EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Institute for Environment and Sustainability

CCU

Joint Research Centre

I. Update on CH4 Inverse Modeling at JRC

Arjo Segers
Peter Bergamaschi

European Commission Joint Research Centre, Institute for Environment and Sustainability, Climate Change Unit, I-21027 Ispra (VA), Italy

CH4 inversions - upgrade

upgrade of the 4D-var CH4 system

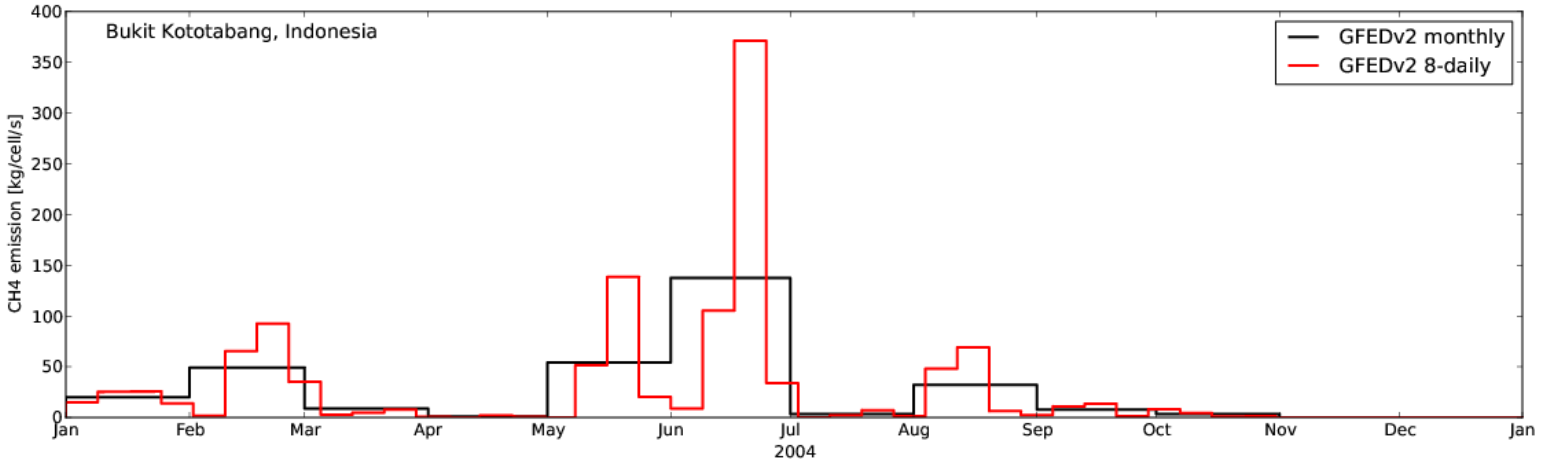
- new scripting ('pycasso')
  - easier running same code on different platforms
  - decreases number of scripts (input.sc, output.sc)
  - less (!) settings visible for users

- brought back to live : TIPP
  - TM Input Pre-Processor (Jan Fokke Meirink)
  - converts raw emissions into single format
  - extended with various Edgar 4.0 formats
  - support GFED 8-daily biomass burning emissions

CH4 inversions - 8-daily fire emissions

extended 4D-var system

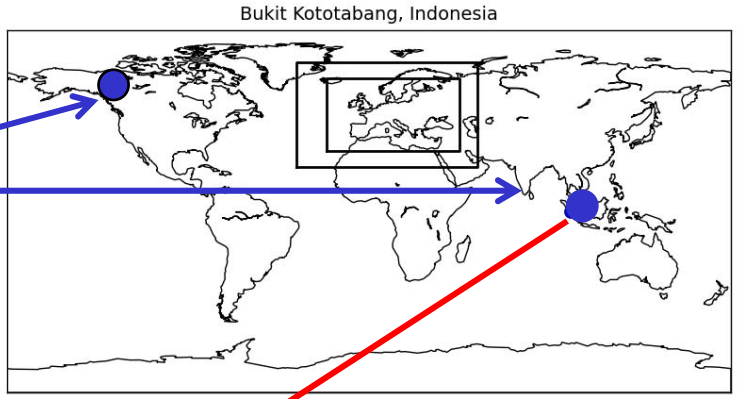
- **Default optimization on monthly averages**
- **Added 8-daily profile on top of monthly average**

