Adjoint Clinic:
Introduction for new users.

International GEOS-Chem Meeting 10

Yanko Davila, Daven Henze, CU Boulder

Overview:

1. Introduction to adjoint modeling
2. Resources
3. Obtaining the code
4. Model structure and key files
5. Examples
    a. Inverse Modeling
    b. Sensitivity Test
    c. Finite Difference
6. The GitLab online interface
7. GCHP adjoint

2. Resources

GEOS-Chem adjoint wiki:

http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_Adjoint

Code management and distribution (aka GitLab):

http://adjoint.colorado.edu:8080/

User's Guide

https://tinyurl.com/456pvhj5

Background papers and presentations, code diagram, plotting tools, GitLab tutorial

https://tinyurl.com/mr2332df

3. Obtaining the Code

1) Sign up for the geos-chem-adjoint mailing list to receive updates about code bugs and developments
2) Review the policies on code use and distribution
3) Request a GIT account by sending an email to the mailing list
4) CHANGE YOUR PASSWORD by going to http://adjoint.colorado.edu:8080/profile/account
5) Generate your SSH Keys by going to http://adjoint.colorado.edu:8080/profile/keys
6) git clone git@adjoint.colorado.edu:yanko.davila/gcadj_std.git

4. Model structure and key files

code/
  define.h

adjoint/
define_adj.h: pre-processor include settings like observation operators
inverse_driver.f: top-level routine (inverse model)
geos_chem_adj_mod.f: top-level routine (adjoint model)
modified/
geos_chem_mod.f

runs/
v10-01b/
geos4/ - (default config for CO pseudo obs benchmark*)
geos5/ - (default config for tropchemistry global finite difference benchmark*)
ucx/ - (default config for fullchemistry global finite difference benchmark)
OptData/
diagadj/
adjtmp/
STR_ID
RXN_ID
input.gcadj
input.geos
run
* only available in v35n and earlier

.git/ - (Contain all the information about commits, history, log, etc.) DON'T REMOVE
.gitignore DON'T REMOVE

Key output files:
runs/v10-01b/ucx/
OptData/
cfn.** : cost function
gctm.gdt.** : gradients
gctm.sf.** : scaling factors
diagadj/
gctm.adj.** : tracer adjoints
gctm.fd.** : finite difference results
gctm.fdglob.** : finite difference results
ems.adj.** : non-normalized emissions sensitivities
gctm.iteration : iteration vs function evaluation
log : cost function from penalty term

The run script

```
#################################################################
# Set run parameters.
# - X
# - XSTOP
# - RNAME
#
# These need to be set and checked  before every run.
# X = 0 creates observations.  X_STOP < X only
# creates a backup of the program directory.
#################################################################
# Set the start (or current ) iteration number
X=1

# Set the stopping iteration number
XSTOP=3
# Give every run a unique name (default is $PBS_JOBID)
RNAME=gcadj_std

# Specify Type of Run "DEFAULT, HDF, SAT_NETCDF, LIDORT"
TYPE=DEFAULT

# Set Compiler Options. For a list of all compiler options
# type 'make help' on the code directory.
# example IFORT_OPT="DEBUG=yes TRACEBACK=yes"

IFORT_OPT=""

# Recompile geos on every iteration.
# NOTE: If you have IPO set to yes we recommend not to recompile.
# IPO optimization make linking process slow.

RECOMPILE=NO

# Save packages to temporary storage before and after execution
# Note: Need to set DSAVE below
SAVE=NO

# Archive packages to permanent storage after execution
# Note: Need to set DARCHIVE below
ARCHIVE=NO

# Set compiler, if necessary
#source /usr/projects/compilers/intel/9.1.043/bin/ifortvars.sh
```

