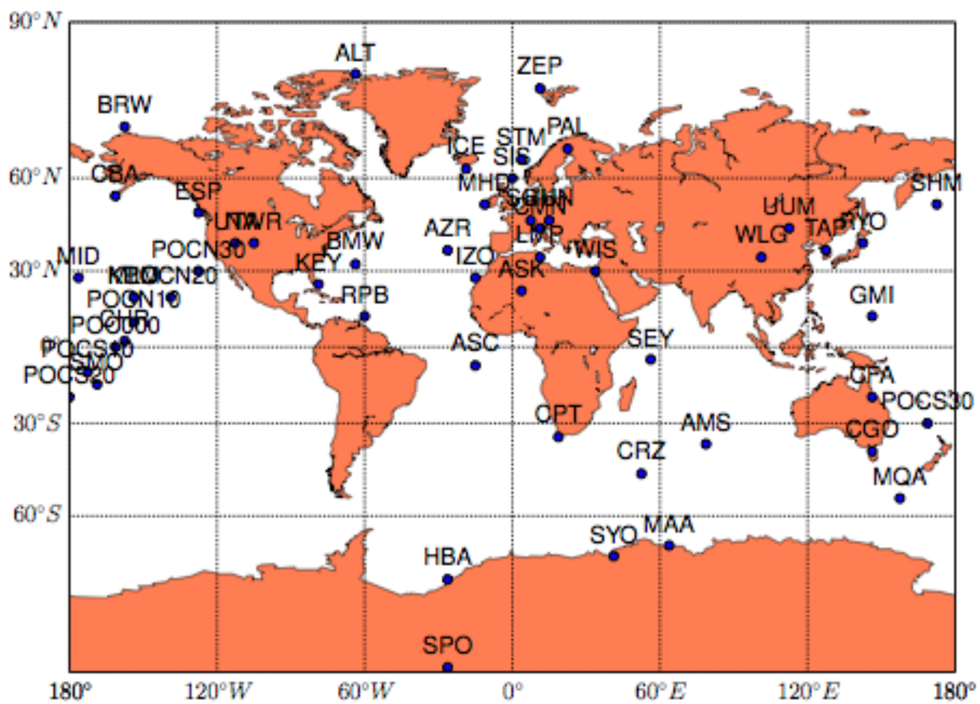
Additional exemples of Data Assimilation With ORCHIDEE & Tool developmenrts

Philippe Peylin, Cedric Bacour, Natasha MacBean, Sebastien Leonard, Pascal Maugis, Fabienne Maignan, Frederic Chevallier, Vladislav Bastrikov, Sylvain Kuppel, Diego Santaren,…..

LSCE

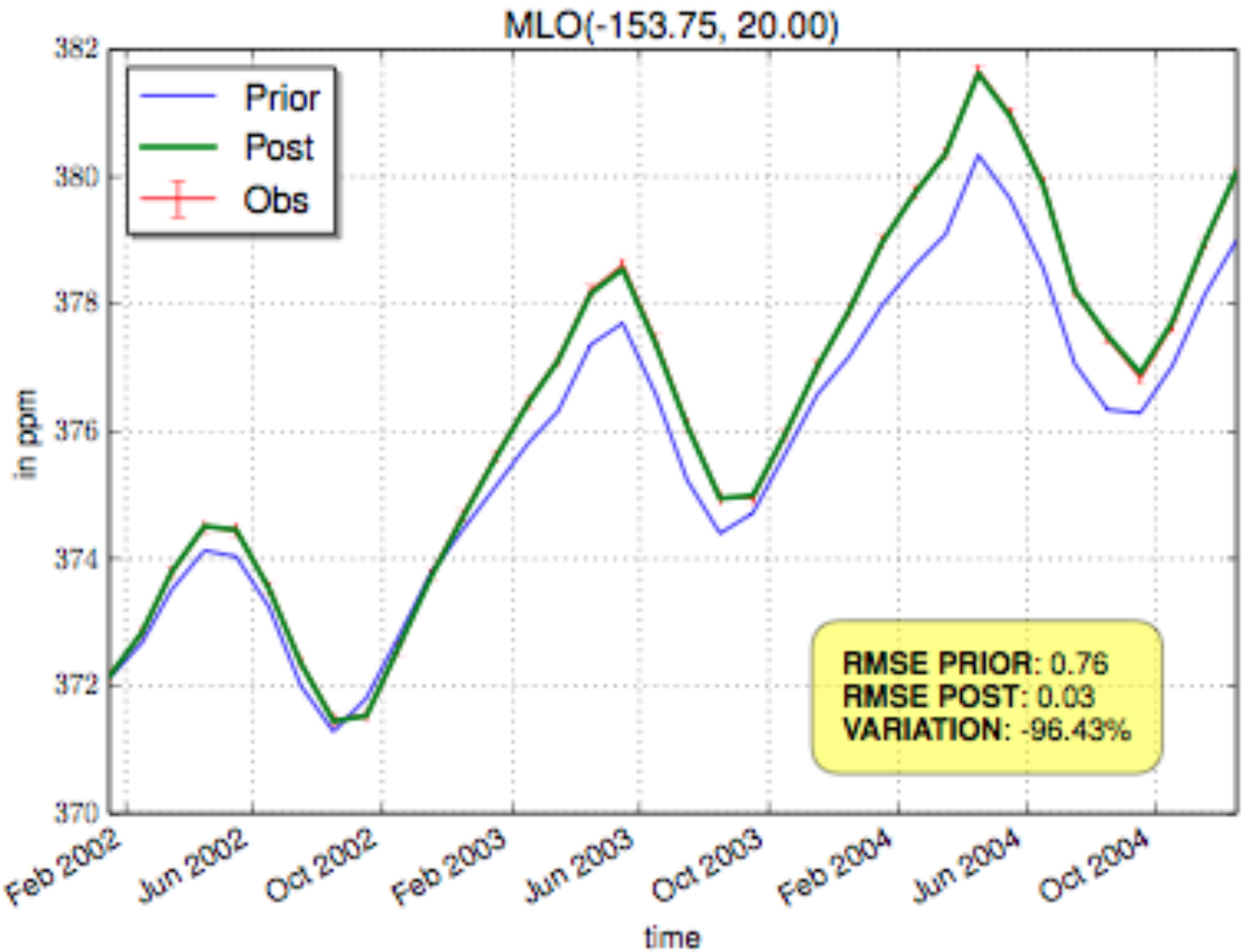
Assimilation of atmospheric CO2 data

- **Pseudo-Obs tests:**
  - Create pseudo obs with some parameter values (TRUTH)
  - Start with a different prior value
  - Variational optimization