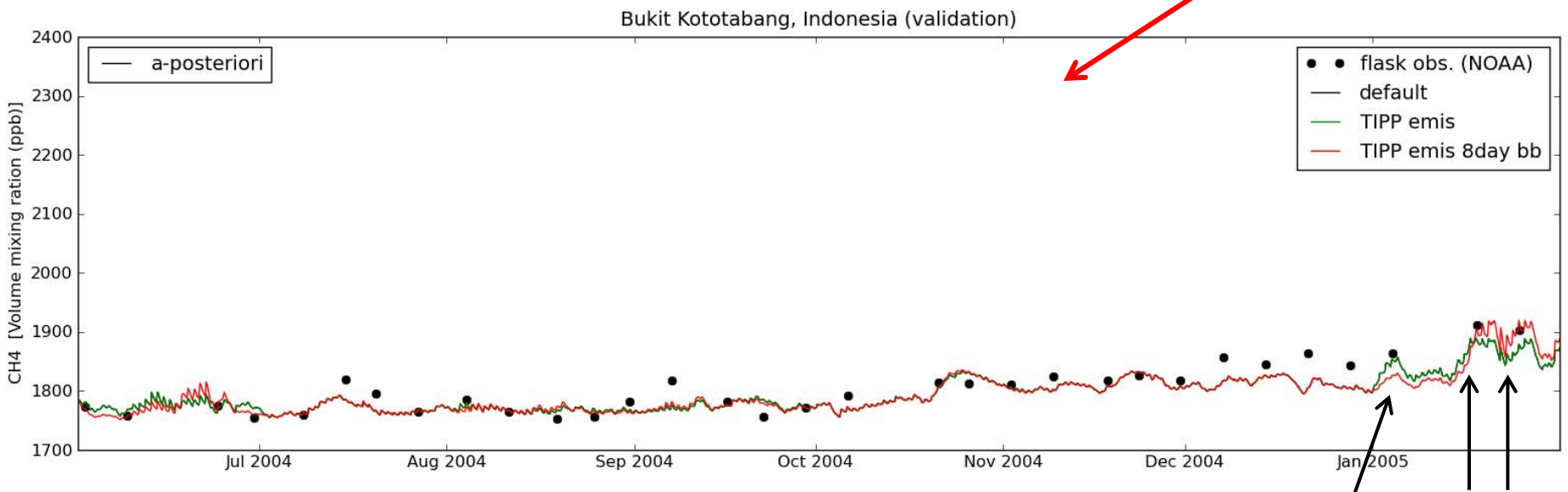


- **results after inversion should be not too different!**

CH4 inversions - 8-daily fire emissions

- **Tested for year 2004**
- **Difference 'monthly' vs 'monthly+8daily' visible at 2 NOAA stations only**





- **optimization of monthly averages only !**
- **** coming soon ** 4D-var with time scales other than monthly**

II. Update on meteo format

Arjo Segers

European Commission Joint Research Centre, Institute for Environment and Sustainability, Climate Change Unit, I-21027 Ispra (VA), Italy

TM5 meteo files

- **from previous TM meeting:**

```
-----------------------------------------  ----------- ---------
 timer                                      system_clock       (%)
 -----------------------------------------  ----------- ---------

 root                                           2566.69
   field2d                                        37.36 (  1.5 %)
   field3d                                      1392.87 ( 54.3 %)
   other                                        1136.46 ( 44.3 %)

 field3d                                        1392.87
   field3d read                                 1322.58 ( 95.0 %)
   field3d transform hori                         66.50 (  4.8 %)
   field3d transform vert                          3.57 (  0.3 %)
   other                                           0.22 (  0.0 %)
-----------------------------------------  ----------- ---------
```

