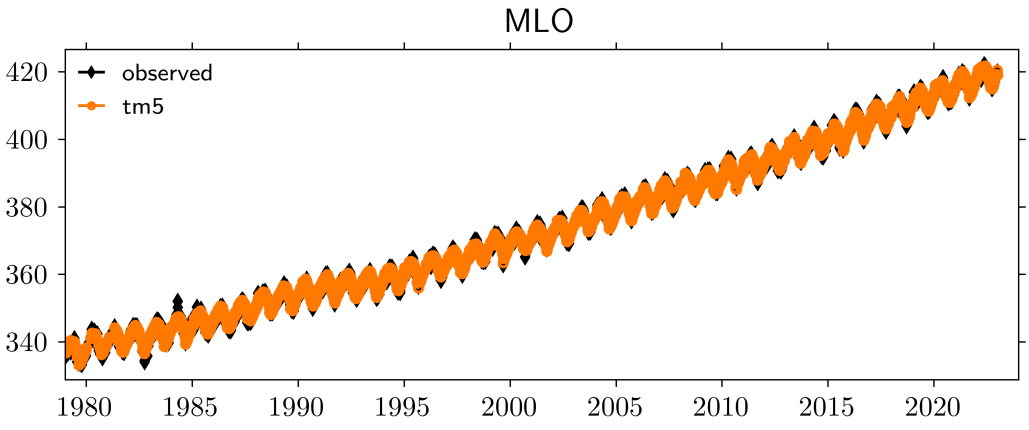
Development of a long-window/short window data assimilation system

Joram Hooghiem, John B. Miller, Aleya Kaushik, Auke van der Woude, Anne-Wil van den Berg, Marnix van de Sande, Remco de Kok, Ingrid Luijkx, Wouter Peters

October 21, 2024

45 year inversion of CO_2 transport $3° \times 2° \times 34$



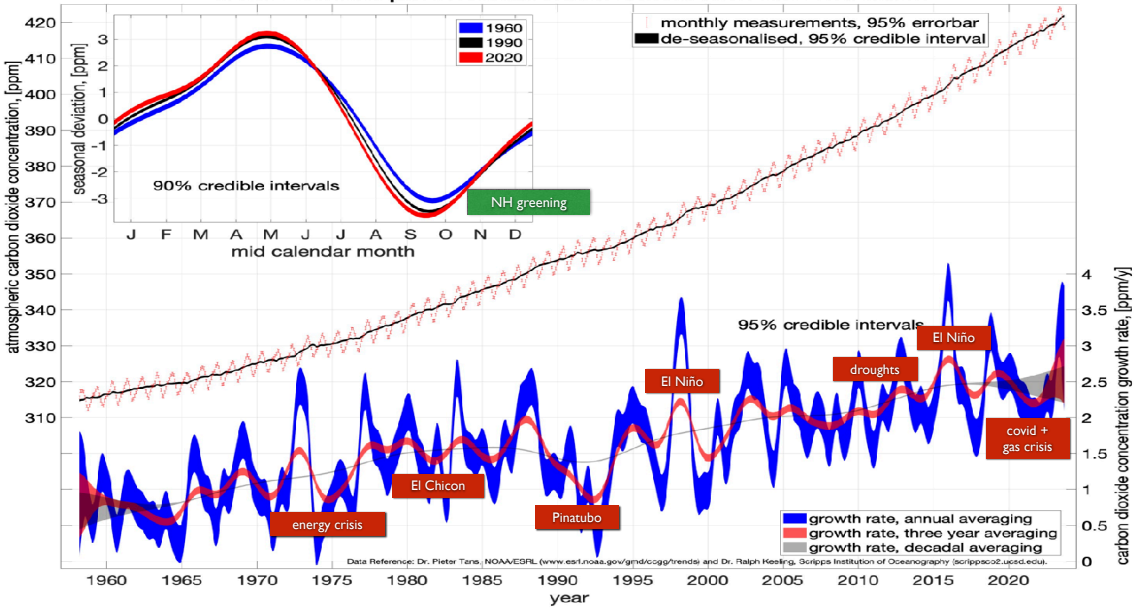
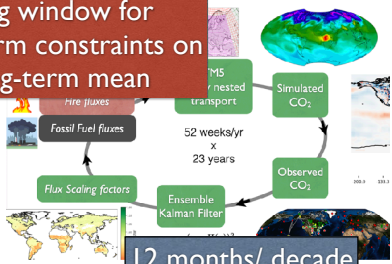