Setting X and XSTOP

Sensitivity	Finite difference	Pseudo obs
X=1	X=1	X=0
XSTOP=1	XSTOP=3	XSTOP=10

Input.gcadj (default configuration for Finite Difference $\frac{\partial O3}{\partial NO}$)

```
GEOS-CHEM ADJOINT SIMULATION v35n
------------------------------+---------------------------------------------
%%% ADJOINT SIMULATION MENU %%%
Do adjoint run    LADJ         : T
Run decoupled adj LADJ_ONLY    : F
Selecet one simulation type    :---
Inverse problem   L4DVAR       : F
Kalman filter     L3DVAR       : F
Sensitivity       LSENS        : T
 => spot finite diff  FD_SPOT  : F
 => global finite diff FD_GLOB : T
------------------------------+---------------------------------------------
%%% FORWARD MODEL OPTIONS %%%
adjoint chemistry LADJ_CHEM    : T
aerosol thermo    LAERO_THEM   : T
 => ISORROPIAII                : F
------------------------------+---------------------------------------------
%%% ADJOINT MODEL OPTIONS %%%
Include a priori term APSRC    : F
 => offdiagonal                : F
Compute DFP inverse Hessian    : F
Compute BFGS inverse Hessian   : F
Include rxn rate sensitivities : F
Delete chk files LDEL_CHKPT    : T
Scale up and FILL adj transport: F
------------------------------+---------------------------------------------
%%% DIRECTORIES %%%
Optimization output            : OptData/
Temporary adjoint dir adjtmp   : adjtmp/
Diagnostics ouptut             : diagadj/
-------------------------------------------------------------------------------
%%% CONTROL VARIABLE MENU %%%
Initial conditions LICS        : T
... OR emissions   LADJ_EMS    : F
 => strat prod/loss LADJ_STRAT : F
 => reaction rates LADJ_RRATE  : F
>------------------------------<
          FOR LICS             :
NSOPT: number of tracers opt   : 1
  => opt these tracers------>  : TRC# trc_name SF_DEFAULT REG_PARAM ERROR
Tracer  #1                     : 1    NO       1          1         1
>------------------------------<
          FOR LADJ_EMS         :
NNEMS: ems groups implemented  : 39
Emission entries ------------> : EMS# ems_name       opt  SF_DEFAULT REG_PARAM ERROR CORR_LX CORR_LY
Emission #1                    : 1    IDADJ_ENH3_an  T    1          1         1       100     100
Emission #2                    : 2    IDADJ_ENH3_na  T    1          1         1       100     100
Emission #3                    : 3    IDADJ_ENH3_bb  T    1          1         1       100     100
Emission #4                    : 4    IDADJ_ENH3_bf  T    1          1         1       100     100
Emission #5                    : 5    IDADJ_ESO2_an1 T    1          1         1       100     100
Emission #6                    : 6    IDADJ_ESO2_an2 T    1          1         1       100     100
Emission #7                    : 7    IDADJ_ESO2_bf  T    1          1         1       100     100
Emission #8                    : 8    IDADJ_ESO2_bb  T    1          1         1       100     100
Emission #9                    : 9    IDADJ_ESO2_sh  T    1          1         1       100     100
Emission #10                   : 10   IDADJ_EBCPI_an T    1          1         1       100     100
Emission #11                   : 11   IDADJ_EBCPO_an T    1          1         1       100     100
Emission #12                   : 12   IDADJ_EOCPI_an T    1          1         1       100     100
Emission #13                   : 13   IDADJ_EOCPO_an T    1          1         1       100     100
Emission #14                   : 14   IDADJ_EBCPI_bf T    1          1         1       100     100
Emission #15                   : 15   IDADJ_EBCPO_bf T    1          1         1       100     100
Emission #16                   : 16   IDADJ_EOCPI_bf T    1          1         1       100     100
Emission #17                   : 17   IDADJ_EOCPO_bf T    1          1         1       100     100
Emission #18                   : 18   IDADJ_EBCPI_bb T    1          1         1       100     100
Emission #19                   : 19   IDADJ_EBCPO_bb T    1          1         1       100     100
Emission #20                   : 20   IDADJ_EOCPI_bb T    1          1         1       100     100
Emission #21                   : 21   IDADJ_EOCPO_bb T    1          1         1       100     100
Emission #22                   : 22   IDADJ_ENO_so   F    1          1         1       100     100
```

```
Emission #23                  : 23   IDADJ_ENO_li   F    1          1         1     100     100
Emission #25                  : 25   IDADJ_ENO_an   F    1          1         1     100     100
Emission #26                  : 26   IDADJ_ENO_bf   F    1          1         1     100     100
Emission #27                  : 27   IDADJ_ENO_bb   F    1          1         1     100     100
Emission #28                  : 28   IDADJ_ECO_an   F    1          1         1     100     100
Emission #29                  : 29   IDADJ_ECO_bf   F    1          1         1     100     100
