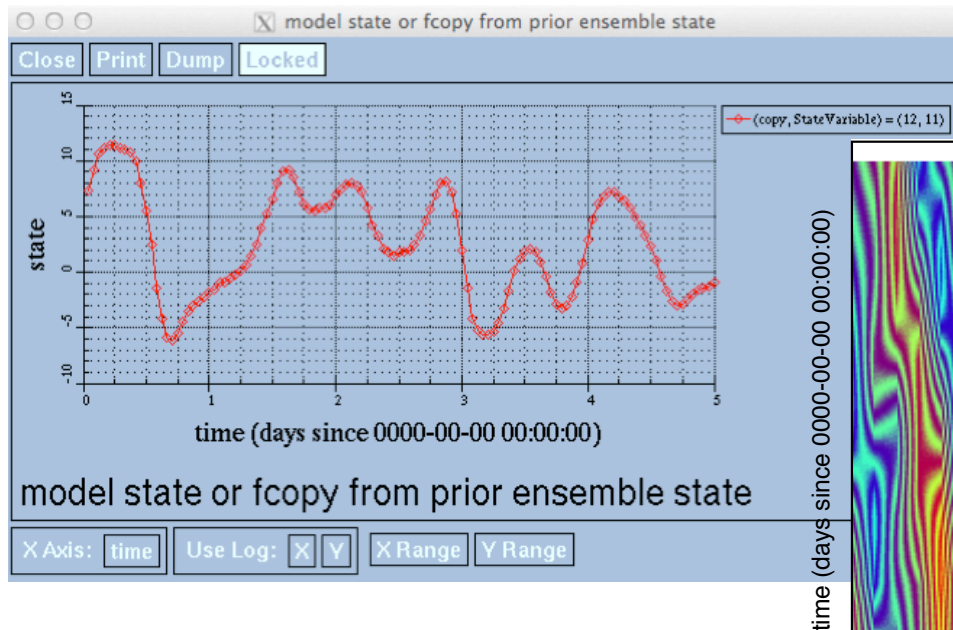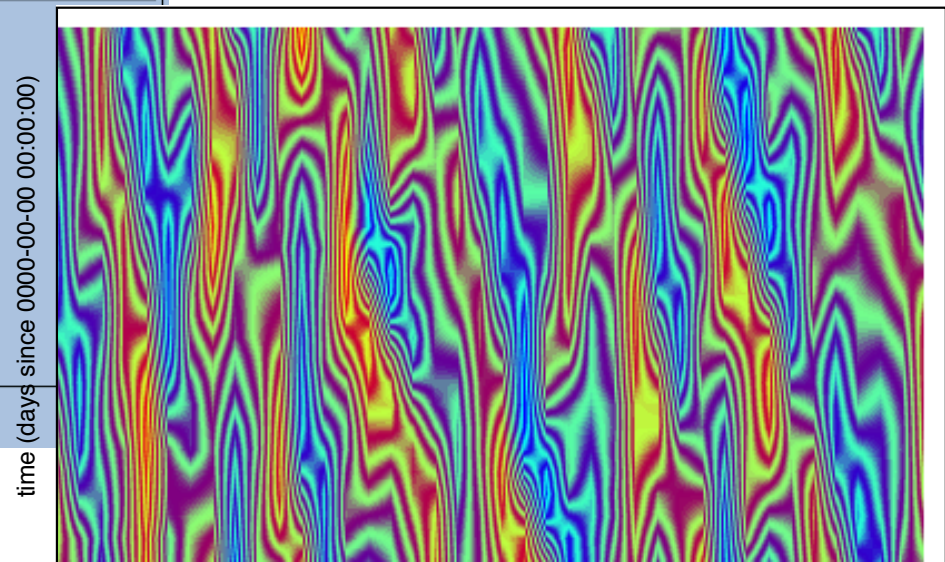
2. Ncview: a quick and surprisingly useful netCDF viewer.
   http://meteora.ucsd.edu/~pierce/ncview_home_page.html

   Displays spatial slices, animations, time series …

model state or fcopy

State Variable ID (indexical)

prior ensemble state
Range of model state or fcopy: -6.18328 to 11.6954 (null)
Range of State Variable ID: 1 to 40 indexical
Range of time: 0 to 1 days since 0000-00-00 00:00:00
Current ensemble member or copy: 1 nondimensional
Frame 1 in File Prior_Diag.nc