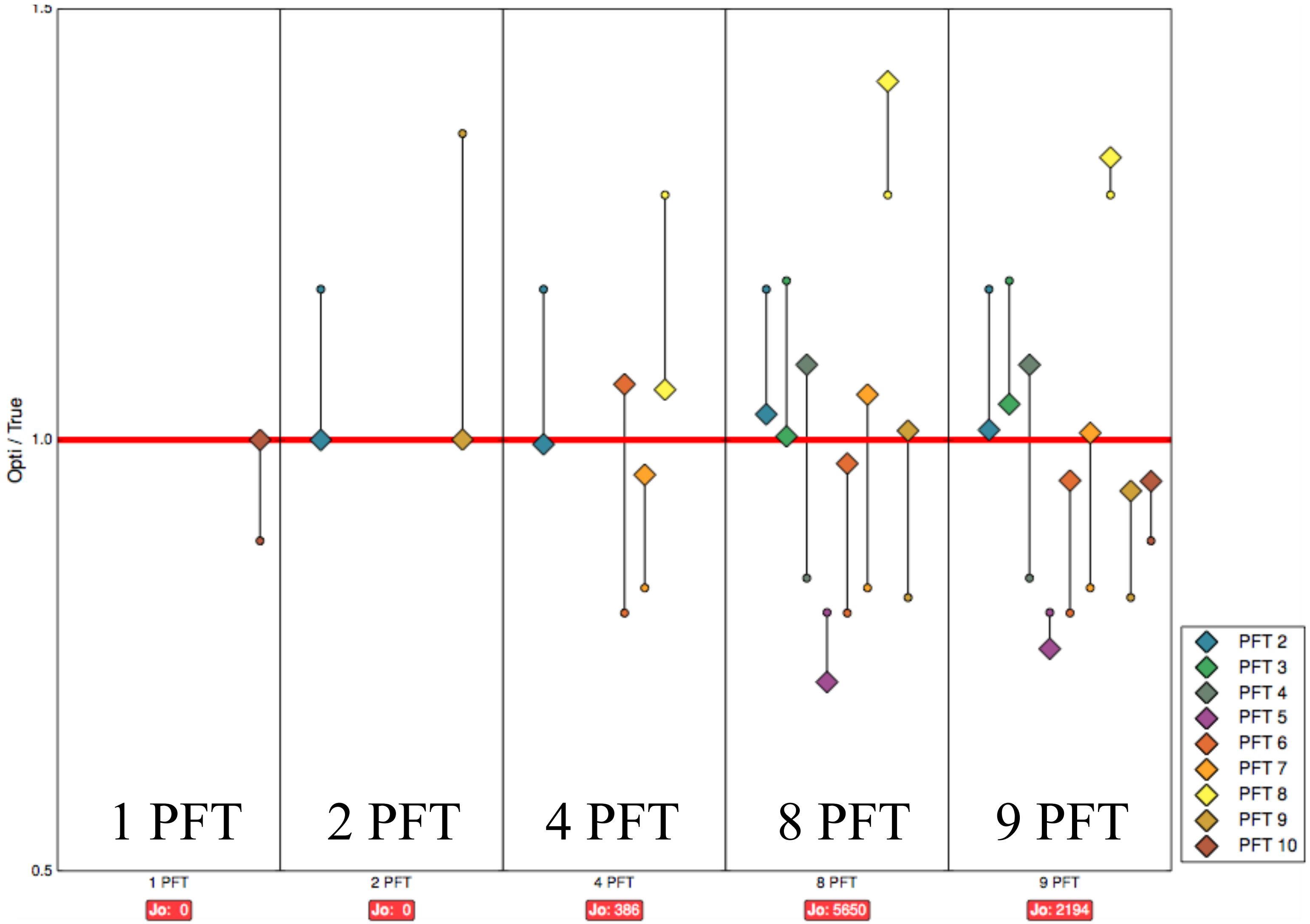


Optimization of Max Photosynthetic capacity

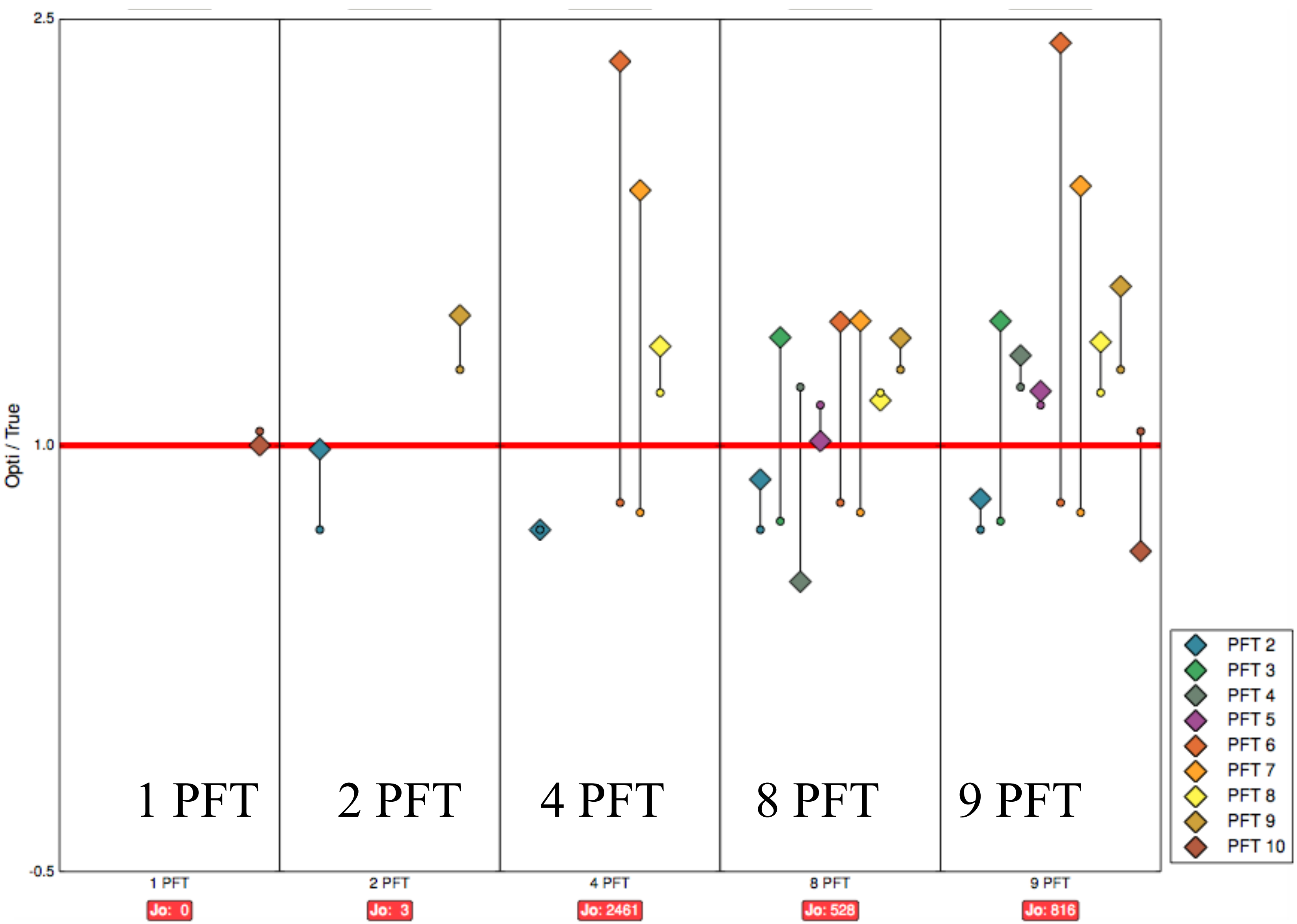
➔ **Perfect Fit to the data with only 1 or 2 PFTs.**



Optimization of Max Photosynthetic capacity

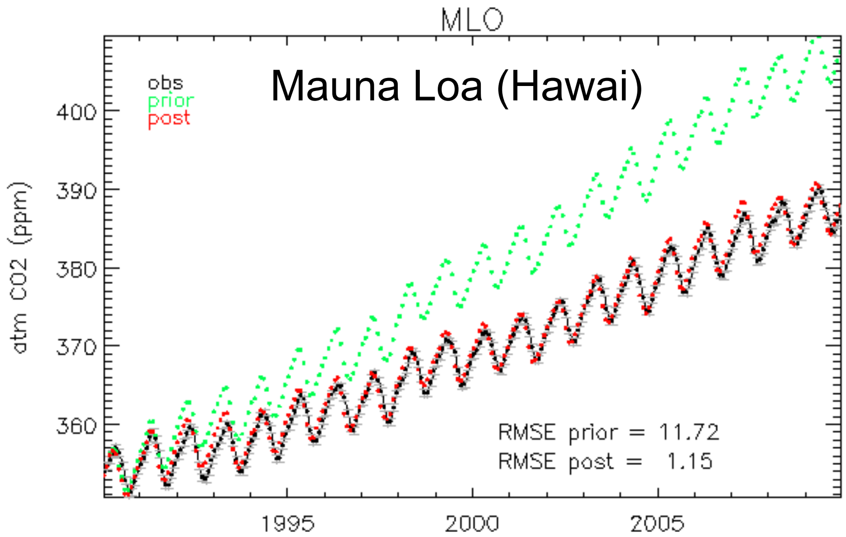


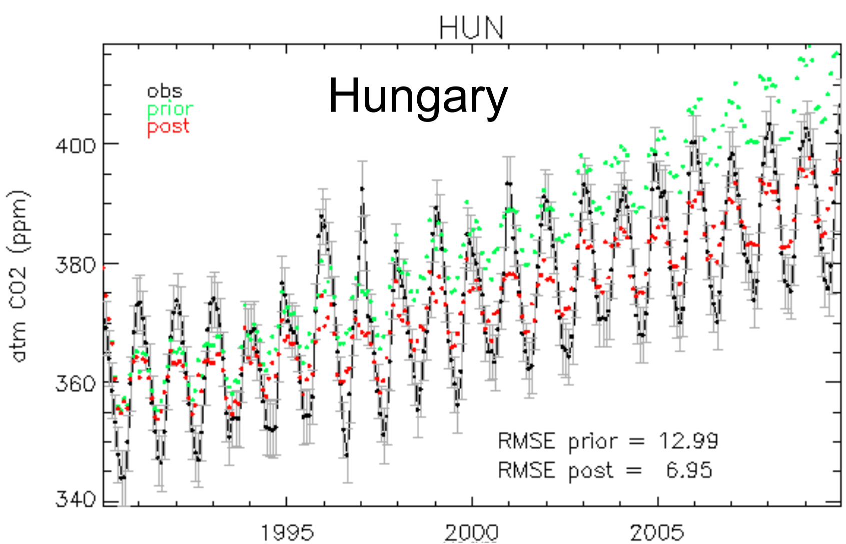
Optimization of “soil water stress” parameter



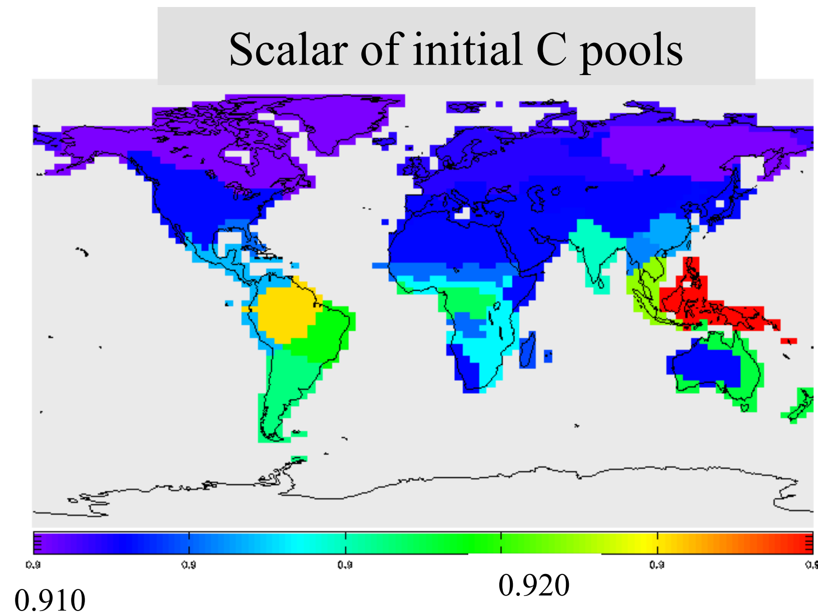
Assimilation of real Atmospheric data

Parameter optimisation





Main changes concern the initial soil carbon pool sizes

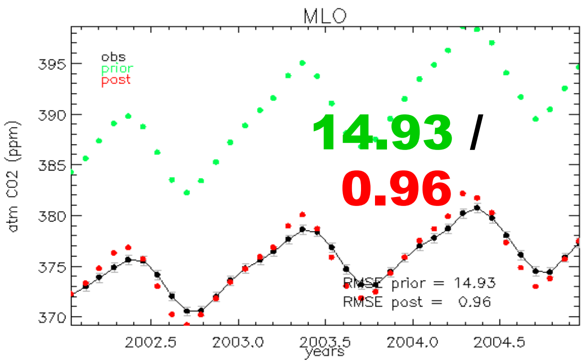


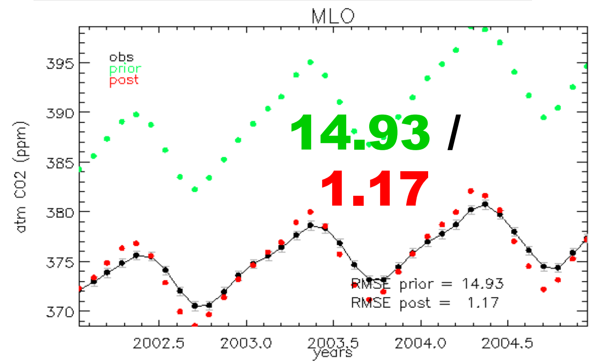
Simultaneous optimization

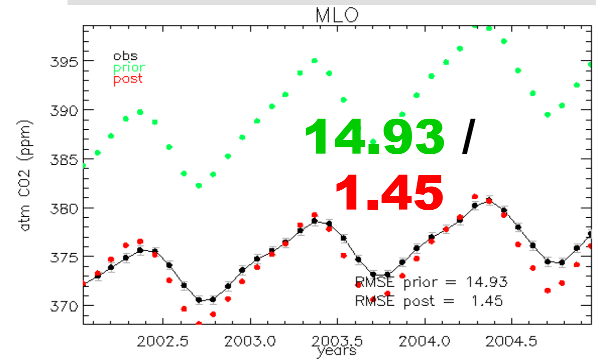
RMS prior / RMS poste

	Atmospheric CO_2	Atmospheric CO_2 + MODIS NDVI	Atmospheric CO_2 + NDVI + FluxNet data

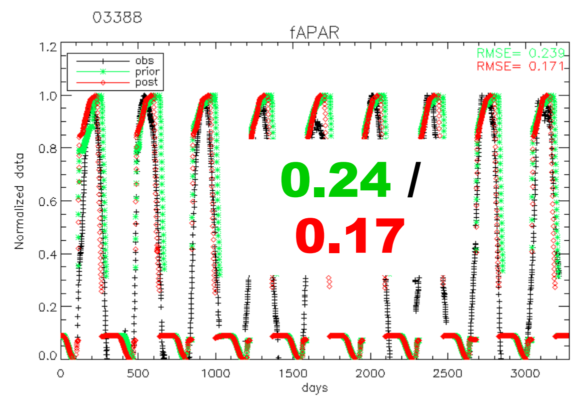
Mauna Loa

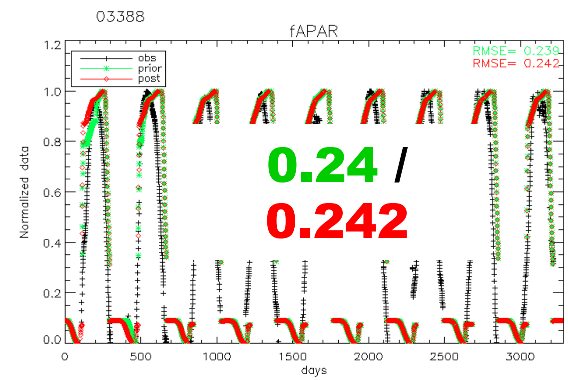






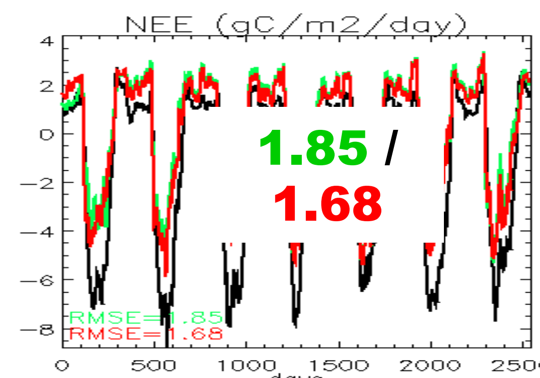
MODIS-NDVI (Temp Dec. For.)





