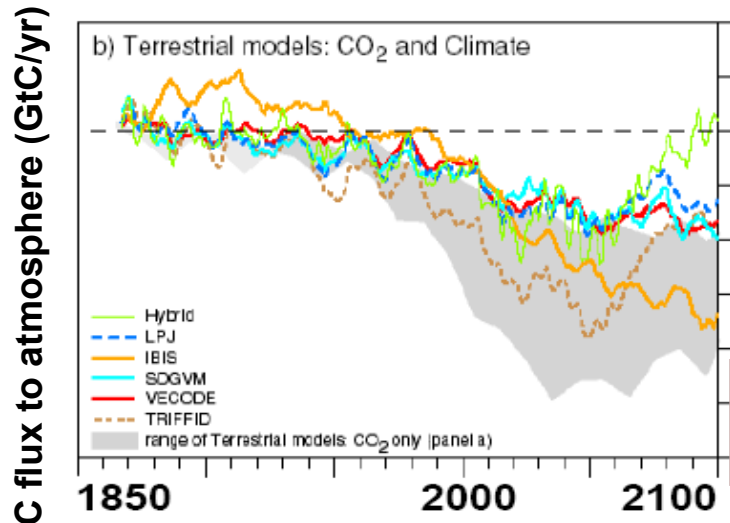


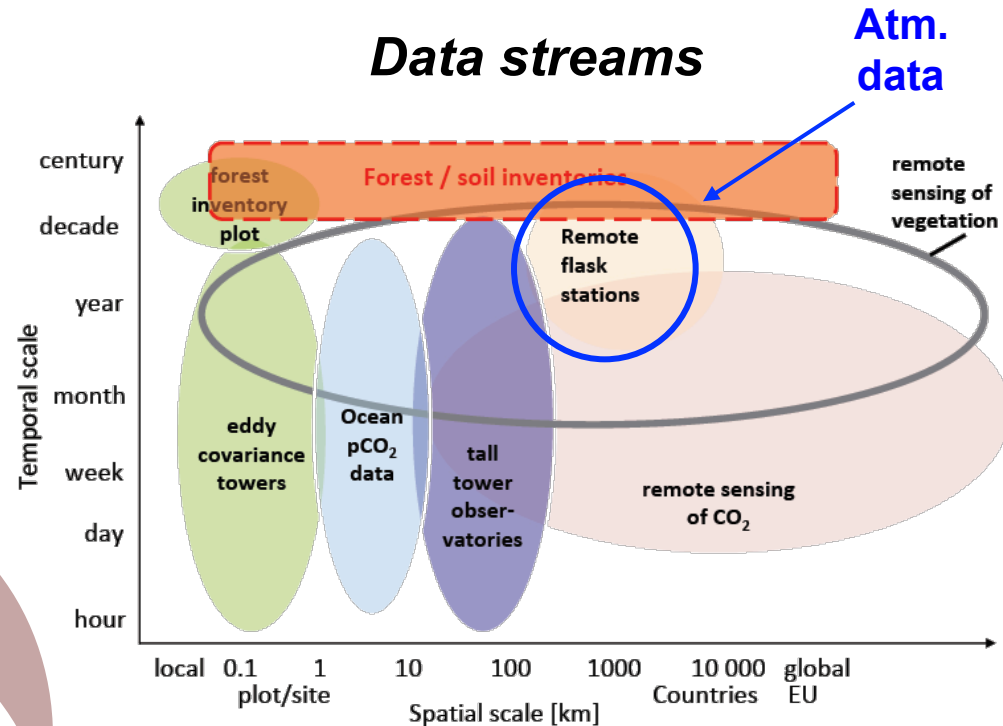
Needs for a Data Assimilation System

Large uncertainty from land to predict global C-balance (C4MIP)



OPTIMISATION OF PARAMETERS
→ Data Assimilation

Optimized ecosystem models
→ reduce the spread ?