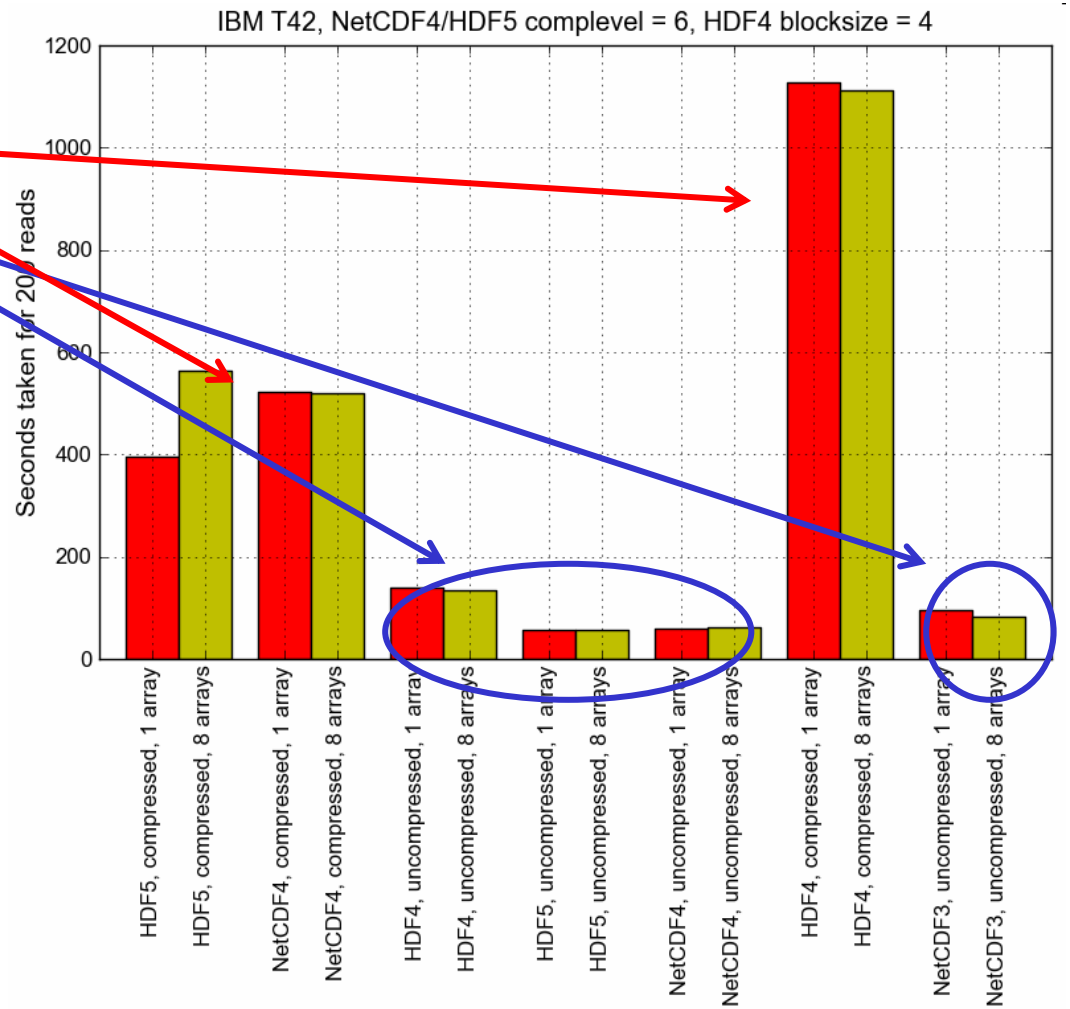
- **In some applications, reading meteo files took about 50% of total time!**

TM5 meteo files

- **Experiments by Sourish:**
  - **compressed**
  - **uncompressed**

- **reading files is slow due to:**
  - **hdf**
  - **compression**

- **is this a problem ?**
  depends on:
  - **application**
    - **1 tracer 4D-var**
    - **full chemistry**
  - **file system (access time)**
  - **installed hdf libraries ?**



TM5 meteo files

- current format:
  - hdf
  - internal compression
    - saves 25% of disk space
    - useful when we only used workstations with small disks
  - multiple 3D records in a file
  - 8-10 years old ?