FluxNet DE-HAI (Temp Dec. For.)

➔Performances of the fit decrease with the number of data streams



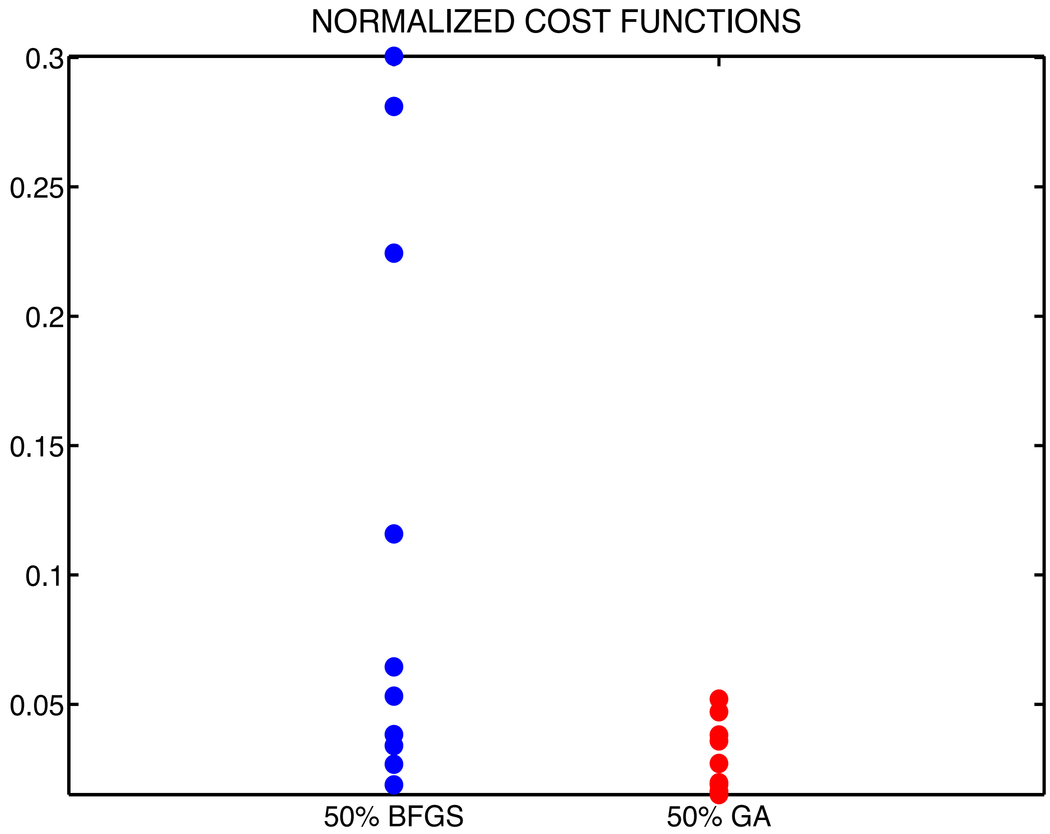
Ensemble vs Variational optimization

- **Create Pseudo-Data with randomly perturbed parameters (within allowed range)**
- **Cost function (J) with no prior term**
- **10 members**
  - Variational scheme: 10 different first guest X
  - Genetic Algo. : 10 different experiments
- **Use J_{ref} with ORCHIDEE standard param.**

Genetic Algo. vs Variational : performances !

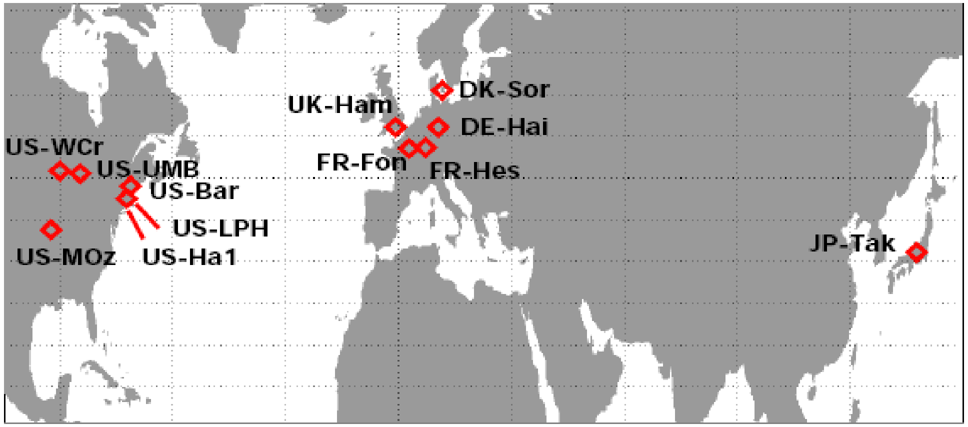
➤ Random perturbation over 50% of the parameter range allowed variation

J / J_{ref}



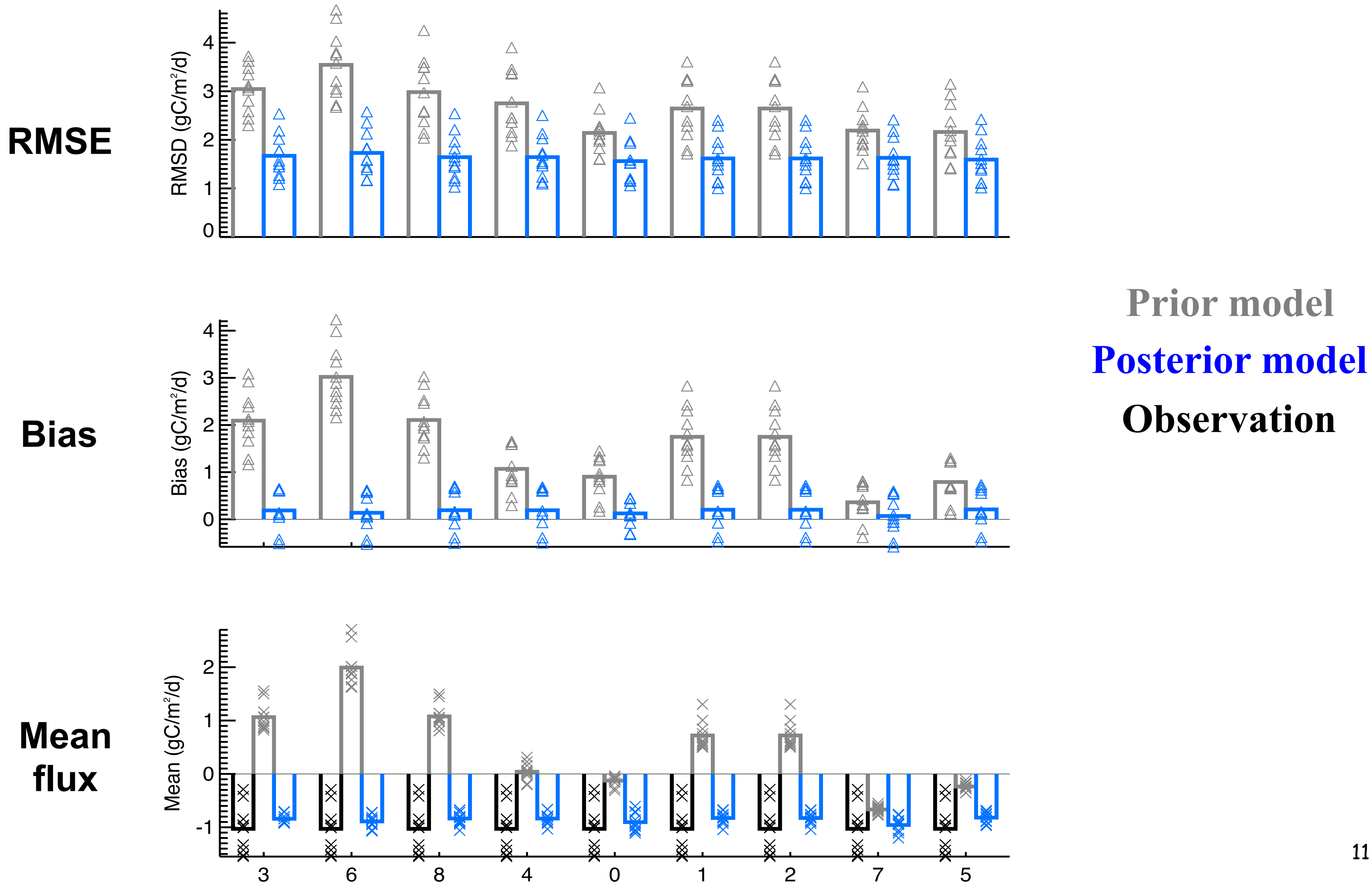
Varitional **Genetic Algo.**

- How does the minimization performance change with more site-data in the cost function ?
- 11 sites for a given Plant Functional Type
  Using observed daily measurements (NEE, LE)



- First test with only the Variational scheme
  (9 optimizations with 9 random first guest)