Figure courtesy of Prof C.E. Rasmussen, from https://mlg.eng.cam.ac.uk/carl/words/carbon.html

CarbonTracker Europe long-window short-window (LWSW)

long window for long-term constraints on long-term mean

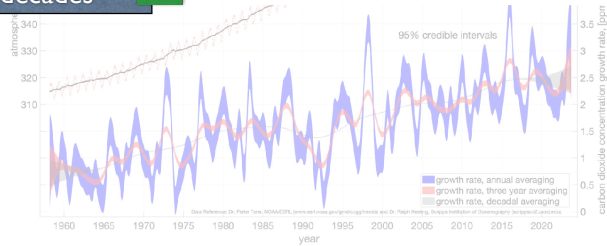


+

12 months/ decade
x
4 decades

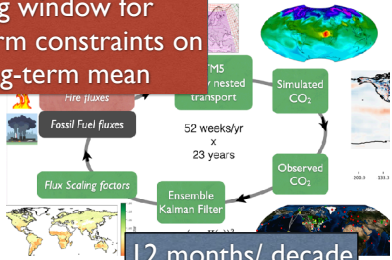
λ1

NEW



CarbonTracker Europe long-window short-window (LWSW)

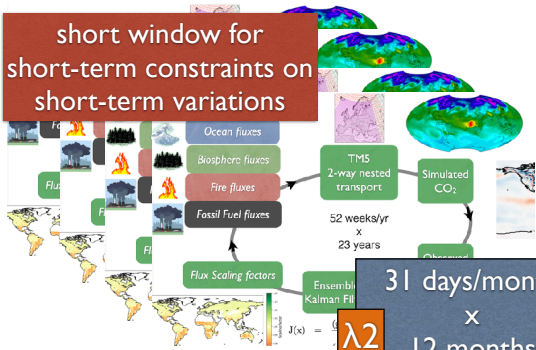
long window for long-term constraints on long-term mean



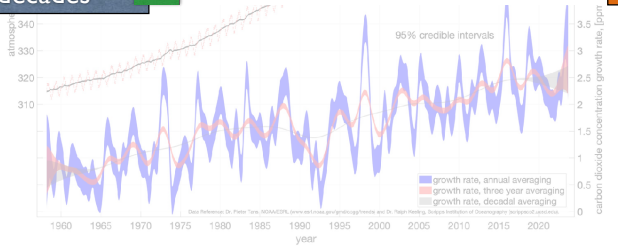
12 months/ decade
x
4 decades
λ1

+

short window for short-term constraints on short-term variations



31 days/month
x
12 months
λ2



NEW

Long window inversion system

- Single window
- Multi-decadal estimates
- Cheap to run
- no adjoint
- no propagation

Long-Window algorithm

```
Setup state
Create ensemble
Transport ensemble (TM5)
Ensemble Kalman Filter
Exit
```

- Global $3^\circ x 2^\circ x 34$
- 45 years CO_2-only
- Wall-clock time: 1 week

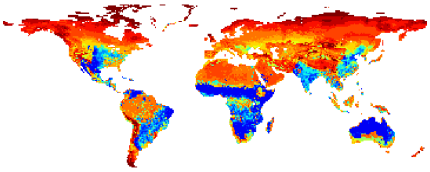
LW constraint: atmospheric data

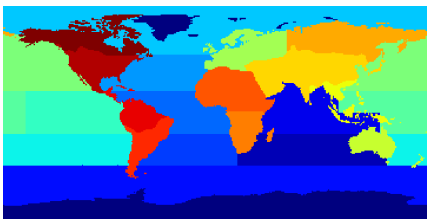
Setting up the sate

- Scaling factors λ (ocean,GPP ,TER-clim)
- $\boldsymbol{P}_{\mathrm{s}}$ Spatially like CT(E): 30 ocean basins / 134 ecoregions
- $\boldsymbol{P}_{\mathrm{t}}$ New temporal covariance
  1. daily

$$\boldsymbol{P} = \boldsymbol{P}_{\mathrm{s}} \otimes \boldsymbol{P}_{\mathrm{t}}$$