Emission #30                  : 30   IDADJ_ECO_bb   F    1          1         1     100     100
Emission #31                  : 31   IDADJ_EISOP_an F    1          1         1     100     100
Emission #32                  : 32   IDADJ_EISOP_bf F    1          1         1     100     100
Emission #33                  : 33   IDADJ_EISOP_bb F    1          1         1     100     100
Emission #33                  : 33   IDADJ_ENO_ac   T    1          1         1     100     100
Emission #34                  : 34   IDADJ_ECO_ac   T    1          1         1     100     100
Emission #35                  : 35   IDADJ_ESO2_ac  T    1          1         1     100     100
Emission #36                  : 36   IDADJ_ESO4_ac  T    1          1         1     100     100
Emission #37                  : 37   IDADJ_EH2O_ac  T    1          1         1     100     100
Emission #38                  : 38   IDADJ_EOCPI_ac T    1          1         1     100     100
Emission #39                  : 39   IDADJ_EBCPI_ac T    1          1         1     100     100
Number emis time group MMSCL  : 1
>------------------------------<
        FOR LADJ_STRAT        :
NSTPL: strat prod & loss trcs : 0
Read reactions from STR_ID file: T
Strat prod & loss trc entries : ID# trc_name    opt   SF_DEFALUT  REG_PARAM ERROR
>------------------------------<
         FOR LADJ_RRATE       :
NRRATES: num of rxn rates     : 0
Read reactions from RXN_ID file: T
 ...or use these Rxn rates    : ID#   rxn_name  opt SF_DEFAULT REG_PARAM ERROR
-------------------------------------------------------------------------------
%%% OBSERVATION MENU %%%
%%%  for PSUEDO_OBS  %%%
%%%   or LSENSE      %%%
Observation frequency OBS_FREQ : 60
Limit number of observations?  : F
  =>  Forcing time till        : 20050701 050000
COST FUNCTION options for LSENS:---
  =>  tracer kg/box            : T
  =>  tracer ug/m3             : F
  =>  tracer ppb               : F
  =>  tracer ppm free trop     : F
  =>  species ppb w/averaging  : F
  =>  tracer ug/m3 pop weight  : F
  =>  tracer ug/m2/hr          : F
  =>  deposition based?        : F
      =>  dry dep (not kpp)    : F
      =>  dry dep (kpp)        : F
      =>  tracer wet LS dep    : F
      =>  tracer wet CV dep    : F
      =>  molec/cm2/s          : F
      =>  kgN/ha/yr            : F
      =>  eq/ha/yr             : F
      =>  kg/s                 : F
  =>  Regional mask?           : F
      =>  binary punch file?   : F
         =>  bpch mask name    : usa_mask.geos.4x5
      OR  netcdf file      ?   : F
         =>  nc mask file name : /home/hyungmin/Class_1/Class1.nc
         =>  nc mask var  name : NPS_16
 >------------------------------<
NOBS: number of tracers to obs : 1
  => obs these tracers------>  : TRC# tracer_name
Tracer  #1                     : 2    O3
>------------------------------<
NOBS_CSPEC: # of species to obs: 0
  => obs these species------>  : species_name
Species #1                     : O3
-------------------------------------------------------------------------------
%%% FINITE DIFFERENCE MENU %%%
fd perturbation        FD_DIFF : 0.1
Numerator of derivative to test:---
 => longitude degree   LONFD   : 32
 => latitude  degree   LATFD   : 21
 => OR pick box by grid index? : T
```

```
   => longidute index  IFD     : 41
   => latitude index   JFD     : 32
 => altidude index     LFD     : 1
 => tracer (STT #)     NFD     : 2
Denomenator of deriv.   to test:
 => w/LEMS: emis group MFD     : 1
 => w/LEMS: sector     EMSFD   : 1
 => w/LICS: tracer     ICSFD   : 1
 => w/LSTR: tracer     STRFD   : 1
 => w/LRRATE: rate     RATFD   : 1
------------------------------------------------------------------------------
%%% DIAGNOSTICS MENU %%%
General                         : T
 => print debug  LPRINTFD       : F
 => jsave, jsave2               : F
 => adjoint traj LADJ_TRAJ      : F
   => w.r.t. scale factors?     : T
 => save iteration diags LITR   : T
 => sense w.r.t absolute emis   : F
CO satellite diganostics        : F
 => H(model)                    : F
 => h(obs)                      : F
 => H(model)-h(obs)             : F
 => adjoint forcing             : F
 => model bias                  : F
 => observation count           : F
 => DOFs                        : F
TES NH3 diagnostics             :---
 => BLVMR                       : F
HDF diagnostics                 :---
 => Level 2                     : F
 => Level 3                     : F
 ------------------------------------------------------------------------------
%%% CRITICAL LOAD MENU %%%
Critical Load obs               : F
 => N deposition                : T
 => Acidity deposition          : F
Critical Load file              : Exceedence.nc
GEOS-Chem file                  : Annual_Deposition.nc
------------------------------------------------------------------------------
END OF FILE                     :
--------------------------------+---------------------------------------------
```