Results for 9 Variational optimizations



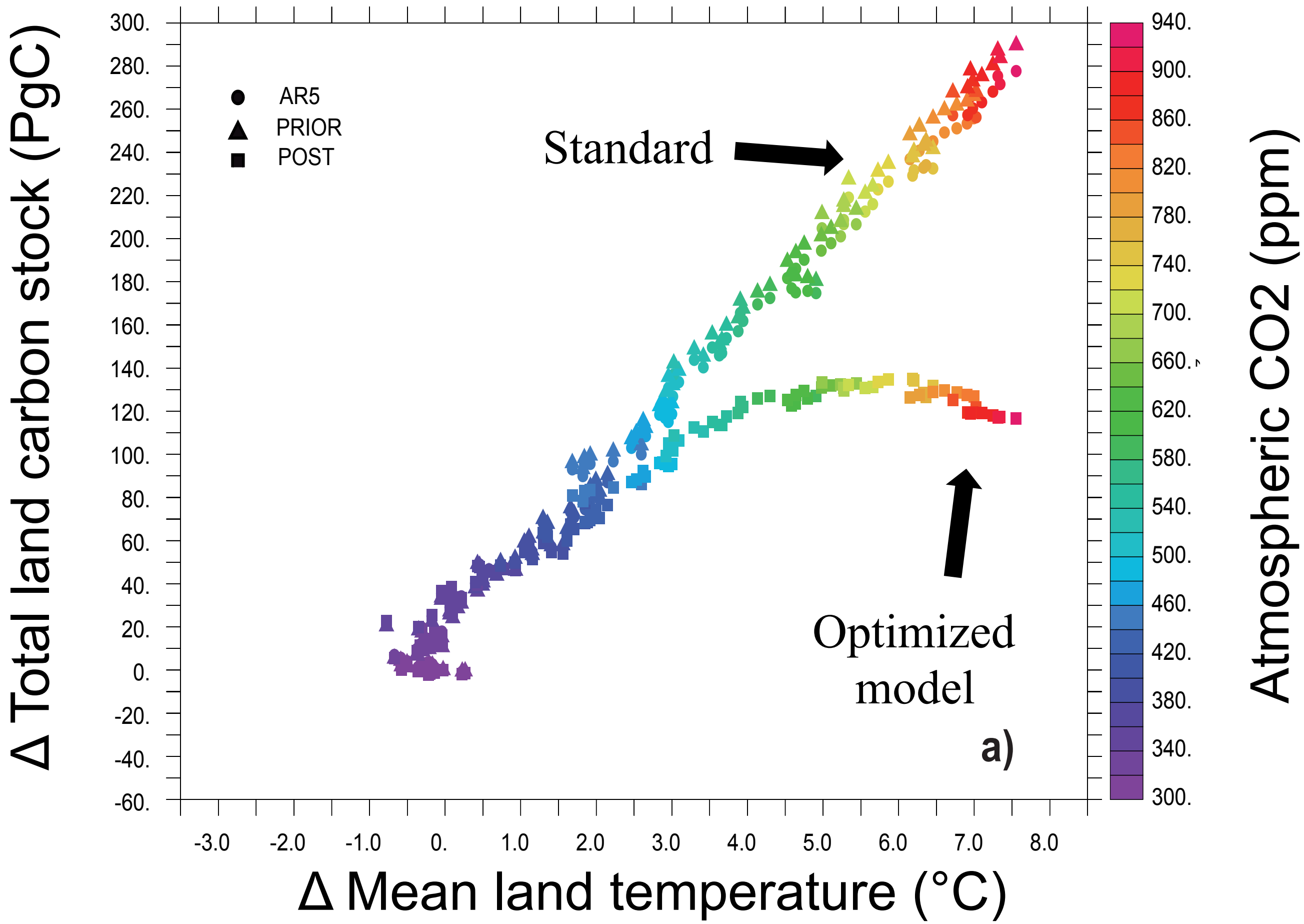
Several applications…

- **Impact the global carbon fluxes**
  - Current fluxes
  - Long term fate of the land carbon stocks

- **Results of the optimization already used in main versions of ORCHIDEE :**
  - TRUNC : VCMAX
  - DOFOCO: Respiration parameter

Impact on prognostic simulations (ISI-MIP)

- Using CMIP5 climate scenario (HadGEM2) bias corrected with RCP8.5 CO2 concentration
- Run ORCHIDEE: Standard vs optimized parameters

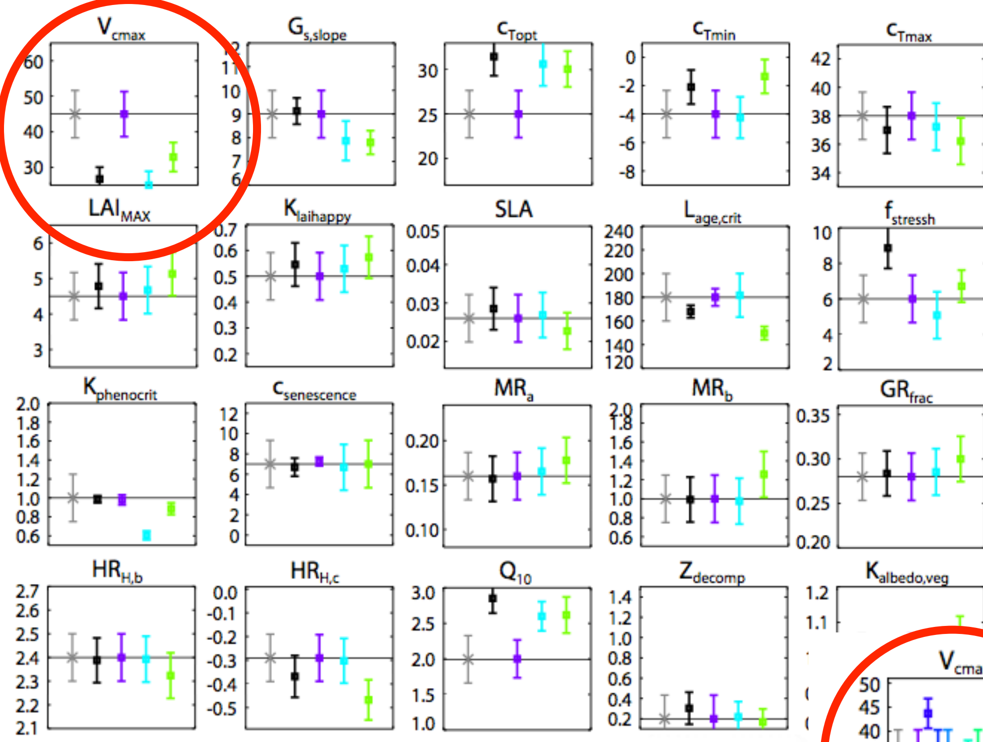


LSCE

Use of Optimization results from FluxNet data to improve parameters

Example of VCMAX at high latitude

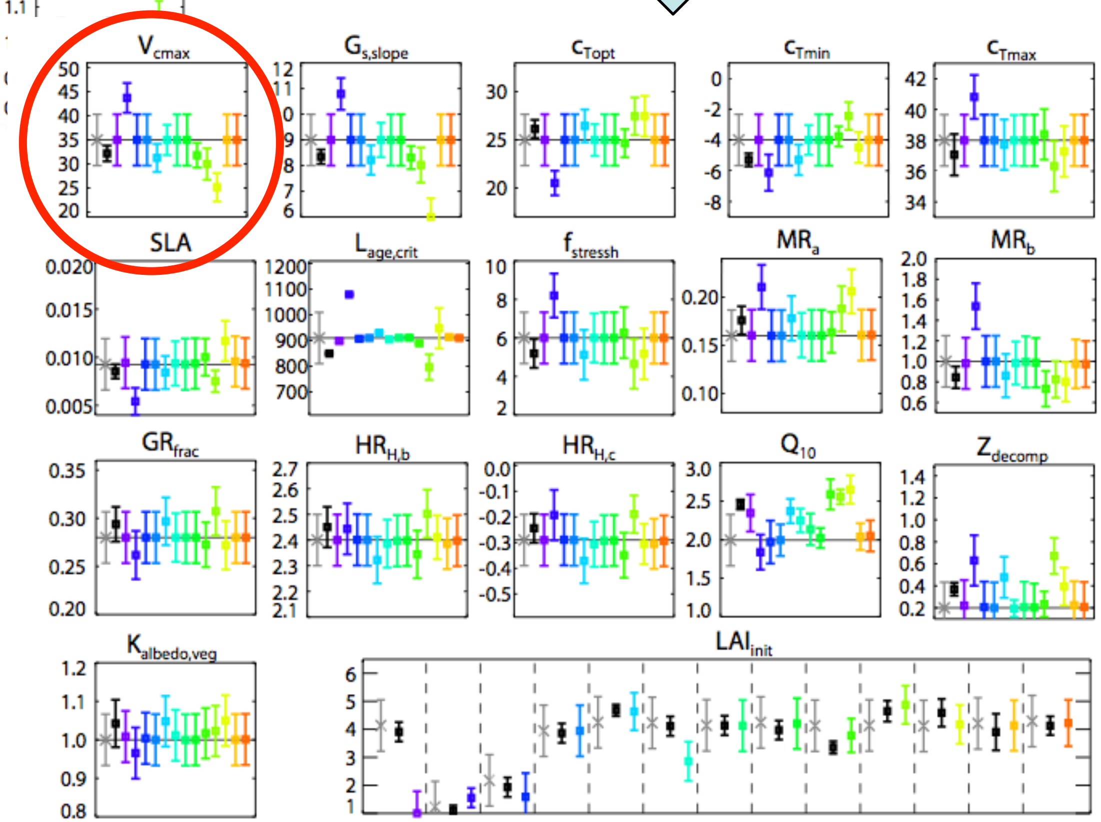
Boreal PFTs:



Broadleaf Deciduous

vs

Needleleaf Evergreen



From Krinner 2003

Table 1. PFTs and PFT-Specific Parameters in ORCHIDEE[a]

PFT	$V_{cmax,opt}$	T_{opt}	λ_{max}	z_{root}	α_{leaf}	h	A_c	T_s	H_s
TrBE	50	37	10	1.25	0.12	25	910	-	0.3
TrBR	60	37	10	1.25	0.14	25	180	-	0.3
TeNE	37.5	27	5	1.	0.14	15	910	-	-
TeBE	37.5	32	5	1.25	0.14	15	730	-	-
TeBS	37.5	28	5	1.25	0.14	15	180	12.5	-
BoNE	37.5	25	4.5	1.	0.14	10	910	-	-
BoBS	37.5	25	4.5	1.	0.14	10	180	5	-
BoNS	35	25	4	1.25	0.14	10	180	7	-
NC3	70	$27.5 + 0.25T_l$	2.5	0.25	0.20	0.2	120	4	0.2
NC4	70	36	2.5	0.25	0.20	0.2	120	5	0.2
AC3	90	$27.5 + 0.25T_l$	6	0.25	0.18	0.4	150	10	0.2
AC4	90	36	3	0.25	0.18	0.4	120	10	0.2

In 1.9.6

From Kattge VCMAX paper

Table 3 Means and statistical moments for observed leaf nitrogen content in natural environment ($N_{a,nat}$, $g\,m^{-2}$), predicted carboxylation capacity at 25 °C (V^{25}_{max}, $\mu mol\,m^{-2}\,s^{-1}$) and nitrogen use efficiency of carboxylation capacity derived by division of V^{25}_{max} by $N_{a,nat}$ (NUE, $\mu mol\,CO_2\,g\,N^{-1}\,s^{-1}$)

		$N_{a,nat}$					V^{25}_{max}						NUE			
PFT	$n_{Na,nat}$	Mean	SD	SE	Sk	Ku	Mean	SD	SE	Sk	Ku	BQ	Mean	SE	Sk	Ku
1 Tropical trees (oxisols)	371	2.17	0.80	0.04	0.66	0.30	29.0	7.7	0.4	0.61	0.56	*62	14.02	2.26	1.72	4.72
2 Tropical trees (nonoxisols)	107	1.41	0.56	0.05	1.76	5.41	41.0	15.1	1.5	1.88	6.45	**94	29.60	2.54	0.54	2.45
3 Temperate broadleaved. evergreen trees	65	1.87	0.93	0.11	0.88	0.14	61.4	27.7	3.4	0.89	0.18	41	33.75	2.32	1.4	3.00
4 Temperate broadleaved deciduous trees	404	1.74	0.71	0.04	0.77	0.78	57.7	21.2	1.1	0.78	0.83	35	33.79	2.37	2.94	14.93
5 Evergreen coniferous trees	220	3.10	1.35	0.09	0.74	1.38	62.5	24.7	1.7	0.77	1.53	29	20.72	1.78	1.38	3.93
6 Deciduous coniferous trees	27	1.81	0.64	0.12	1.08	0.49	39.1	11.7	2.3	1.08	0.61	53	22.05	1.61	0.53	0.61
7 Evergreen shrubs	130	2.03	1.05	0.09	1.60	2.65	61.7	24.6	2.2	1.68	3.19	52	32.09	4.24	0.64	1.23
8 Deciduous shrubs	179	1.69	0.62	0.05	0.61	0.47	54.0	14.5	1.1	0.67	0.76	160	33.14	4.38	1.27	3.21
9 C3 herbaceous	254	1.75	0.76	0.05	1.42	2.94	78.2	31.1	2.0	1.44	3.10	42	45.29	2.57	1.79	8.83
10 C3 crops	***209	1.62	0.61	0.04	0.41	0.31	100.7	36.6	2.5	0.43	0.40	120	62.75	3.65	3.13	27.42

$n_{Na,nat}$: number of data, SD and SE: standard deviation and standard error, Sk: skewness, Ku: kurtosis, BQ: estimates of V^{25}_{max} by Beerling & Quick (1995), for comparison. *Tropical evergreen trees; **Tropical deciduous trees; ***All N_a data. PFT, plant functional type.