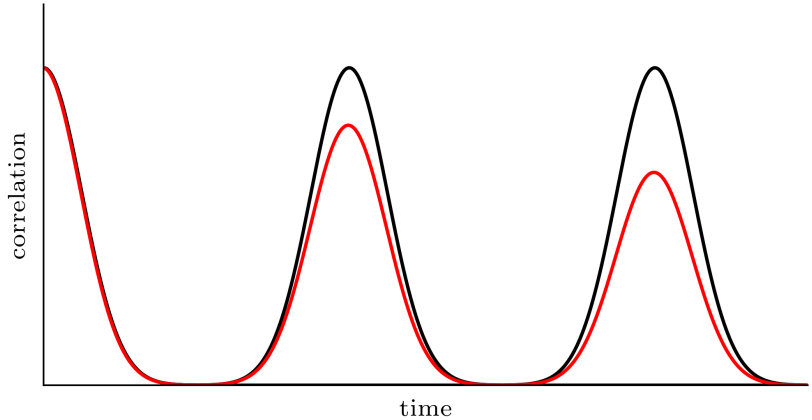
$N \approx 5 \times 10^6$ (Back of the envelope ddof 5×10^5
150 ensemble members



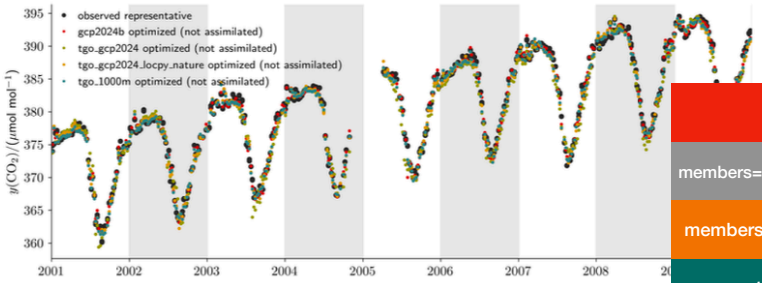


Cyclic covariance with trend

year-to-year correlation is dampened



Is 150 members enough?



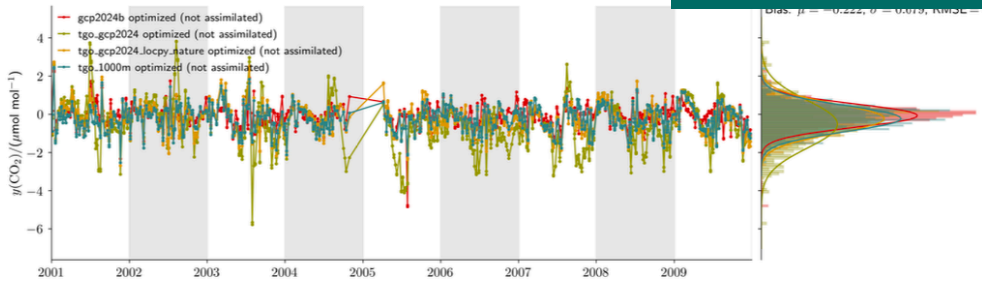
ALT, Alert, Nunavut
82°27′N, 62°30′W, 185 masl
surface-flask
NOAA Global Monitoring Laboratory, United States

GCB2024b (μ = -0.068, RMSE = 0.609)

members=150, non-localized (μ = -0.561, RMSE = 1.347)

members=150, localized (μ = -0.221, RMSE = 0.802)

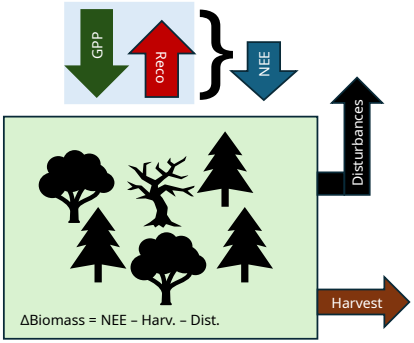
1000 members (μ = -0.222, RMSE = 0.714)

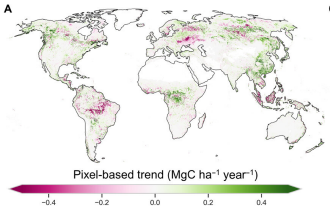


LW constraint: Δ-biomass
(Auke van der Woude & Marnix van de Sande)

Use long-term constraint: Δ-biomass

A simplified view of forest carbon fluxes



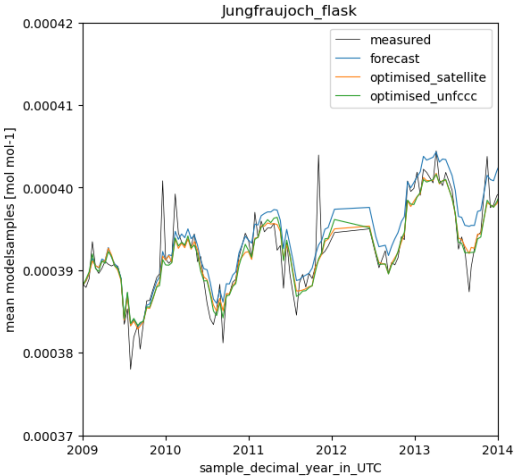


(Xu et al. 2021)