3. Many other graphical/analysis programs can read netCDF.
    (Note that we use *udunits* metadata convention.)

4. netCDF Operator (NCO) tools allow operations on netCDF files:
    (http://nco.sourceforge.net)
    Selecting hyperslices of fields,
    Differencing netCDF file,
    Averaging, etc.

netCDF Operators (NCO) Software Stack
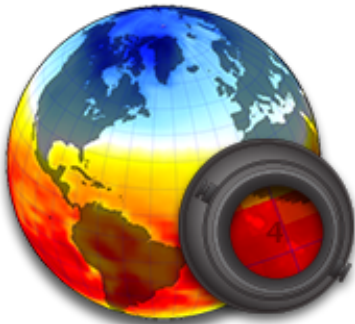
NASA GISS: Panoply
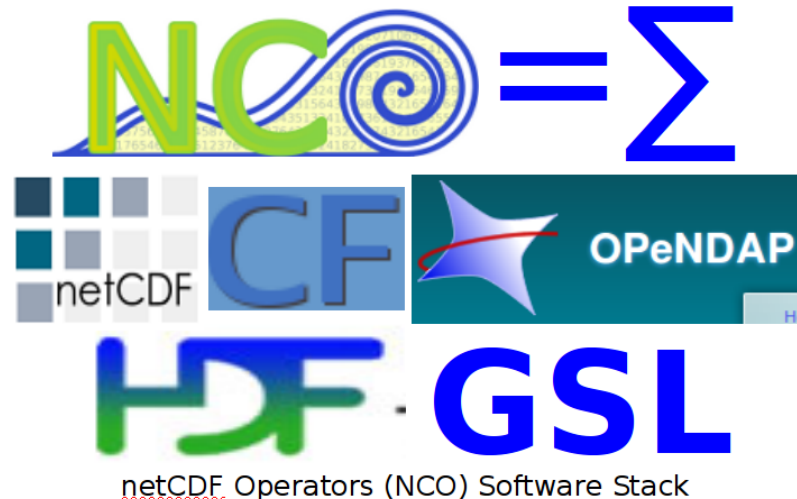
Observation-space files:

Quick recap of 'standard' observation sequence file names
(all names are actually specified in namelists):

- obs_seq.in   input to *perfect_model_obs*
- obs_seq.out   **output** from *perfect_model_obs*, also **input** to *filter*
- obs_seq.final   output from *filter*

Observation sequence file output by *filter* has prior, posterior, observed value
(and truth for OSSEs). For an overview, check out the DART webpage section:
www.image.ucar.edu/DAReS/DART/DART_Observations.php#obs_seq_overview

Contents of *obs_seq.final* controlled by filter_nml:
1. *obs_sequence_in_name* = 'obs_seq.out'
        Name of input observation sequence file.
2. *obs_sequence_out_name* = 'obs_seq.final'
        Name of output observation sequence file.
3. *num_output_obs_members* = ##
        Output this many individual ensemble estimates.

Observation-space diagnostics:

The observation sequence file is not in a particularly user-friendly format.
To aid in the evaluation and interpretation, a program named **obs_diag** must be run to produce a netCDF file with results that can be plotted in a manner of your choosing. DART has Matlab functions/scripts that create high-quality graphics.

See tutorial section 18 for full coverage of viewing / diagnosing obs sequences.
Also covered in:
http://www.image.ucar.edu/DAReS/DART/DART_Documentation.php#obs_diagnostics

Here are a few of the Matlab functions available in *<dart>*/diagnostics/matlab

- plot_rank_histogram.m
- plot_evolution.m
- plot_rmse_xxx_evolution.m
- two_experiments_evolution.m   (works with more than two, actually)
- plot_profile.m
- plot_bias_xxx_profile.m
- plot_rmse_xxx_profile.m
- two_experiments_profile.m   (works with more than two, actually)

Observation-space diagnostics:

SOME of the information in the observation sequence files can be converted to netCDF and easily plotted. A program named *obs_seq_to_netcdf* must be run to produce the netCDF.

Here are a few of the Matlab functions available in *<dart>*/diagnostics/matlab.