From Krinner 2003

Table 1. PFTs and PFT-Specific Parameters in ORCHIDEE[a]

PFT	$V_{cmax,opt}$	T_{opt}	λ_{max}	z_{root}	α_{leaf}	h	A_c	T_s	H_s
TrBE	50	37	10	1.25	0.12	25	910	-	0.3
TrBR	60	37	10	1.25	0.14	25	180	-	0.3
TeNE	37.5	27	5	1.	0.14	15	910	-	-
TeBE	37.5	32	5	1.25	0.14	15	730	-	-
TeBS	37.5	28	5	1.25	0.14	15	180	12.5	-
BoNE	37.5	25	4.5	1.	0.14	10	910	-	-
BoBS	37.5								
BoNS	35								
NC3	70								
NC4	70								
AC3	90								
AC4	90								

In 1.9.6

➔Used these iformation to modify VCMAX for Boreal Broadleaf summergrean and Boreal Evergreen Needleleaf

paper

($N_{a,nat}$, g m^{-2}), predicted
y derived by division of

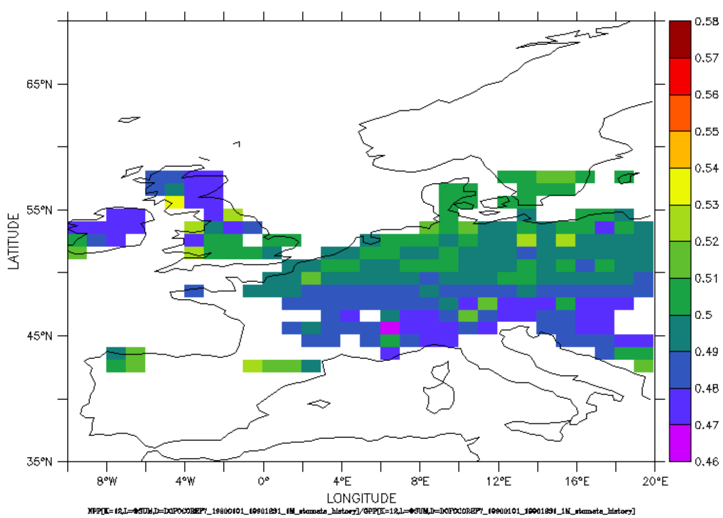
		$N_{a,nat}$					V_{max}^{25}						NUE			
PFT	$n_{Na,nat}$	Mean	SD	SE	Sk	Ku	Mean	SD	SE	Sk	Ku	BQ	Mean	SE	Sk	Ku
1 Tropical trees (oxisols)	371	2.17	0.80	0.04	0.66	0.30	29.0	7.7	0.4	0.61	0.56	*62	14.02	2.26	1.72	4.72
2 Tropical trees (nonoxisols)	107	1.41	0.56	0.05	1.76	5.41	41.0	15.1	1.5	1.88	6.45	**94	29.60	2.54	0.54	2.45
3 Temperate broadleaved. evergreen trees	65	1.87	0.93	0.11	0.88	0.14	61.4	27.7	3.4	0.89	0.18	41	33.75	2.32	1.4	3.00
4 Temperate broadleaved deciduous trees	404	1.74	0.71	0.04	0.77	0.78	57.7	21.2	1.1	0.78	0.83	35	33.79	2.37	2.94	14.93
5 Evergreen coniferous trees	220	3.10	1.35	0.09	0.74	1.38	62.5	24.7	1.7	0.77	1.53	29	20.72	1.78	1.38	3.93
6 Deciduous coniferous trees	27	1.81	0.64	0.12	1.08	0.49	39.1	11.7	2.3	1.08	0.61	53	22.05	1.61	0.53	0.61
7 Evergreen shrubs	130	2.03	1.05	0.09	1.60	2.65	61.7	24.6	2.2	1.68	3.19	52	32.09	4.24	0.64	1.23
8 Deciduous shrubs	179	1.69	0.62	0.05	0.61	0.47	54.0	14.5	1.1	0.67	0.76	160	33.14	4.38	1.27	3.21
9 C3 herbaceous	254	1.75	0.76	0.05	1.42	2.94	78.2	31.1	2.0	1.44	3.10	42	45.29	2.57	1.79	8.83
10 C3 crops	***209	1.62	0.61	0.04	0.41	0.31	100.7	36.6	2.5	0.43	0.40	120	62.75	3.65	3.13	27.42

$n_{Na,nat}$: number of data, SD and SE: standard deviation and standard error, Sk: skewness, Ku: kurtosis, BQ: estimates of V_{max}^{25} by Beerling & Quick (1995), for comparison. *Tropical evergreen trees; **Tropical deciduous trees; ***All N_a data. PFT, plant functional type.

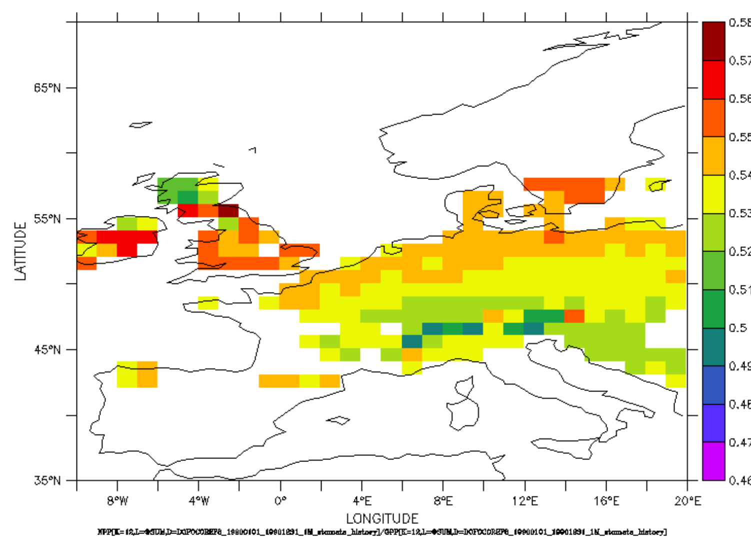
DA to calibrate species-param. in DOFOCO

- **DOFOCO group used DA to calibrate their model**
- **Assimilation of GPP & NPP extrapolated data over EU**
- **Optimization of the NPP/GPP ratio (i.e. an effective respiration parameter for Rauto) for 26 PFTs**

Prior values : Oak



Posterior values: Oak



Ongoing developments….

Tangent linear (Adjoint) of ORCHIDEE “TRUNC”

- Effort carried by Pascal Maugis
- Using TAF automatic differentiator !
- Long and tedious work → will imply few code structure changes

Optimization tools

→ Provide a “TOOL” for ORCHIDEE group

- Based on any possible version of ORCHIDEE
- Using finite difference or TL model for the sensitivity calculations
- Using eithe a Monte-Carlo (GA) or Variational approach
- PYTHON language; Based partly on PYVAR algorithm (F. Chevallier)
- Will include error calculation and error propagation
- Coupling with LMDZ to use atmospheric data
- Vladislav Bastrikov and Sebastien Leonard are working on it ! (co-supervised by P. Peylin, N. MacBean, C. Bacour,….)
- Expected completion : Summer 2014!

Conclusions / Recomandations…

- Parameter optimization with “Statistical tools” is promising for ORCHIDEE group
- Thresholds and non linearities complicate the optimization (Monte carlo schemes may be needed)
- Do not overtune parameters because of Model errors!
  ➔ NEED to properly define parameter range of variations & errors (sensitivity analysis help to select parameters)
- Using site data to improve global features is at high risk !
- Do not use only one optimized parameter value if several parameters were tune (error correlations)

Additional slides

Summary of C-cycle multi-data assimilation

- **Error estimation** → propagation on C flux & stock uncertainties
- **Flux Evaluation:**
  - Forest Biomass stock is improved against FAO data
  - GPP improve against "FluxNet-derived" products
- Using multiple data streams constrains different flux scales
- Insights on model improvement from posterior misfit
- Large impact on future land carbon budget

- Full analysis C sources and sinks underway: **www.carbones.eu**
- Need to use of other data streams → biomass
- Ongoing CCDAS intercomparison : GEOCARBON project