UNFCCC

(adapted from Van der Woude et al. 2024)

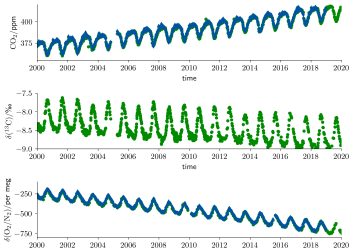
Proof of concept: Reasonable assimilated mole fractions after optimisation with biomass datasets

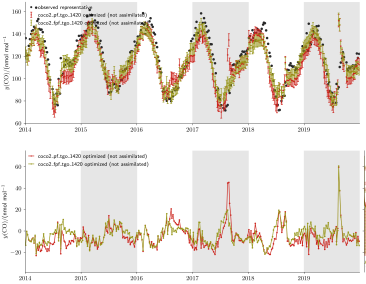


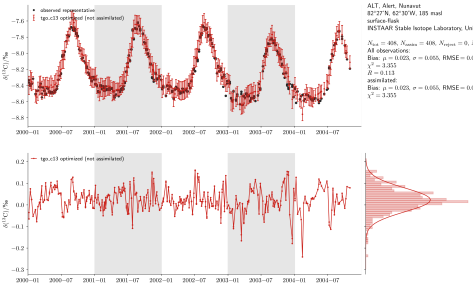
Satellite Xu et al. (Auke)

UNFCCC (Marnix)

A single framework for CO_2, CO, O_2,$\delta(^{13}C)$,$\Delta(^{14}C)$



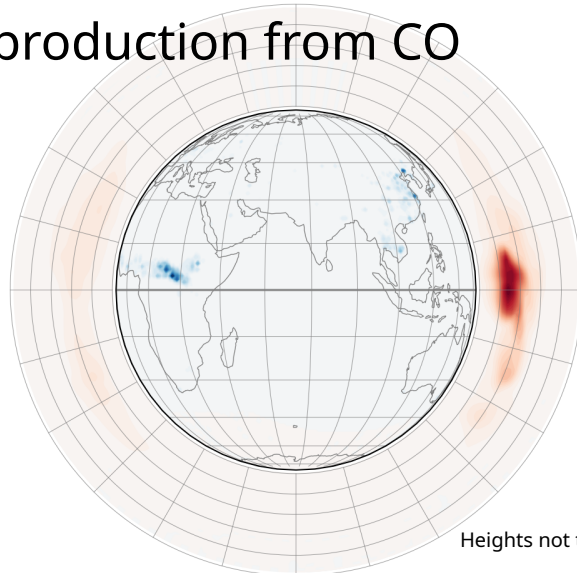




SW: effect of CO on CO_2 inversions
(Remco de Kok)

Inversion with CO_2 production from CO

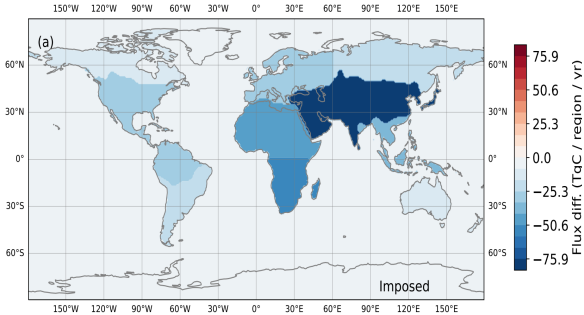
- Data from Bo Zheng (Tsinghua Univ.)
- 2000-2021
- Biomass burning (BB) and fossil fuel (FF) (each ~0.2 PgC/yr)
- Compare “3D” run with normal “REF” GCP run



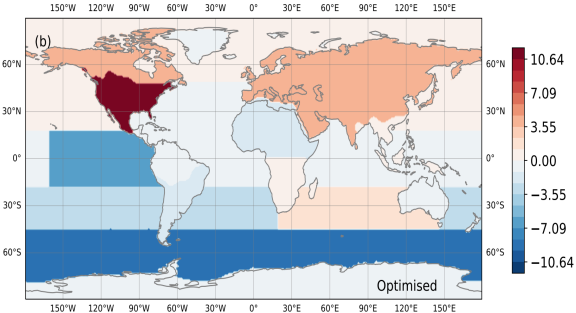
Heights not to scale

Inversion result (3D – REF): annual mean

Difference FF + BB (imposed)