Stratospheric production and loss

```
Strat prod & loss trc entries  : ID# trc_name    opt   SF_DEFALUT  REG_PARAM ERROR
Tracer  #1                     : 1    NOx_p      T     1           1         1
Tracer  #2                     : 2    Ox_p       T     1           1         1
Tracer  #3                     : 3    PAN_p      T     1           1         1
. . .
```

Reaction Rates

```
Rxn rate entries               : ID#   rxn_name                    opt   SF_DEFAULT   REG_PARAM   ERROR
Rate #1                        : 1     NO+O3                       T     1            1           1
Rate #2                        : 2     OH+O3                       T     1            1           1
Rate #3                        : 3     HO2+O3                      T     1            1           1
Rate #4                        : 4     NO2+O3                      T     1            1           1
Rate #5                        : 5     O3+MO2                      T     1            1           1
Rate #6                        : 6     2OH->O3+H2O                 T     1            1           1

. . .
```

define_adj.h

```
!*****************************************************************************
!  Include file "define_adj.h" specifies C-preprocessor "switches" that are
!  used to include or exclude certain sections of ADJOINT code, mostly for
!  controlling observation datasets used. The reason they are pre-processor
!  switches instead of logical flags is so that we can omit the code which
!  requires installation of hdf libaries and such. All are independent of
!  each other, but not of simulation and tracer type
!  (adj_group, 6/08/09)
!
!  List of "Switches"
!  ========================================================================
!  (1 ) TES_NH3_OBS         : Use NH3 data from TES
!  (2 ) PM_ATTAINMENT       : Compute PM attainment
!  (3 ) SOMO35_ATTAINMENT   : Compute ozone attainment
!  (4 ) SCIA_KNMI_NO2_OBS   : Use NO2 obs from SCIA KNMI retrieval
!  (5 ) IMPROVE_SO4_NIT_OBS : Use sulfate-nitrate from IMPROVE network
!  (6 ) CASTNET_NH4_OBS     : Use amonia from CASTNET network
!  (7 ) TES_O3_OBS          : Use O3 obs from TES
!  (8 ) SCIA_DAL_NO2_OBS    : Use NO2 obs from SCIA Dalhousie retrieval
!  (9 ) SCIA_DAL_SO2_OBS    : Use SO2 obs from SCIA Dalhousie retrieval
!  (10) MOPITT_V3_CO_OBS    : Use v3 CO obs from MOPITT
!  (11) MOPITT_V4_CO_OBS    : Use v4 CO obs from MOPITT
!  (12) SCIA_BRE_CO_OBS     : Use CO obs from SCIA Bremen retrieval
!  (13) AIRS_CO_OBS         : Use CO obs from AIRS (UMBC) retrieval
!  (14) PSEUDO_OBS          : Generate pseudo obs if no data selected
!  (15) LOG_OPT             : Optimized log of scaling factors
!  (16) SOMO35_ATTAINMENT   : Ozone attainment
!  (17) PM_ATTAINMENT       : PM    attainment
!  (18) LIDORT              : Online radiative forcing calculations
!  (19) GOSAT_CO2_OBS       : Use CO2 obs from GOSAT retrieval
!  (20) MODIS_AOD_OBS       : Use AOD obs from MODIS
!  (21) IMPROVE_BC_OC_OBS   : Use BC and OC aerosol obs from IMPROVE
!  (22) MOPITT_V5_CO_OBS    : Use v5 CO obs form MOPITT
!  (23) MOPITT_V6_CO_OBS    : Use v6 CO obs form MOPITT
!  (24) TES_O3_IRK          : Use radiative kernels for TES O3
!  (25) OMI_SO2_OBS         : Use OMI L3 SO2
!
! NOTES:
! (1 )  Replace MOPITT_IR_CO_OBS with MOPITT_V3_CO_OBS and MOPITT_V4_CO_OBS
!       (zhe, dkh, 02/04/11)
! (2 ) Add MODIS_AOD_OBS (xxu, dkh, 01/09/12, adj32_011)