Limitations of a CCDAS…

➔ Strongly rely on a given model structure

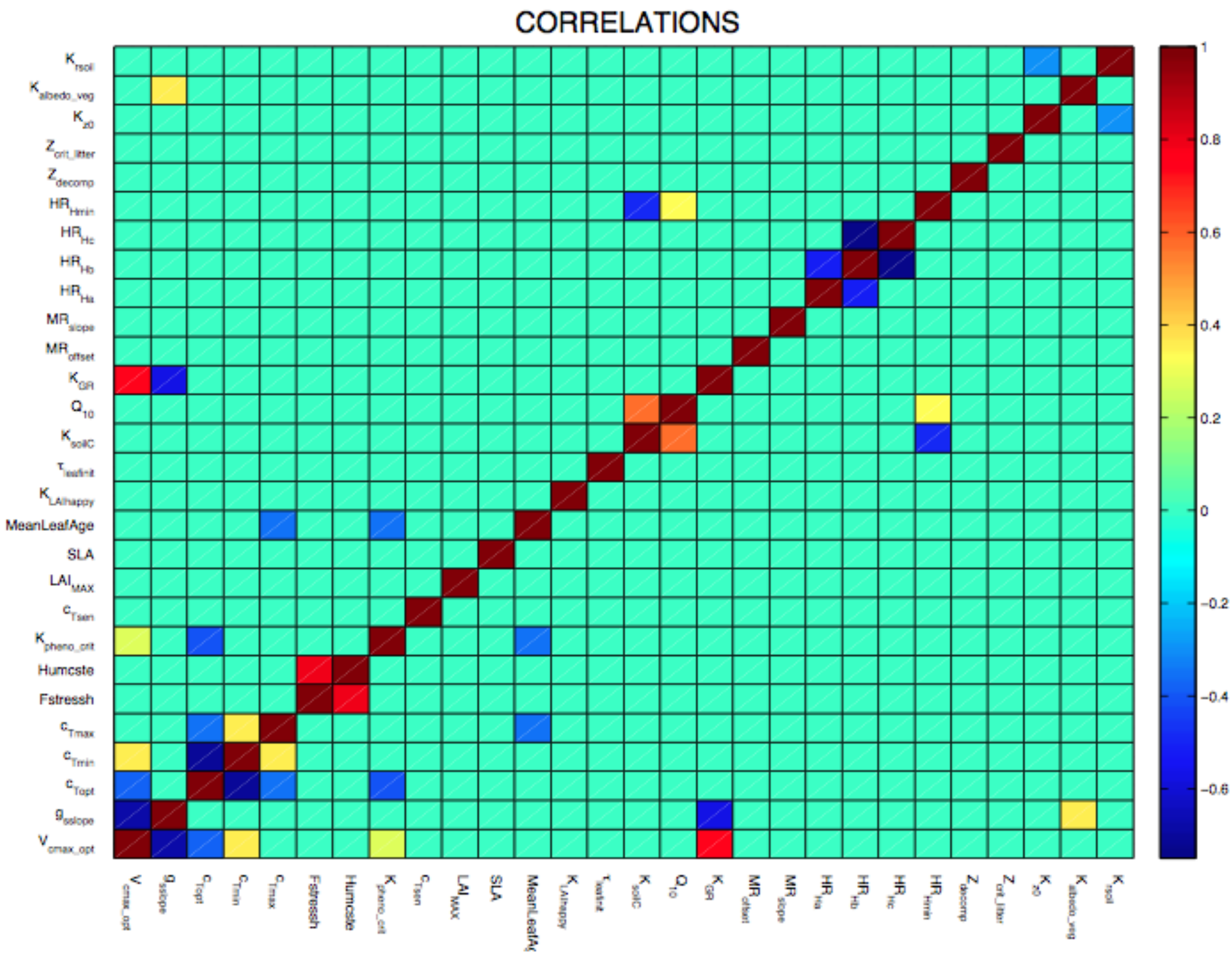
➔ Missing processes in the ecosystem model might lead to
- Wrong parameter estimates
- Poor model predictability (Strong biases)

➔ Initial conditions (C pools) difficult to optimize

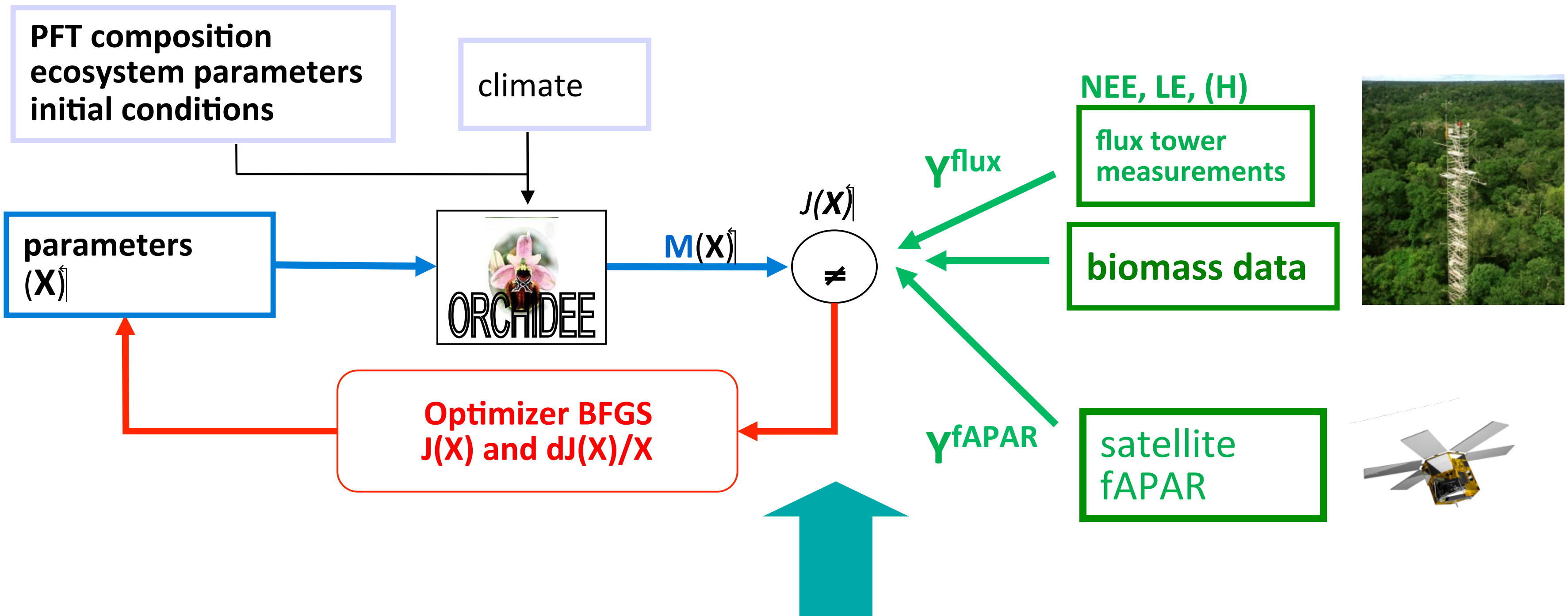
➔ Non-linearities might complicate the parameter optimization

➔ Need to :
- keep independent data for model output validation
- Keep classical Atmospheric inversion

Posterior error covariance matrix



Implementation..

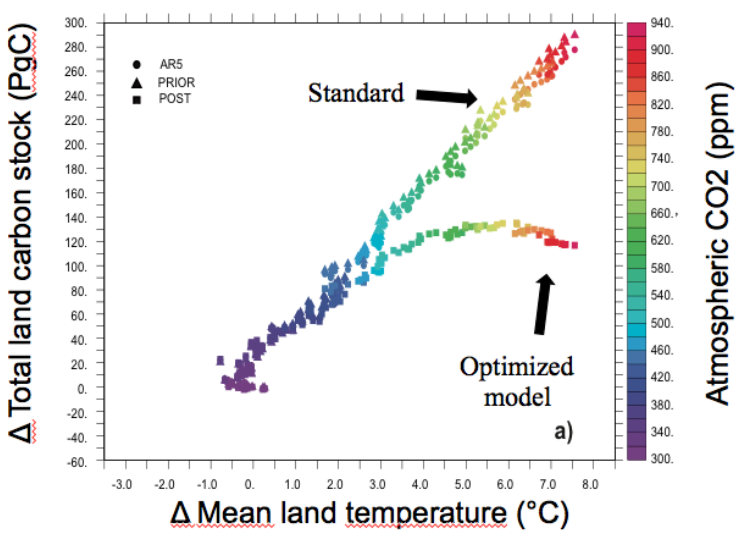


Cost function: $$J(x) = \frac{1}{2}\left[\left(y - M(x)\right)^t R^{-1}\left(y - M(x)\right) + (x - x_b)^t P_b^{-1}(x - x_b)\right]$$

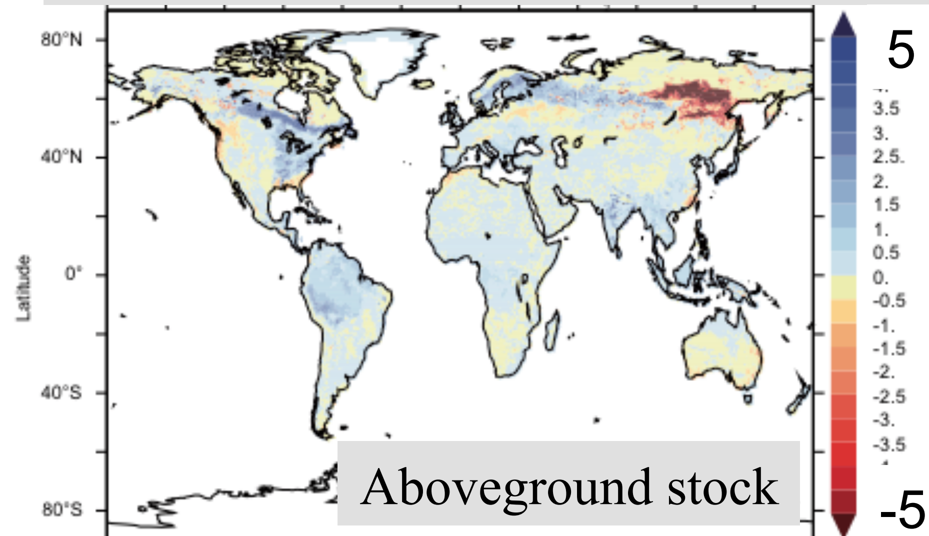
- **Iterative minimization using either:**
  - **Variational approach (with Tangent Linear model for DJ/dx)**
  - **Monte Carlo approach**

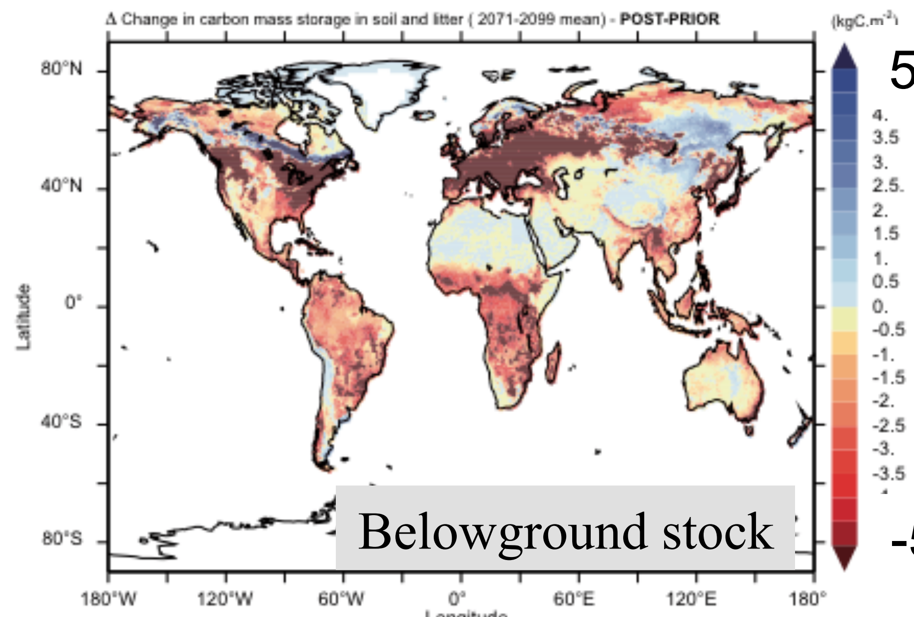
Impact on prognostic simulations (ISI-MIP)

- Large decrease of soil carbon storage above + 3° (changes in input & mineralisation)
- Only small decrease of vegetation carbon stock



2100 minus 2000: Δ C (kgC/m2)



