Data Assimilation Research Testbed Tutorial

Section 16: Diagnostic Output

Version 1.0: June, 2005
DART Diagnostic Output Categories:

1. State space:
   Values of models state vector.
   Output using netCDF format.

2. Observation space:
   Values of the observations.
   DART specific obs_sequence format for now.

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

4. Program diagnostic output:
   Identification for source code version and namelist values.
   Error, warning, message output from modules.
State Space Diagnostic Files:

Available in netCDF (common data format)
(http://my.unidata.ucar.edu/content/software/netcdf/index.html).

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 = 20,
      (include this many of the individual ensemble members);
   4. output_interval = 2, (only output every nth assimilation time);

Output interval for True_State.nc in perfect_model_obs_nml.
Viewing the State Space netCDF files:

1. Standard DART matlab diagnostics:
   a. plot_bin: rank histograms,

![Histograms for Lorenz_96 variables](image)
b. **plot_correl**: correlation $x(t)$ with all other state vars at all times,
c. **plot_ens_err_spread**: rms error and spread,

![Lorenz_96 model Var 1 Ensemble Error Spread for ./Prior_Diag.nc graph](image1)

- **time-mean Ensemble Mean Total Error** = 0.22454
- **time-mean Ensemble Spread** = 0.27079

![Lorenz_96 model Var 13 Ensemble Error Spread for ./Prior_Diag.nc graph](image2)

- **time-mean Ensemble Mean Total Error** = 0.30668
- **time-mean Ensemble Spread** = 0.31104

![Lorenz_96 model Var 27 Ensemble Error Spread for ./Prior_Diag.nc graph](image3)

- **time-mean Ensemble Mean Total Error** = 0.21285
- **time-mean Ensemble Spread** = 0.26805
d. `plot_ens_mean_time_series`: time series of ens. mean,
e. `plot_ens_time_series`: time series all available ensembles,
f. **plot_phase_space**: 3D phase space time evolution,
g. **plot_sawtooth**: truth, prior and posterior time series,
h. plot_total_err: total error for different fields,

![Graph showing Lorenz 96 Total Error over all 40 variables for ./Prior_Diag.nc](image)

- time-mean Ensemble Mean Total Error = 1.9752
- time-mean Ensemble Spread Total Error = 2.1046
i. **plot_var_var_correl**: \( x(t) \) correlation to single variable, all times.
Viewing the State Space netCDF files:

2. ncview: a quick, dirty, but surprisingly useful netCDF viewer. (http://meteora.ucsd.edu:80/~pierce/ncview_home_page.html) Displays spatial slices, time series,...

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

model state or fcopy from prior ensemble state
Viewing the State Space netCDF files:

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

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

Observation sequence file output by filter has prior, posterior, observed value, (and truth for OSSEs).

Contents of obs_sequence.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 final observation sequence.
3. output_obs_ens_mean = .true.,
   Output the ensemble mean observation estimate.
4. output_obs_ens_spread = .true.,
   Output the ensemble spread.
5. num_output_obs_members = 0,
   Output this many individual ensemble estimates.
(See section 18 for viewing obs_sequence files).
Regression confidence factor output:

Controlled by reg_factor_nml:
  1. save_reg_diagnostics = .true., should file be output?
  2. save_reg_diagnostics = .false., name of output file.

File size could be (model size) x (number obs.) x (number 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.
Program Diagnostic Output:

File *dart_log.out*

Written to by all DART executables.

Contains registration information:
- Program start time,
- CVS 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.

Watch out: This file is NOT cleared by DART. Can get longgggggg...