GCA-M2O2-v1 User Guide

2014.03.01

Meemong Lee

Jet Propulsion Laboratory, California Institute of Technology

1. Directory Structure

GCA-M2O2 is an assimilation system developed as an extension to the GEOS-Chem-Adjoint system. The directory is organized in four areas, M2O2_SERVER, L2#, Restart, and GCA-M2O2-v1. The content of each area is described below.

1.1 **M2O2_SERVER** / contains the source codes and executables for downloading the L2# data from the GESDISC-M2O2 web services. Users can download the L2# datasets prior to running GCA-M2O2 or use the datasets for their own assimilation applications. There are three stand-alone L2# download tools, MLSO3, AIRSCO, and ACOSxCO2 where the tool name convention is mission name + component name. Each tool performs two types of web-service interfaces, L2# data request and L2# data downloading. The command-line-interface is described in the **M2O2_SERVER Examples** section of this manual.

1.2 **L2#**/ contains the data sets needed for the assimilation. The data is organized by mission, each mission has sub-folders populated by the daily NetCDF L2# datasets. The file structure is "mission-component/yyyy/mm/". For example, the data for an assimilation of Ozone from the Micro-Limb-Sounder (MLS) of the month of July 2006 should be in the path "L2#/MLS-O3/2006/07/" and the files will be named "MLS-20060701-O3.nc", "MLS-20060702-O3.nc", etc. You will find more information on the **M2O2 Server Examples** Section of this manual.

1.3 **Restart** / contains restart files used by the application examples. The restart files are organized by simulation mode and model resolution. The simulation mode includes CH4, CO2, and Full Chemistry and the resolution and the model resolution includes 2°x2.5° and 4°x5°.

1.4 **GCA-M2O2-v1**/ contains four sub-folders, Code/, Data/, Runs, and V34/ (the version number of the GEOS-Chem-Adjoint system that GCA-M2O2-v1 integrates).

1.4.1 **Code**/ sub-folder is designed to simplify the interface between the GESO-Chem-Adjoint and M2O2 software development. It contains the soft-links to the V34 source code (original, adjoint, modified, and new) and two additional software folders, m2o2code/ and cmscode/. The "run" script is provided to utilize the Geos-Chem-adjoint 's "make" process within the directory structure.

1.4.2 **Data**/ sub-folder contains the input data for running GCA-M2O2 (required by GEOS-Chem) that can be shared by all applications.

1.4.3   **Runs/** sub-folder contains example runs for 3DVAR, 4DVAR, and FWD application areas. The run example includes input configurations, run scripts, and run logs.

1.4.4   **V34/** sub-folder contains the source codes of GEOS-Chem-Adjoint version 34 . All of the source codes except the observation operators are integrated.

2. Assimilation Process

The assimilation process involves three simple steps: 1) verification of L2# dataset availability, 2) preparation of input configuration files and 3) formulation of a run script.

2.1 **Verification of L2# data availability** : Users must refer to the temporal coverage of the L2# data of each mission (specified in M2O2 wiki). The individual M2O2_SERVER utility also lists the temporal coverage of the corresponding L2# data (see section 4 for data download examples). When the "Load from M2O2 Server" option is set to be "T", the GCA-M2O2 program will automatically contact the corresponding web service at GESDISC and load the L2# data at the beginning of each month during the simulation-run. If the simulation period is set to be outside of the available temporal coverage, the program will abort.

2.2 **M2O2 input configuration**: "M2O2 MENU" is required for specifying assimilation options, output options, and L2# data options. The assimilation options include the assimilation duration (i.e., time step per assimilation process) , the name and id # of the tracer to be assimilated, the model uncertainty (fraction), and the maximum iteration count of the optimization process. The output options include saving the prior, posterior, and pseudo observations. The requested outputs are saved as daily NetCDF files for the specified assimilation time step. The L2# data options include automated M2O2 data downloading, the number of missions, and the information of the mission data products. The example M2O2 menu shown in Figure 1 illustrates a 3DVAR assimilation of MLS and TES Ozone data products with the specified levels to be assimilated and filename convention for each dataset.

```
%%% M2O2 MENU %%%
Assimilation duration (hrs)    : 2
tracer name & id               : O3 2
Model uncertainty              : 0.01
max. iteration count           : 8
Save Prior?                    : T PRIOR
Save Posterior?                : T POSTERIOR
Save Pseudo-L2?                : F PSEUDO-L2
Load from M2O2 Server?         : F
No. of missions                : 2
<----------------------  select? column? mission SLEV ELEV dataPath ---------------------->
 1. mission information        : T  F  MLS    30   47   /data/meemong/M2O2/L2#/MLS-O3/YYYY/MM/MLS-YYYYMMDD-O3.nc
 2. mission information        : T  F  TES     1   35   /data/meemong/M2O2/L2#/TES-O3/YYYY/MM/TES-v4-YYYYMMDD-O3.nc
```

Figure 1. Example M2O2 Menu

2.3 **Run script** : GCA-M2O2 employs four types of command line arguments to allow flexible specification of the run options in 4D-var assimilation applications :

- **'-i '** for forward model configuration file specification

- **'-a'** for adjoint model configuration file specification,

- **'-n'** for ptimization loop count specification, and

- **'-v'** for debug message enable specification .