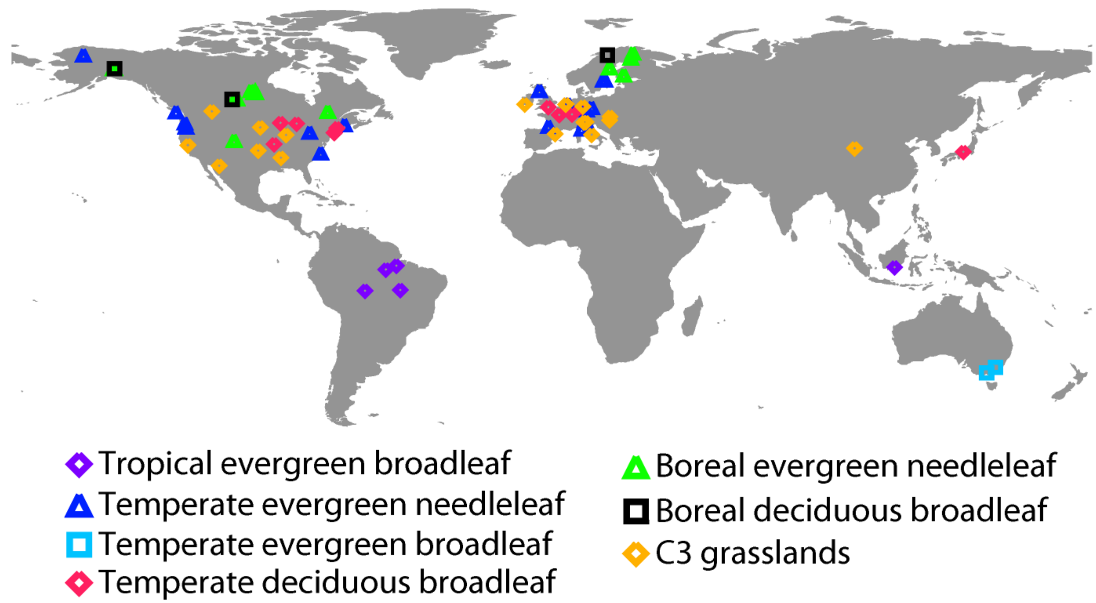


Few Assimilation exemples…

- **In situ FluxNet NEE & LE data**
  - Santaren et al. 2007, Verbeck et al. 2009, Kuppel et al. (2013)
  - Bacour et al. (in prep), Santaren et al. (2013, submitted)
- **NDVI satellite data**
  - McBean et al. (in prep)
- **FluxNet + Above Ground Biomass data**
  - Thum et al. (in prep)
- **Atm. CO2 data & all other data streams**
  - Methodological & 20yr flux reanalysis: in preparation
- **Tree ring data (Width, 13C and 18O)**
  - Launois et al., in preparation

Assimilation of flux NEE & LE data

➔ 78 sites from the FLUXNET network



➔ **Assimilation of daily-averaged NEE & LE**

➔ **Optimization of about** 20 parameters per PFTs

Parameter	Genericity
$V_{cmax,opt}$ $c_{T,min/opt/max}$ $L_{age,crit}$, $f_{stressh}$	PFT
$G_{s,slope}$	PFT
LAI_{MAX}, SLA	PFT
LAI_{init}	**Site**
$K_{lai,alloc}$	PFT
$K_{phenocrit}$, c_{senes}	PFT
MR_a, MR_b, GR_{frac}	PFT
Q_{10}, HR_b, HR_c Z_{decomp}	PFT
K_{soilC}	**Site**
$K_{albedo,veg}$	PFT

Future R&D needs…

Include Forest Biomass data in the CCDAS

- In situ biomass data with forest age to calibrate models
- Satellite biomass data to assess:
  - forest age distribution
  - inpact of disturbances (fires, land cover changes)

Joint assimilation of Carbone and Water cycles data

- To assess the impact of drought, weather extremes on C cycle
- To assess Water Use Efficiency
- Toward an full "Earth System Model" Data Assimilation procedure

C reanalysis over longer period *(i.e. 100yr; ERACLIM2)*

Toward Earth Climate Model optimization..

- Long term research..

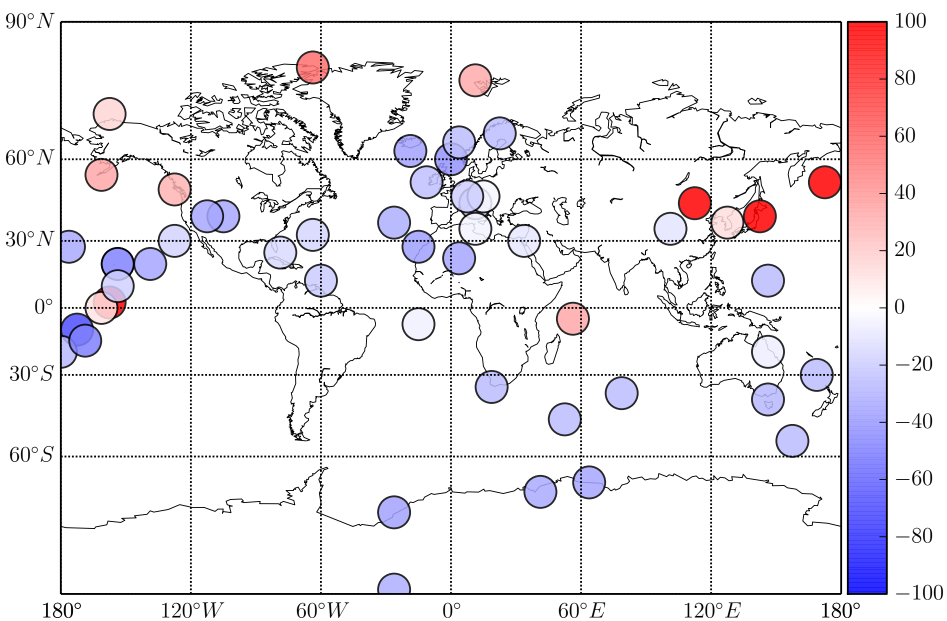
Fit to the CO2 data: Amplitude and Phase

Signal decomposition:

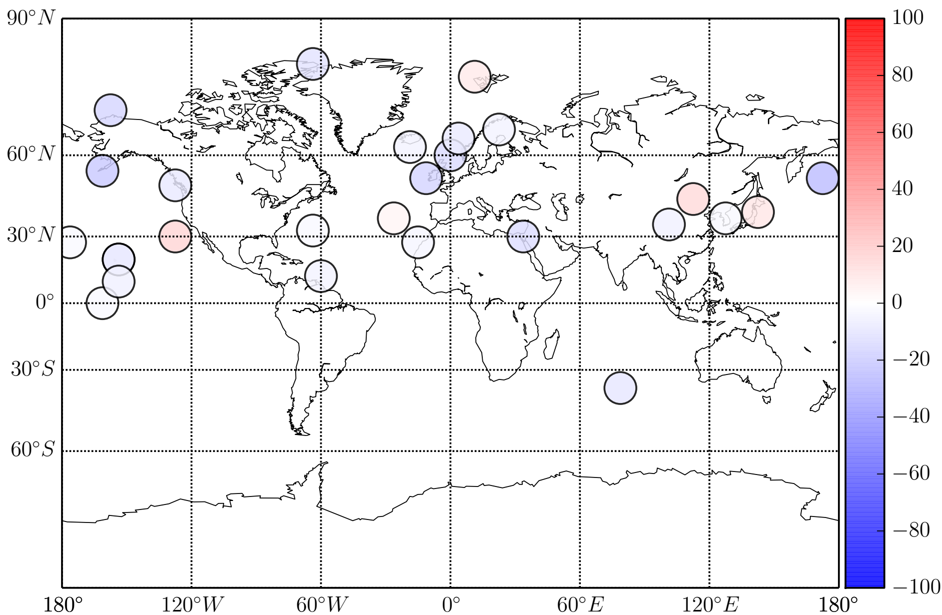
- Amplitude : monthly max – min
- Phase: Length of Carbon uptake period