! (3 ) Addd IMPROVE_BC_OC_OBS (yhmao, dkh/ 01/16/12, adj32_013)
! (4 ) Add MOPITT_V5_CO_OBS (zhej, dkh, 01/16/12, adj32_016)
! (5 ) Add CH4 obs operators (kjw, dkh, 02/12/12, adj32_023)
! (6 ) Add MOPIT_V6_CO_OBS and drop support for MOPITT v3 and v4 (zhe, dkh
06/2015)
!*****************************************************************************
!
!=============================================================================
! Undefine all "switches" so that they cannot be accidentally reset
!=============================================================================

#undef TES_NH3_OBS
#undef PM_ATTAINMENT
#undef SOMO35_ATTAINMENT
#undef SCIA_KNMI_NO2_OBS
#undef IMPROVE_SO4_NIT_OBS
#undef CASTNET_NH4_OBS
#undef TES_O3_OBS
#undef TES_O3_IRK
#undef SCIA_DAL_NO2_OBS
#undef SCIA_DAL_SO2_OBS
#undef SCIA_BRE_CO_OBS
#undef AIRS_CO_OBS
#undef GOSAT_CO2_OBS
#undef PM_ATTAINMENT
#undef SOMO35_ATTAINMNET
#undef PSEUDO_OBS
#undef LOG_OPT
#undef LIDORT
#undef LBKCOV_ERR
! (xxu, dkh, 01/09/12, adj32_011)
#undef MODIS_AOD_OBS
! (yhmao, dkh, 01/13/12, adj32_013)
#undef IMPROVE_BC_OC_OBS
! (zhej, dkh, 01/16/12, adj32_016)
#undef MOPITT_V5_CO_OBS
#undef MOPITT_V6_CO_OBS
! (kjw, dkh, 02/12/12, adj32_023)
#undef TES_CH4_OBS
#undef SCIA_CH4_OBS
#undef MEM_CH4_OBS
#undef LEO_CH4_OBS
#undef GEOCAPE_CH4_OBS
!mkeller
#undef OMI_NO2_OBS
! ( ywang, 04/21/15)
#undef OMI_SO2_OBS

!----------CO observations------
! pick any combination
! => MOPITT CO
!    => MOPITT v5
!    => MOPITT v6
! => AIRS CO
! => SCIA Bremen CO
!#define  MOPITT_V5_CO_OBS 'MOPITT_V5_CO_OBS'
!#define  MOPITT_V6_CO_OBS 'MOPITT_V6_CO_OBS'
!#define  AIRS_CO_OBS     'AIRS_CO_OBS'
!#define  SCIA_BRE_CO_OBS   'SCIA_BRE_CO_OBS'

!---------aerosol-related----------
!NH3 observations
! = > TES_NH3_OBS
!SO2 observations
! => SCIA_DAL_SO2_OBS
!Aerosol observations
! => PM_ATTAINMENT
! => IMPROVE_SO4_NIT_OBS
! => IMPROVE_BC_OC_OBS
! => CASTNET_NH4_OBS
!#define TES_NH3_OBS         'TES_NH3_OBS'
!#define SCIA_DAL_SO2_OBS    'SCIA_DAL_SO2_OBS'
!#define PM_ATTAINMENT       'PM_ATTAINMENT'
!#define IMPROVE_SO4_NIT_OBS 'IMPROVE_SO4_NIT_OBS'
!#define IMPROVE_BC_OC_OBS   'IMPROVE_BC_OC_OBS'
!#define CASTNET_NH4_OBS     'CASTNET_NH4_OBS'
!#define MODIS_AOD_OBS       'MODIS_AOD_OBS'

!--------ozone-related-------------
! => SOMO35_ATTAINMENT
! => TES O3
! => TES O3 IRKs
!#define SOMO35_ATTAINMENT 'SOMO35_ATTAINMENT'
!#define TES_O3_OBS    'TES_O3_OBS'
!#define TES_O3_IRK    'TES_O3_IRK'

!--------CH4 Observations------------
! => TES CH4
! => SCIA CH4
! => MEM CH4
! => Generic LEO instrument CH4
! => GEOCAPE CH4
!#define TES_CH4_OBS         'TES_CH4_OBS'
!#define SCIA_CH4_OBS        'SCIA_CH4_OBS'
!#define MEM_CH4_OBS         'MEM_CH4_OBS'
!#define LEO_CH4_OBS         'LEO_CH4_OBS'
!#define GEOCAPE_CH4_OBS     'GEOCAPE_CH4_OBS'

!-------NO2 observations------------
! => SCIA_KNMI_NO2_OBS
! => SCIA_DAL_NO2_OBS
!#define SCIA_KNMI_NO2_OBS   'SCIA_KNMI_NO2_OBS'
!#define SCIA_DAL_NO2_OBS    'SCIA_DAL_NO2_OBS'

!-------OMI NO2 tropospheric columns
!#define OMI_NO2_OBS 'OMI_NO2_OBS'

!-------CO2 observations------------
! => GOSAT_CO2_OBS
!#define GOSAT_CO2_OBS

!-------SO2 observations------------
! => OMI_SO2_OBS
!#define OMI_SO2_OBS         'OMI_SO2_OBS'

!------other options-----------------
!#define PSEUDO_OBS    'PSEUDO_OBS'
!#define LOG_OPT       'LOG_OPT'
!#define LIDORT        'LIDORT'
!#define LBFGS_INV     'LBFGS_INV'
!#define LBKCOV_ERR    'LBKCOV_ERR'
```