Difference BIO + OCE (optimised)



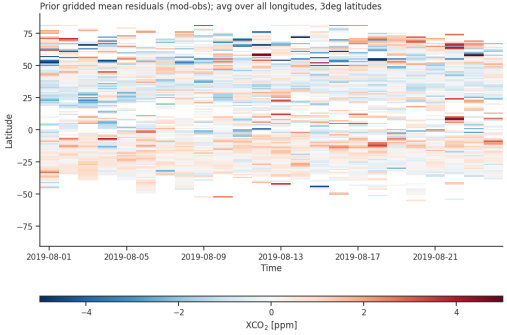
LW-SW: Carbon exchange over the Amazon during the wildfire season of 2019
(Anne-Wil van den Berg)

Long-Window Short-Window CO

Long Window

1. Close large scale CO - CO_2 and budget
2. Background observation network
3. Reduce mean bias

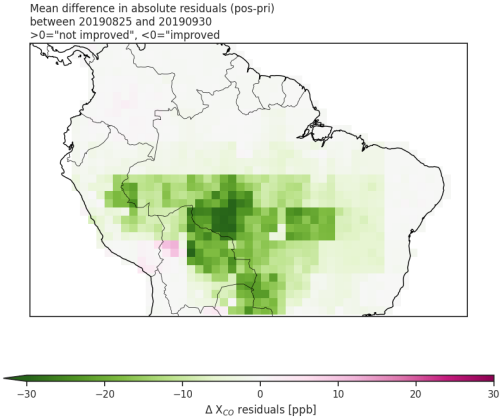
Validation: satellite OCO-2 data



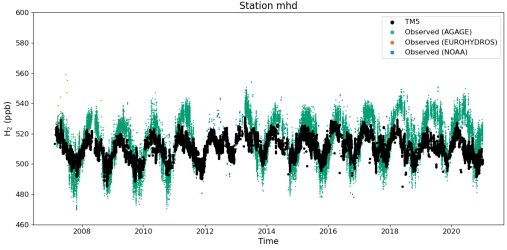
Short-Window

1. Focus on Amazon wildfire 2019
2. Small 3 day windows cycles
3. Optimization with MOPIT XCO data

MOPIT-XCO posterior improvement



Towards hydrogen inversions: Git merge Firmin Stroo

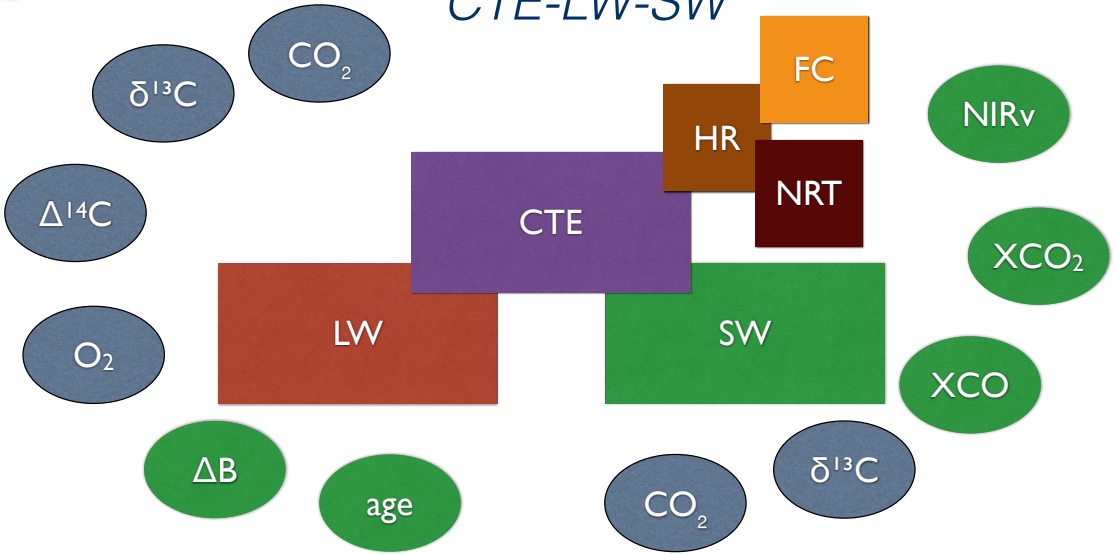


CTDAS main
TM5-MP main

H_2 inversions

Full chemistry in EnKF?