Measure of the improvement: 1 – RMSE_poste/RMSE_prior

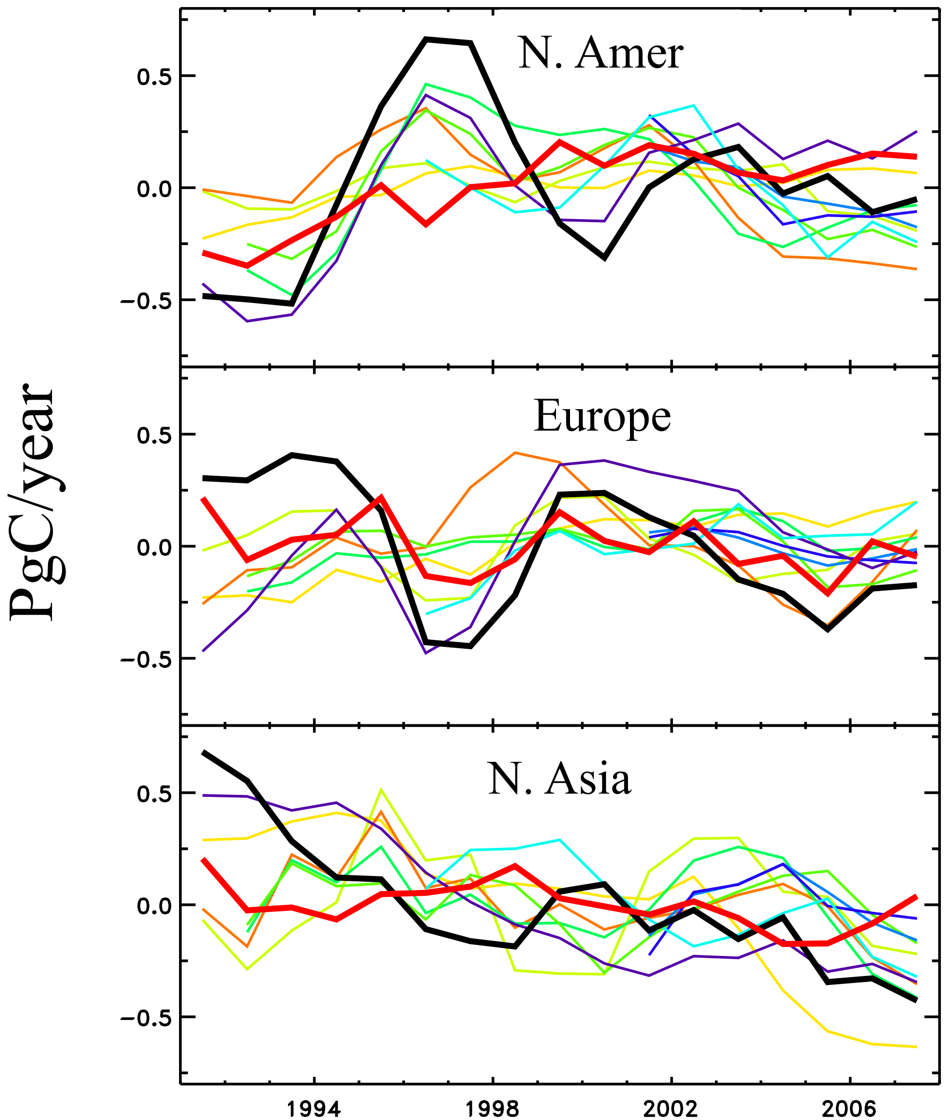
Amplitude improvement



Phase improvement



Estimated land C fluxes: long term trend



Comparison with Independant Atmos. Inversions

CCDAS-Parameter optim

CCDAS-fluxes optim

JENA_s96

LSCE_var

CTrac_US

CTrac_EU

C13CCAM

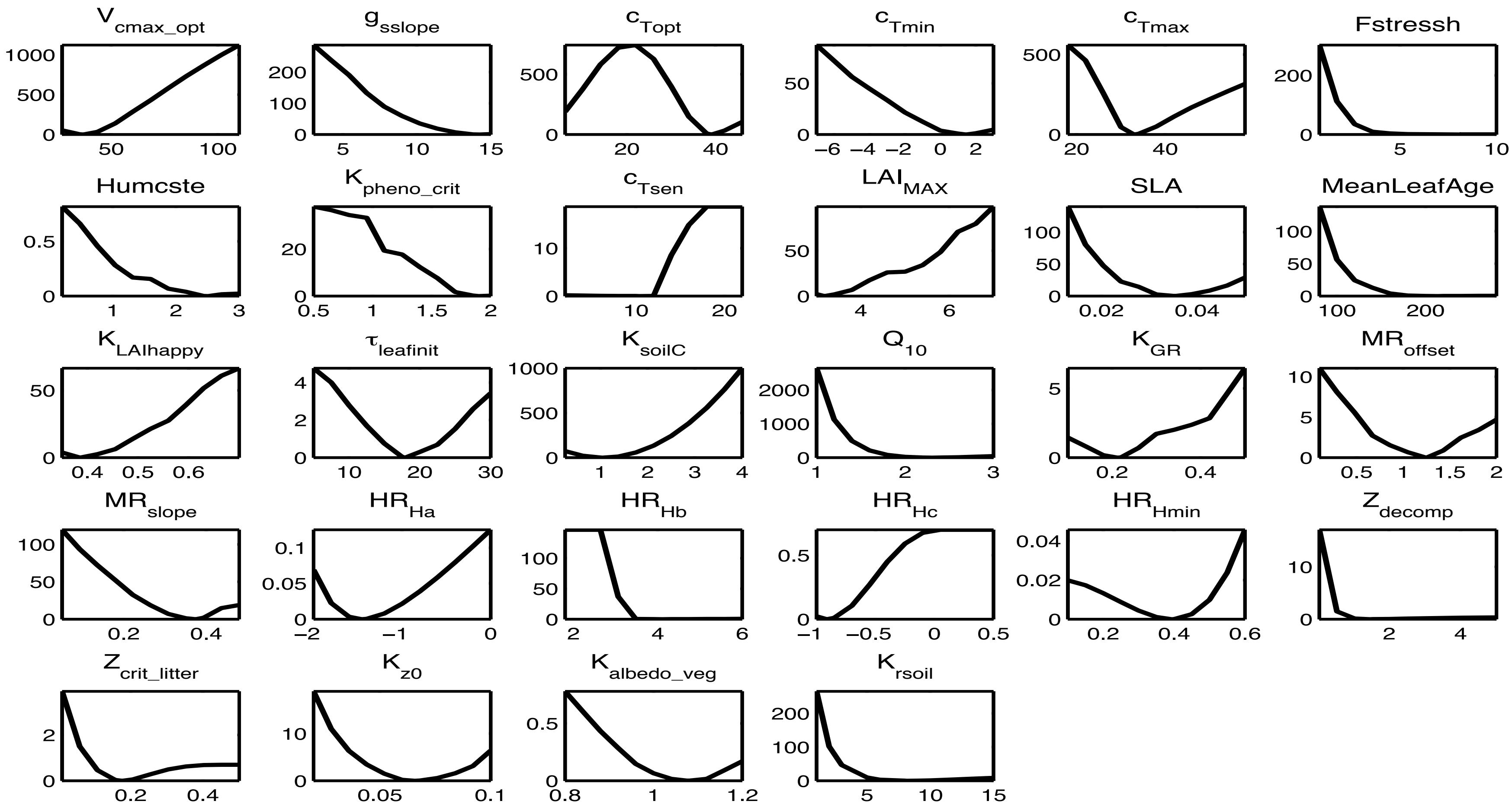
C13MATCH

TRCOM

RIGC

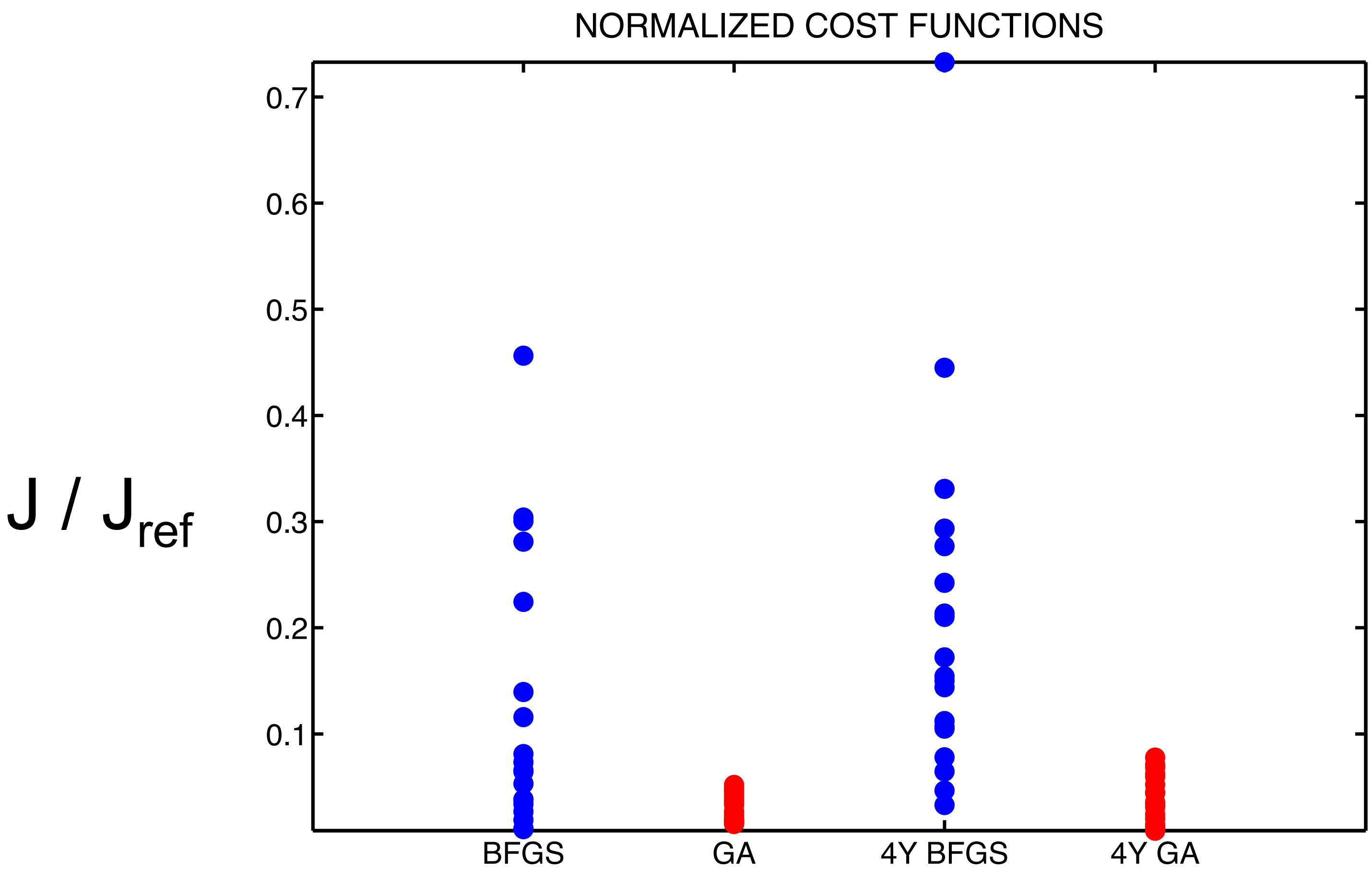
JMA

Cross section of J at "True" parameters (used for pseudo-data)



X range: Allowed range of variation for each parameter

Genetic Algo. vs Variational : performances !

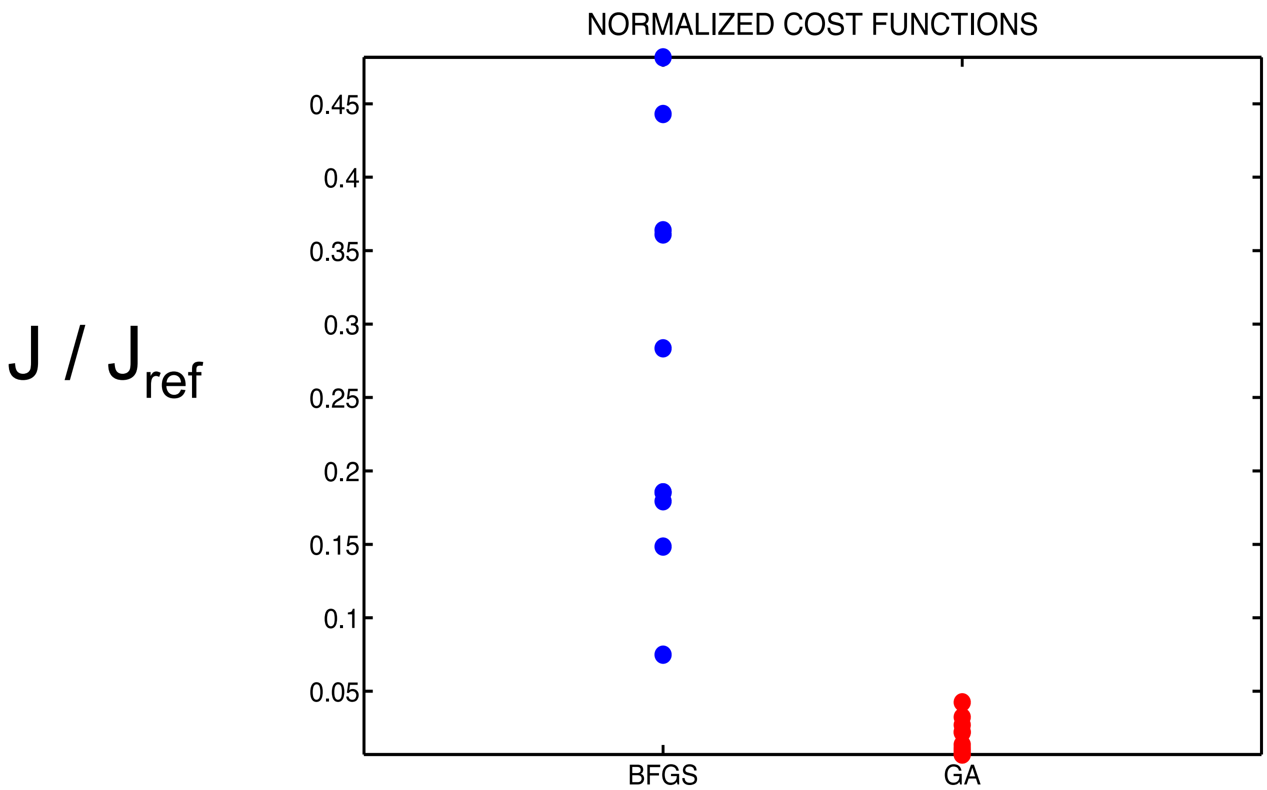


1 year of data ***4 year of data***

Varitional **Genetic Algo.**

Genetic Algo. vs Variational : performances !

➢ Random perturbation over 100% of the parameter range allowed variation



Varitional **Genetic Algo.**