5. Examples

a. Inverse Modeling

Test if we can correct an error in anthro NO_x emissions using "observations" of O_x

- First generate a set of "observations" of O_x using base case (scaling factors = 1) NO_x emissions.

- Inversion begins with initial guess of NO_x emission = 1.5 * base NO_x emissions

- All other emissions assumed to be correct (so only optimize anthro NO_x emissions)

- Inversion uses Ox observations in every grid cell, at every hour, over the course of 1 day (so it better work!).

- Use the adjoint model to complete 10 iterations -- check to see if NO_x emission scaling factors converge to 1.0

- trop chemistry, geos5, 4x5, v35

b. Sensitivity Test

Sensitivity of Ox with respect to NOx using initial conditions. See how much Ox changes with changes in NOx emissions.

- Sensitivity of Ox in level 1 at hour 6 with respect to NOx initial conditions in level 1

- trop chemistry geos5, 4x5, v35

C. Finite Difference

Compare adjoint sensitivities to sensitivities calculated using finite differences in each column of the model (horizontal transport off)

- Sensitivity of Ox in level 1 at hour 6 with respect to NOx initial conditions in level 1

- trop chem, geos5, 4x5, v35

This is a standard test we do for every model version to benchmark code, see the google spreadsheet
https://docs.google.com/spreadsheet/ccc?key=0As2MewHKyWpDdDlTeFdPanU2RzlWc0luaFpEQXlyclE#gid=0

6. Useful GIT commands:

Initial download:

```
git clone git@adjoint.colorado.edu:yanko.davila/gcadj_std.git
```

Status of project vs the current repository:

```
git status
```

Check difference of files (differences have colors for easy reading)

```
git diff --color <wildcard> [<wildcard>] <path>/foo_mod.f
```

Checkout specific version

```
git checkout <wildcard>
```

Replacing a file with the newest version from the repository:

```
git checkout origin/master -- <path>/foo_mod.f
```

Merging changes in a file:

```
git merge -m <wildcard>
```

Comitting

```
git commit -a
```

Tagging a version

```
git tag -a TAGNAME
```

Deleting a tag

```
git tag -d TAGNAME
git push origin :refs/tags/TAGNAME
```

List the history of a file:

```
git log -- <path>/foo_mod.f
```

Add a file to the repository

```
git add <file_name>
```

Delete a file from the repository

```
git rm <file_name>
```

There are several wildcards that you can use on git for example:
"origin/master" - Latest version on the repository
"HEAD" - Latest version as of your last download
"v33i" - Specific TAG, find all tag names on GitLab or by typing " git tag "
"32d5c926e" - Specific COMMIT, find all commit numbers on GitLab

7. GCHP adjoint

Installation and run guide
Last updated: 06/03/2022, see most current at https://tinyurl.com/3ayyy8ez

Note: these instructions are specific to the NAS Pleiades system, where it is assumed users have directories on the following filesystems: /swbuild, /nobackup . The make step proceeds much faster on /swbuild. It is also expected that users have familiarity with submitting batch jobs to the queue ono this system. For more information about running simulations on Pleiades, see https://www.nas.nasa.gov/hecc/support/kb/running-jobs-with-pbs-121

1. Obtaining the code

First, change directory to the location where you would like to download the source code. On Pleiades, it is useful do request access to the /swbuild directory, where the code will build and compile much more quickly.

```
cd /swbuild/[username]
```

There are two ways to download the code. If you plan to work on code development, then you should first have a github account with your ssh keys setup (more on that at https://docs.github.com/en/authentication/connecting-to-github-with-ssh). Then use the following:

```
git clone git@github.com:TerribleNews/GCHPctm_adj.git ./GCHP-adj
```

Alternatively, if you just plan to use the code, you can use the following:

```
git clone https://github.com/TerribleNews/GCHPctm_adj.git ./GCHP-adj
```

Next enter the code directory and checkout a dev branch:

```
cd GCHP-adj/
git checkout dev
git submodule update --init –recursive
```

2. Setting up a run directory

The next step is to create a run directory using the createRunDir.sh scripts in `GCHP-adj`/run. Here we will assume that you plan to use a directory called /nobackup/[user]/GCHPtest and use gchp_c24_adj as the run name, but you can replace those with something else if you'd prefer. Replace [user] with your username everywhere below.

```
mkdir /nobackup/[user]/GCHPtest
cd run
./createRunDir.sh gchp_c24_adj

============================================================
GCHP RUN DIRECTORY CREATION
============================================================

------------------------------------------------------------
Choose simulation type:
------------------------------------------------------------
   1. Full chemistry
```

```
   2. TransportTracers
   3. CO2 w/ CMS-Flux emissions
> 3

------------------------------------------------------------
Choose meteorology source:
------------------------------------------------------------
  1. MERRA-2 (Recommended)
  2. GEOS-FP
> 1

------------------------------------------------------------
Enter path where the run directory will be created:
------------------------------------------------------------
> /nobackup/[user]/GCHPtest

  -- This run directory has been set up for 20140901 - 20141001.

  -- The default frequency and duration of diagnostics is set to
monthly.

  -- You may modify these settings in runConfig.sh.

------------------------------------------------------------
Do you want to track run directory changes with git? (y/n)
------------------------------------------------------------
> n

Created /nobackup/[user]/GCHPtest/gchp_c24_adj
```

3. Build, compile, and install the code

The following instructions assume you are working in a bash shell. It should be fairly easy to translate these scripts to other shells. If you're not using bash by default, then first launch a bash shell

```
bash
```

First, test to make sure your bash shell has access to the module command:

```
which module
```

If the output contains "which: no module in …" then you will need to issue the following command:

```
source /usr/share/modules/init/bash
```