- testing a new format:
  - NetCDF4
    - = based on HDF5
    - ! bug in HDF5 for IBM AIX machines (ecmwf ...); therefore tests with 'classic' NetCDF
  - 4D records
  - uncompressed

TM5 meteo files

- test:
  - TM5 from trunk (version `last week`)
  - L60, glb6x4 / eur3x2 / eur1x1
  - test ‘chemistry’ (5 transported tracers)
  - 12 hour simulation
  - ecmwf/c1a, single cpu

	HDF		NC	
timer	system_clock	(%)	system_clock	(%)
root	102.42		87.46	
init	2.18	(2.1 %)	2.21	(2.5 %)
step start	13.32	(13.0 %)	9.44	(10.8 %)
step init	41.27	(40.3 %)	30.20	(34.5 %)
step run	45.12	(44.1 %)	45.12	(51.6 %)
other	0.53	(0.5 %)	0.49	(0.6 %)

→ about 12% reduction of time

TM5 meteo files

- **test:**
  - **v3.0 chemistry benchmark**
  - **ecmwf/c1a, 4 mpi tasks, 2 openmp threads**

	HDF		NC	
timer	system_clock	(%)	system_clock	(%)
root	2276.34		2018.85	
init	2.86	(0.1 %)	2.76	(0.1 %)
done	1.30	(0.1 %)	1.23	(0.1 %)
step start	25.19	(1.1 %)	18.98	(0.9 %)
step init	918.70	(40.4 %)	682.77	(33.8 %)
step run	1325.78	(58.2 %)	1310.80	(64.9 %)
step done	0.02	(0.0 %)		
other	2.49	(0.1 %)	2.31	(0.1 %)

→ **about 12% reduction of time**

TM5 meteo files

- test:
  - 4D-var CH4
  - JRC cluster (fast local disk, one user)
  - not tested with NC files yet

	HDF		NC	
timer	system_clock	(%)	system_clock	(%)
tracer model	208.87			
model init	5.00	(2.4 %)		
timestep init	32.17	(15.4 %)		
timestep run	169.96	(81.4 %)		
timestep done	1.70	(0.8 %)		
other	0.05	(0.0 %)		

→ reduction of time max. 15%

TM5 meteo files - proposal

Proposal for coming months:

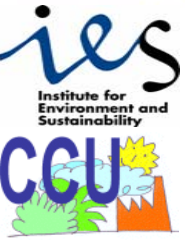
- **wait for upgrade of HDF5/NetCDF4 libs at ECMWF**

- **large scale test of run times:**
  - **different applications**
  - **different machines**

  could volunteers be identified by meeting ?

  (support for coding and configuration is available...)

- **next meeting:**
  - **decide if improvement is worthwhile the investment**
  - **if so, decide on switch/conversion to new format**



III. Scripting

Arjo Segers

European Commission Joint Research Centre, Institute for Environment and Sustainability, Climate Change Unit, I-21027 Ispra (VA), Italy

pycasso scripting

Evolution of the TM5 scripting ...

- 'run-tm5'
  - 2002-2010
  - bourne shell (/bin/sh)
  - to build source code from base/proj directories
  - many settings hard coded in many files:
    - run-tm5, run-tm5.jb, config.compiler, configure, ...

- 'runtm5.py'
  - 2009-...
  - python
  - translation of 'run-tm5'
  - many configurations now controlled by rcfile settings

- 'pycasso' branch
  - based on 'runtm5.py'
  - *all* settings now controlled by rcfile settings
  - get rid of runtime files 'times.in', 'runtime.rc', etc

pycasso - setup/submit scripts

Split between 'building' and 'running' an exectable

- **Before: one script to do all:**
  `./runtm5.py  tm5.rc`
- **Now first build:**
  `./setup_tm5  tm5.rc`
- **and then from run directory:**
  /scratch/run> ./submit_tm5 tm5.rc
- **... but for the impatient those could be combined again:**
  ./setup_tm5 tm5.rc --submit --queue
- **Advantage: to reproduce a run it is sufficient to store:**
  - **executable**
  - **rcfile**
  - **submit script**

  without the need to re-build the executable again

pycasso - build

- building a code : pycasso scripts
  - 'python compile and setup script organizer‘
  - set of scripts to compile a code and setup a run directory without any reference to TM5
  - ... but with all tasks necessary for TM5 included