Improve:

- Uncertainty estimates
- C land budget estimates
- Future climate predictions
- Process understanding

Main actors working on Data Assimilation



Philippe



Natasha



Cedric



Vladislav



Pascal



Catherine



Nina

A dedicated web site
<https://orchidas.lsce.ipsl.fr/>



orchidas : [Home](#)

Welcome to the ORCHIDEE Data Assimilation Systems website

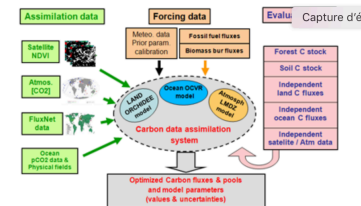
ORCHIDEE Data Assimilation Systems were designed at [IPSL/LSCE](#) in order to optimise the carbon, hydrology and energy-related parameters of the [ORCHIDEE](#) Land Surface Model, using various data sources (e.g. in situ flux measurements, satellite products, atmospheric CO₂ measurements, carbon inventory data, etc.).

The aim of the different assimilation procedures is to minimise a misfit function that measures the mismatch between

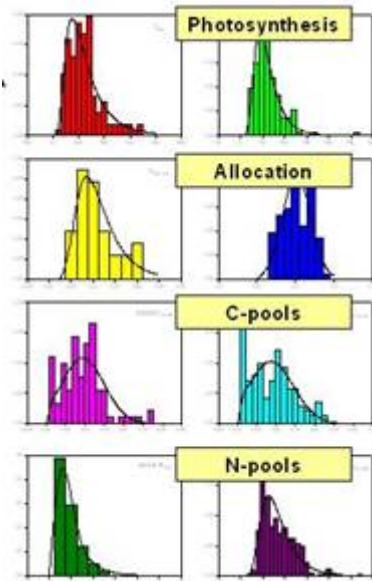
- the model outputs and the various data streams, and
- a *priori* knowledge of the parameter values

taking into account uncertainty of both components in a statistically robust framework. In this way, we combine our current understanding of the system (models) with the most up-to-date, detailed process information (observations), in order to provide the best estimate of the variables being studied.

Given this information, the [ORCHIDEE](#) Data Assimilation Systems allow the derivation of optimized posterior model parameter values and uncertainties. These



Bayesian Calibration of process-based models



Prior pdf for the parameters (ϑ)

Likelihood of the data

Bayes' Theorem

$$P(\vartheta|D) = P(\vartheta) P(D|\vartheta) / P(D)$$

Posterior pdf for the parameters

Scaling constant
($= \int P(\vartheta) P(D|\vartheta) d\vartheta$)

Formalism in the case of Gaussian errors...

Baye's theorem:
$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}).p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}$$

Assuming Gaussian Error statistics lead to minimize the cost function $J(\mathbf{x})$ to obtain the mean of $p(\mathbf{x}|\mathbf{y})$

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

\mathbf{x} : state vector ;

\mathbf{x}_b : mean prior value of state vector

\mathbf{y} : observation vector ;

\mathbf{H} : linear observation operator

\mathbf{B} / \mathbf{R} : Background / Observation error covariance matrix

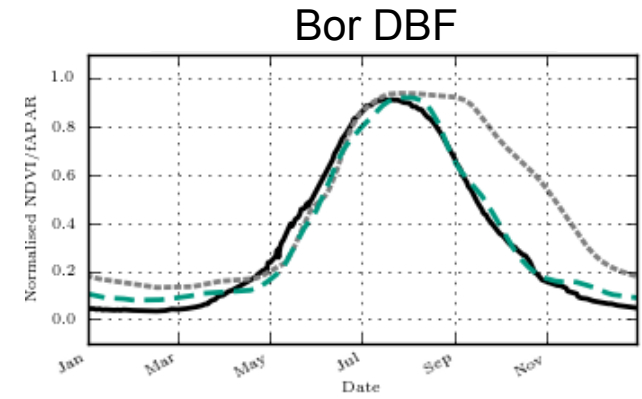
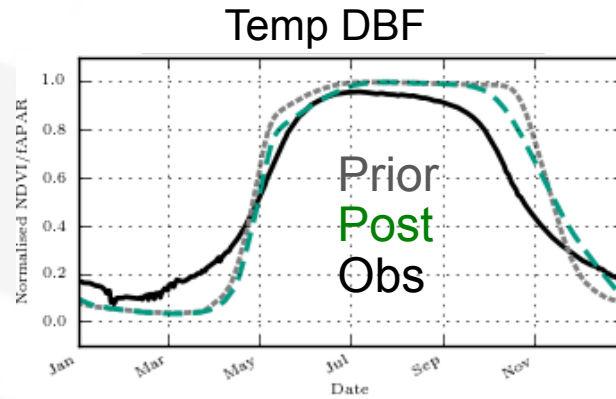
Past and Ongoing work