Before you can compile, you will need to set your environment on Pleaides. In your /swbuild/[user]/GCHP-adj directory, create a file called env.rc with the following contents, making sure to define ESMF_ROOT environment variable as typically done for your account when running GCHP, and to define the path to cmake:

```
module purge

# load ifort, sgi mpi, and netcdf
module load comp-intel/2020.4.304
module load mpi-hpe/mpt.2.25
module load hdf4/4.2.12 # required by netcdf
```

```
module load hdf5/1.8.18_mpt # also required by netcdf?
module load netcdf/4.4.1.1_mpt
module load python3 # maybe required for f2py?
# Define ESMF_ROOT with correct path for your account, for example
#export ESMF_ROOT=/nobackupp12/clee59/software/
Export ESMF_ROOT=???
export OMP_STACKSIZE=500m
# define path to cmake, for example
export PATH=/nobackupp12/clee59/software/cmake-3.16.3-Linux-x86_64/bin/:$PATH
ulimit -s unlimited

export FC=ifort
export CC=icc
export CXX=icpc
```

Next, load these environment variables with

```
source env.rc
```

In order the run the adjoint model, we must first run the forward model. In the case of GCHP, we have separated this into two separate executions, but only a single executable. First, the adjoint-enabled binary must be run in forward mode, then (usually) in the same directory, the same binary is run in adjoint mode. Building the adjoint-enabled binary is just like building gchp normally, but you will supply the flags -DADJOINT=yes and -DREVERSE_OPERATORS=yes to cmake. Build your binary by creating a build directory, sourcing the env.rc file created in the step above and then running make -j 4 and make install. Replace [user] with your username.

```
cd /swbuild/[user]/GCHP-adj
mkdir build
cd build
CMAKE_PREFIX_PATH=/nasa/hpe/mpt/2.25_rhel79/lib cmake ../ \
-DRUNDIR=/nobackup/[user]/GCHPtest/gchp_c24_adj -DADJOINT=yes \
-DREVERSE_OPERATORS=yes
make -j 4
make install
```

4. Running the code: default forward and adjoint simulation

Copy the run scripts to your run directory

```
cd /nobackup/[user]/GCHPtest/gchp_c24_adj
cp runScriptSamples/gchp.{adjoint,pleiades}.run ./
```

These are the run scrips for running the forward model and adjoint model, respectively, on NAS Pleiades. You will need to update the job commands at the top of the file, such as the group_list, your email address, etc.

You will also need to give an environment setup file to the job scripts. Possibly many of these exports can be eliminated but this has yet to be confirmed experimentally (let us know though if you have time!). Create a file in your home directory called gchp.ifort20_TOSS3_pleiades.env with the following contents:

```
#!/bin/bash
# Based on seastham's home/pleiades.basrch
```

```
# Builds NetCDF on Pleiades using the standard compiler set, ready for
GCHP

# load ifort, sgi mpi, and netcdf
module load comp-intel/2020.4.304
module load mpi-hpe/mpt.2.25
module load hdf4/4.2.12 # required by netcdf
module load hdf5/1.8.18_mpt # also required by netcdf?
module load netcdf/4.4.1.1_mpt


export ESMF_COMM=mpi
export ESMF_COMPILER=intel

# Tell GCHP where the MPI binaries, libraries and so on can be found
export MPI_ROOT=$( dirname $( dirname $( which mpiexec ) ) )

# Set up the compilers
export FC=ifort
export F90=$FC
export F9X=$FC
export F77=$FC
export CC=gcc
export CXX=g++

export OMP_STACKSIZE=500m
ulimit -s unlimited

# Needed one NetCDF is installed
export NETCDF_HOME=$(nc-config --prefix)

export GC_BIN="$NETCDF_HOME/bin"
export GC_INCLUDE="$NETCDF_HOME/include"
export GC_LIB="$NETCDF_HOME/lib"

# If using NetCDF after the C/Fortran split (4.3+), then you will need
to
# specify the following additional environment variables
export NETCDF_FORTRAN_HOME=$(nf-config --prefix)
export GC_F_BIN="$NETCDF_FORTRAN_HOME/bin"
export GC_F_INCLUDE="$NETCDF_FORTRAN_HOME/include"
export GC_F_LIB="$NETCDF_FORTRAN_HOME/lib"

export PATH=${NETCDF_FORTRAN_HOME}/bin:$PATH
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:${NETCDF_FORTRAN_HOME}/lib

# Add NetCDF to path
export PATH=$PATH:${NETCDF_HOME}/bin
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:${NETCDF_HOME}/lib

# Disable OpenMP
export OMP_NUM_THREADS=1

# Attempt an alias for mpifort if not found
#command -v mpifort >/dev/null 2>&1 || alias mpifort=mpif90
```

```
export PATH=${PATH}:${HOME}/mpi_extra

# Set path to GMAO Fortran template library (gFTL)
export gFTL=$(readlink -f ./gFTL)

# Set ESMF optimization (g=debugging, O=optimized (capital o))
export ESMF_BOPT=O
```

Tell GCHP where your environment variables live:

```
cd /nobackup/dhenze/GCHPtest/gchp_c24_adj
./setEnvironment.sh ~/gchp.ifort20_TOSS3_pleiades.env
```

Now run the forward model with the gchp.pleiades.run script:

```
qsub gchp.pleiades.run
```

After that's complete you can archive the run results as

./archiveRun.sh ../fwd _default

This causes all simulation settings, output and logs to be stored under a directory called fwd_default in the parent directory from the run directory.