- link_obs.m
- plot_obs_netcdf.m
- plot_obs_netcdf_diffs.m
- plot_coverage.m
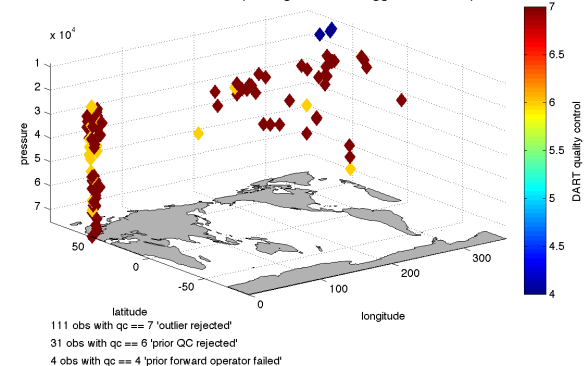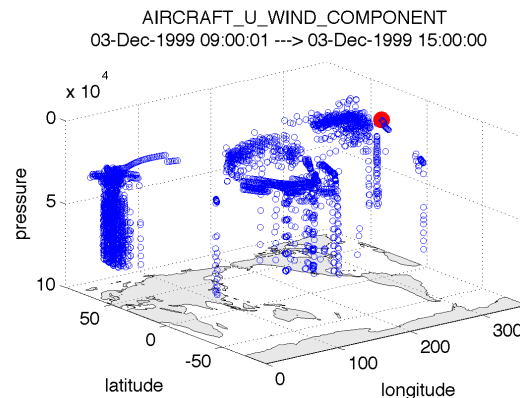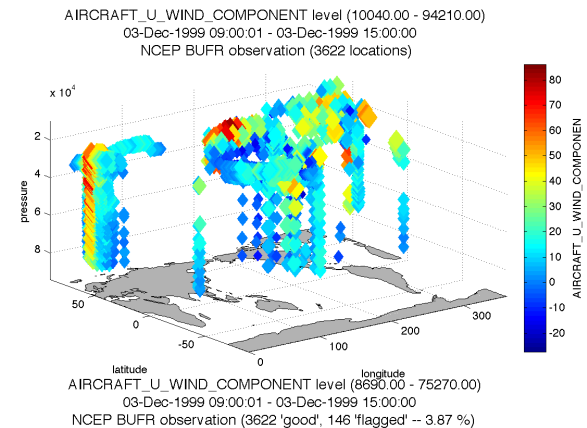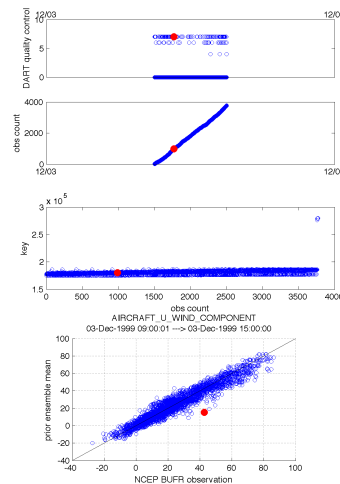
Regression confidence factor output:

Reminder: reg_factor α introduced in Tutorial Section 13 – when running the group filter (with more than 1 group!).

Controlled by reg_factor_nml:
1. save_reg_diagnostics = .true.   Should file be output?
2. reg_diagnostcs_file = 'reg_diagnostics'   Name of output file.

File size could be (model size) X (number of obs.) X (number of assim times).
    Very big, even for small models (only first 4 obs output default).

Normally, modify code in reg_factor_mod.f90 to control:
    Output is at end of select_regression = 1 code block.
    Format is ASCII:
        time in days, time in seconds, obs_index, state_index, α

Plot with Matlab *plot_reg_factor*.

File *dart_log.out*

All DART executables **append** to this file!

Contains:

- registration information
- Program start time,
- version of code for each module used*
- Namelist values for each module**
- Names of output files,
- Diagnostic output for modules (through error_handler()),
- Warnings and fatal errors from DART code.

<u>Fair Warning</u>: This file is **not** cleared by DART. Can get very longgggggg …
You should feel free to delete/rename it before starting the next experiment.

*Hopefully
**may be in a separate file, depending on *utilities_nml* setting

DART Tutorial Index to Sections