Figure 2 shows an example run script for a 4DVAR assimilation application where the forward model configuration file is "m2o2.fwd" and the adjoint model configuration file is "m2o2.adj". The assimilation process is explicitly controlled with the optimization loop count and the verbose flag activating the debug messages during the run.

```
#!/bin/sh
../../Code/geos -i m2o2.fwd -a m2o2.adj -n 1 -v > log.debug
#../../Code/geos -i m2o2.fwd -a m2o2.adj -n 2 > log.2
#../../Code/geos -i m2o2.fwd -a m2o2.adj -n 3 > log.3
#../../Code/geos -i m2o2.fwd -a m2o2.adj -n 4 > log.4
exit
```

Figure 2. Example run script for 4D-var assimilation

3. Application Examples

GCA-M2O2 can be used in three different areas, 1) to generate global concentration maps assimilating sparse observations employing 3D-variational assimilation method; 2) to perform inverse modeling of fluxes or initial conditions employing the 4D-variational assimilation method; and 3) to perform forward modeling and extract pseudo observations for a given observation sample scenario.

3.1 **GCA-v34/Runs/3DVAR** : Global concentration map generation examples can be found under MLS, TES, and MLS+TES sub folders. The examples show global ozone concentration time series maps generated assimilating Microwave Limb Sounder (MLS), Tropospheric Emission Sounder (TES), and MLS+TES Ozone (O3) L2# datasets.

3.2 **GCA-v34/Runs/4DVAR** : Inverse modeling examples can be found under CMS/ sub folder. The examples show CO_2 flux estimation process applied for JPL's Carbon Monitoring System flux project (CMS-flux) assimilating column CO_2 (xCO_2) observations from the Atmospheric Carbon Observing System (ACOS) project (ACOS-xCO_2 L2# data).

3.3 **GCA-v34/Runs/FWD** : Pseudo observation application examples can be found under MLS folder. The pseudo observations imply concentration profiles extracted from the model

simulation closely aligned with the mission observation properties with respect to sampling time & location, and atmospheric-sounding pressure levels. Currently, this option is supported for MLS mission only.

4. M2O2_SERVER Examples

The L2# data can be downloaded via a stand-alone tool which interfaces with the mission-specific L2# server at GESDISC (Goddard Earth System Data and Information Service Center). There are three stand-alone L2# download tools under M2O2_SERVER/, MLSO3, AIRSCO, and ACOSxCO2 where the tool name convention is mission name + component name. The command line interface requires three arguments, 1) directory name for L2# data, 2) start date, and end date. When the requirement arguments are missing or incorrect, the tool will print out a help message. Figure 3 illustrates the usage of MLSO3 which download L2# data of Ozone from MLS mission. The example indicates that the L2# data between 2006/07/01 and 2006/07/31 should be downloaded at /data/meemong/M2O2/L2#/MLS-O3/ folder. The downloaded data will be organized under yyyy/mm/MLS-yyyymmdd-O3.nc as shown in Figure 4.

```
[meemong@prospero M2O2_SERVER]$ ./MLSO3
 Usage : MLSO3 m2o2Dir startDate endDate
    example : MLSO3 /data/meemong/M2O2/L2#/MLS-O3/ 20060701 20060731
             valid date range is from     20040801  to     20121231
 M2O2_ERROR: m2o2Dir argument error(          1 )
[meemong@prospero M2O2_SERVER]$ 
```

Figure 3. MLSO3 Command Line Example

```
[meemong@prospero L2#]$ cd MLS-O3
[meemong@prospero MLS-O3]$ ls
2006
[meemong@prospero MLS-O3]$ cd 2006
[meemong@prospero 2006]$ ls
07  08
[meemong@prospero 2006]$ cd 07
[meemong@prospero 07]$ ls
MLS-20060701-O3.nc  MLS-20060706-O3.nc  MLS-20060711-O3.nc  MLS-20060716-O3.nc  MLS-20060721-O3.nc  MLS-20060726-O3.nc  MLS-20060731-O3.nc
MLS-20060702-O3.nc  MLS-20060707-O3.nc  MLS-20060712-O3.nc  MLS-20060717-O3.nc  MLS-20060722-O3.nc  MLS-20060727-O3.nc
MLS-20060703-O3.nc  MLS-20060708-O3.nc  MLS-20060713-O3.nc  MLS-20060718-O3.nc  MLS-20060723-O3.nc  MLS-20060728-O3.nc
MLS-20060704-O3.nc  MLS-20060709-O3.nc  MLS-20060714-O3.nc  MLS-20060719-O3.nc  MLS-20060724-O3.nc  MLS-20060729-O3.nc
MLS-20060705-O3.nc  MLS-20060710-O3.nc  MLS-20060715-O3.nc  MLS-20060720-O3.nc  MLS-20060725-O3.nc  MLS-20060730-O3.nc
[meemong@prospero 07]$ 
```

Figure 4. L2# Data File Organization