CTE-LW-SW


CO2
δ13C
FC
HR
NRT
NIRv
Δ14C
CTE
XCO2
LW
SW
O2
XCO
ΔB
age
CO2
δ13C

Ongoing

Faster advection on reduced grid

1. domain transform:
   1. xy2yz
   2. recude-advectx-expand single routine
   3. yz2xy
2. x advection now takes as long as y advection
3. `par.nx=4`

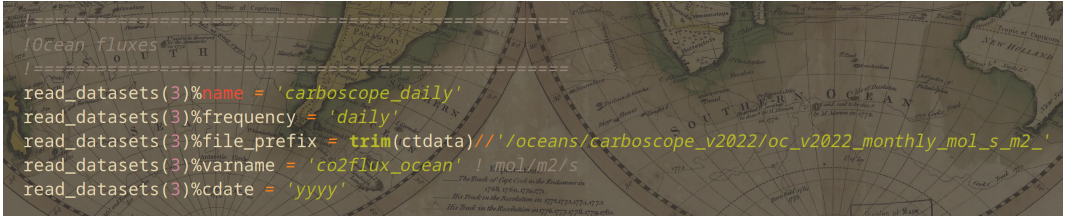
	domain transform	base code
mpi comm	191.60 s	242.01 s
other	50.26 s	189.76 s

Issue when doing 3d interpolation

1. `update_halo` routines do not update the corners of the halo
2. Biases in output when output is in the corner-box of the domain
3. Hurts especially with lots of satellite data
4. Quick fix: before output:
    1. `update_halo`
    2. `update_halo_jband` ! updates also the corner

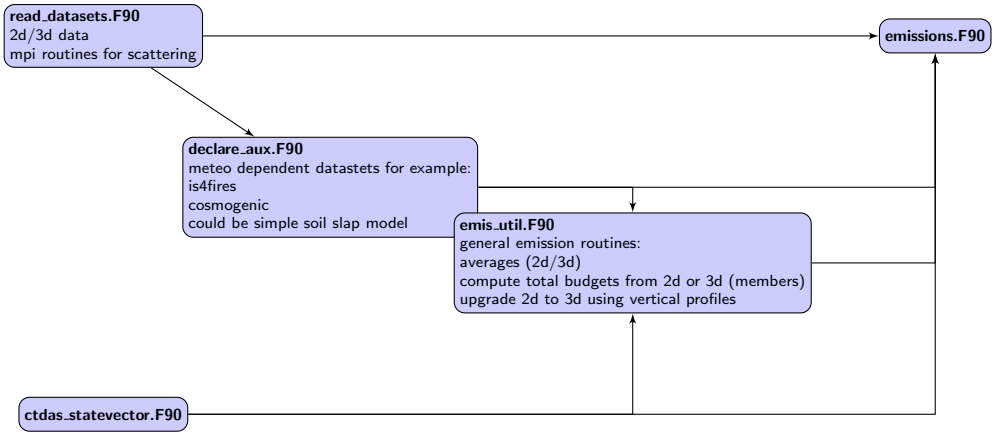
TM5-MP devs that made this possible

1. `read_data.F90` that mimics the meteo "set" routine
   1. Only emissions / tracers in use will be read
   2. Single read-scatter routines for 2d 3d fields



```
!=============================================
!Ocean fluxes
!=============================================
read_datasets(3)%name = 'carboscope_daily'
read_datasets(3)%frequency = 'daily'
read_datasets(3)%file_prefix = trim(ctdata)//'/oceans/carboscope_v2022/oc_v2022_monthly_mol_s_m2_'
read_datasets(3)%varname = 'co2flux_ocean' ! mol/m2/s
read_datasets(3)%cdate = 'yyyy'
```

TM5-MP devs that made it possible



TM5-MP devs that made this possible

no more source files: parse source code

1. rc file pre-modified before expanding variables in rc file
2. chem param written on the fly
3. ... except for init files (planned)

```
self.dacycle = dacycle
self.load_rc(self.filename)      # load the specified rc-file
self.validate_rc()               # validate the contents
self.write_chem_param()          # write the chem_param.F90 source file
```

my open source python projects

vremap
`https://github.com/JJDHooghiem/vremap`
mass conserved vertical remapping of TM5 restart files

1. creating init
2. working with other models
3. developed for Open-IFS-CC

pyenkf
`https://github.com/JJDHooghiem/pyenkf`

1. python api to enkf routine written in fortran
2. mabye cuBLAS in the future?