- All configuration done through rcfile settings:
  - code directories (base,proj,...)
  - which files to copy and which not (source.rc)
  - compiler flags
  - libraries

- special TM5 configurations in module 'user_scripts_tm5' (dims_grid.h, TracerOrder, ...)

pycasso - jobchain

Revised job chain

- job scripts are written by submit script

- as small as possible!
  - queue options
    - specified in rcfile! facilitates many platforms
  - run a single script or executable
  - submit a new job if necessary

... thus no long 'run-tm5.jb' template script anymore ...

- each job in the chain has its own rcfile:

tm5_001.jb tm5_001.rc
tm5_002.jb tm5_002.rc
...

with different values inserted for time range, istart, ...

→ easier to re-run a job

pycasso - MDF

- (extra) MDF module
  - "multiple data formats"
  - single interface to both HDF and NetCDF(4)
  - interface similar to NetCDF F90 interface:

```
call MDF_Create( 'test.hdf', MDF_HDF    , MDF_NEW,  hid, status )
call MDF_Create( 'test.nc‘ , MDF_NETCDF4, MDF_NEW, ncid, status )
```

  - could replace 'file_hdf'
  - single (large) source file
  - created using python scripts
  - available from 'tools' directory

pycasso - macros

- (extra) Macro definitions not via compiler flags:

```
f90 -c tm5.F90 -DMPI -Dwith_zoom -Dwithout_anything ...
```

- ... but defined in include file "tm5.h", written by setup script:

```
! this is tm5.h
#define MPI
#define with_zoom
#define without_anything
```

- Include this file in the top of every TM5 source file:

```
#include "tm5.h“
```

- Advantage:
  - macro definitions now part of source, thus shipped together: always clear which macros where defined to compile a code
  - if macro definitions change, code is re-compiled automatically thanks to makedep.f90

pycasso – grid definition

- (extra) grid definition in rcfile, including reduced grid:

```
     ! global 6x4
region.glb600x400.xcyc                :     1
region.glb600x400.touch_np            :     1
region.glb600x400.touch_sp            :     1
region.glb600x400.xbeg                :  -180
region.glb600x400.xend                :   180
region.glb600x400.ybeg                :   -90
region.glb600x400.yend                :    90
region.glb600x400.im                  :    60
region.glb600x400.jm                  :    45
region.glb600x400.redgrid.nh.n        :     4
region.glb600x400.redgrid.nh.comb     :    60 20 10 5
region.glb600x400.redgrid.sh.n        :     4
region.glb600x400.redgrid.sh.comb     :    60 20 10 5
```

- specify al list of regions to be used:

```
    regions     :  glb600x400  eur300x200  eur100x100
```

- no need for ‘proj/grid’ and ‘RedGrid’ files anymore

pycasso - timers

- (extra) timing routines

  – **added lines to the code:**

```
integer   ::  itim
call GO_Timer_Def( itim, 'main' )
...
call GO_Timer_Start( itim )
... ! this is the main task
call GO_Timer_Done( itim)
```

- at end of the run a profile is written:

```
-------------------------  ------------ ---------
timer                      system_clock        (%)
-------------------------  ------------ ---------
tracer model                     208.87
  model init                       5.00 (  2.4 %)
  timestep init                   32.17 ( 15.4 %)
  timestep run                   169.96 ( 81.4 %)
  timestep done                    1.70 (  0.8 %)
  other                            0.05 (  0.0 %)
.........
```

pycasso - jobchain

- **Where used ?**
  - **JRC 4D-var CH4 code**
  - **TM5/base/branches/pycasso**

- **Recently tested:**
  - **trunk source with pycasso scripts : ok**
  - **pycasso source with trunk scripts : ok**
  - **pycasso source with pycasso scripts : ok**

- **If no objections: merged into the trunk soon ...**