Next, manually move a checkpoint file to be a new restart file for the adjoint:

```
mv gcchem_internal_checkpoint
gcchem_internal_restart.20141001_060000z.nc4
```

You may also want to remove the forward run's gchp.log from the run directory (you have a copy saved in fwd_default if needed)

rm gchp.log

Then, in the same directory, start the adjoint run:

```
qsub gchp.adjoint.run
```

After that's complete you can archive the run results as

./archiveRun.sh ../adj _default

The results of the adjoint are in

OutputDir/GEOSChem.Adjoint.20140901.nc4

OutputDir/GEOSChem.SFEmissions.20140901.nc4

These files contain the initial condition scaling factor sensitivities and emissions scaling factor sensitivities. Emissions sensitivities are included for each hour and sector.

5. Running the code: global FD test
We recommend starting with step 4 (default run) before conducting the FD test. This will ensure things work prior to changing any settings. Also, the instructions within step 4 for setting up p

To conduct a global finite difference test, you will have to complete 4 model runs. The instructions here assume you are doing all of these from the same directory, in the order described below. Alternatively (not shown), it's possible to set them up in separate run directories; one would need to copy checkpoint and restart files appropriately.

As the current CO2 adjoint only is developed for transport, you will also need to go through both runConfig.sh and runConfig_adj.sh and turn off chemistry, convection, and deposition:

```
# geoschem_config.yml
Turn_on_Chemistry=false
Turn_on_Transport=true
Turn_on_Cloud_Conv=false
Turn_on_PBL_Mixing=false
Turn_on_Non_Local_Mixing=false
Turn_on_Dry_Deposition=false
Turn_on_Wet_Deposition=false
```

5.1 Unperturbed forward run:

Adjust the settings in the Adjoint Variables menu of the GCHP.rc in your run directory to match column 1 of Table 1.

Next submit the job:

```
qsub gchp.pleaides.run.
```

When the run completes:

```
./archiveRun.sh ../fd_glob_0
```

This last step causes all data to be stored under a directory called fd_glob_0 in the parent directory from the run directory.

5.2 Adjoint run:

Adjust the settings in the Adjoint Variables menu of the GCHP.rc in your run directory to match column 2 of Table 1. Note: the batch script will execute runConfig_adj.sh, and

the latter automatically sets the MODEL_PHASE to ADJOINT. Changing it manually yourself before executing will not have an impact.

Next submit the job:

qsub gchp.adjoint.run.

When the run completes:

./archiveRun.sh ../fd_glob_adj

5.3 Positive perturbed forward runs:

Adjust the settings in the Adjoint Variables menu of the GCHP.rc in your run directory to match column 3 of Table 1. Next submit the job:

qsub gchp. pleaides.run.

When the run completes:

./archiveRun.sh ../fd_glob_1

5.3 Negative perturbed forward runs:
Lastly, adjust the settings in the Adjoint Variables menu of the GCHP.rc in your run directory to match column 4 of Table 1. Next submit the job:

qsub gchp. pleaides.run.

When the run completes:

./archiveRun.sh ../fd_glob_2

Table 1: Settings in GCHP.rc for GLOBAL FD test

MODEL_PHASE: FORWARD	MODEL_PHASE: ADJOINT	MODEL_PHASE: FORWARD	MODEL_PHASE: FORWARD
FD_TYPE: GLOBAL	FD_TYPE: GLOBAL	FD_TYPE: GLOBAL	FD_TYPE: GLOBAL
FD_STEP: 0	FD_STEP: 0	FD_STEP: 1	FD_STEP: 2
FD_SPEC: CO2	FD_SPEC: CO2	FD_SPEC: CO2	FD_SPEC: CO2
# #FD_LAT: 44.65	# #FD_LAT: 44.65	# #FD_LAT: 44.65	# #FD_LAT: 44.65
# #FD_LON: -63.58	# #FD_LON: -63.58	# #FD_LON: -63.58	# #FD_LON: -63.58
IFD: 3	IFD: 3	IFD: 3	IFD: 3
JFD: 56	JFD: 56	JFD: 56	JFD: 56
LFD: 1	LFD: 1	LFD: 1	LFD: 1