- **Technical / Methodological work**
 - Gradient based vs Monte-carlo optimisation method
 - Computation of Tangent Linear and Adjoint models of ORCHIDEE
 - Simultaneous vs sequential data assimilation
- **Scientific work**
 - **Carbon cycle focussed**
 - FluxNet data assimilation (several papers)
 - NDVI / Fluorescence DA
 - Multi-data streams including [CO₂]
 - **Water / Energy cycle focussed**
 - Assimilation of Surface Soil moisture
 - Surface temperature downscaling

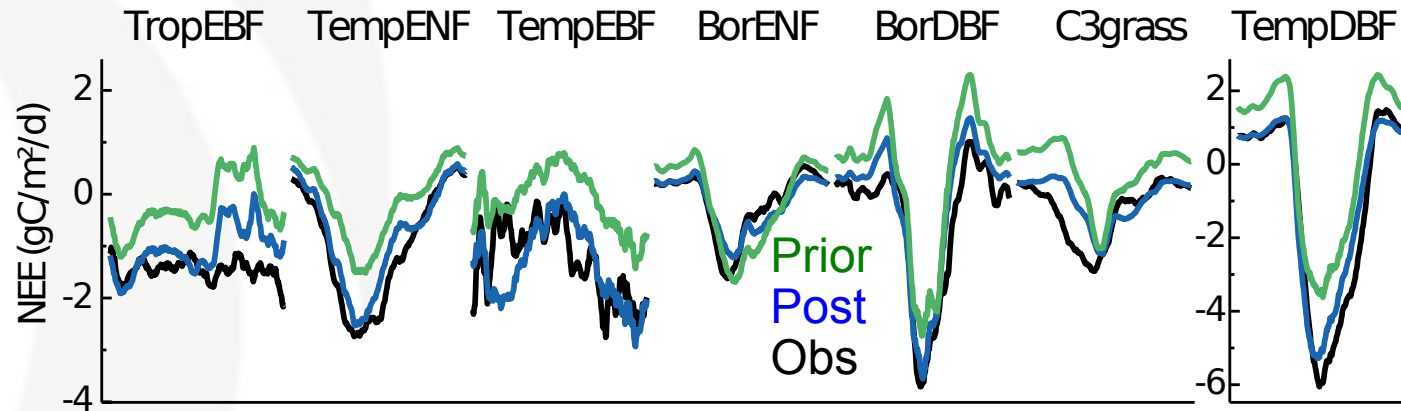


Stepwise approach (20 yr): a compromise!

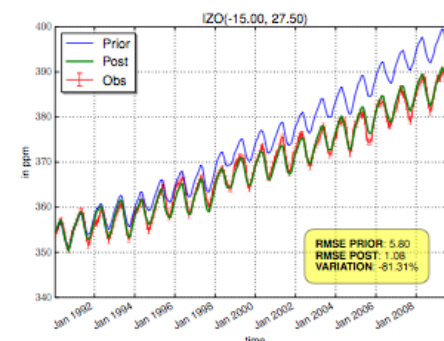
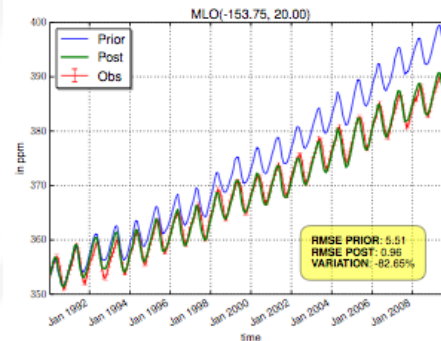
Step 1:
MODIS-NDVI
4 params /PFT



Step 2:
75 fluxnet data
≈ 20 params /PFT

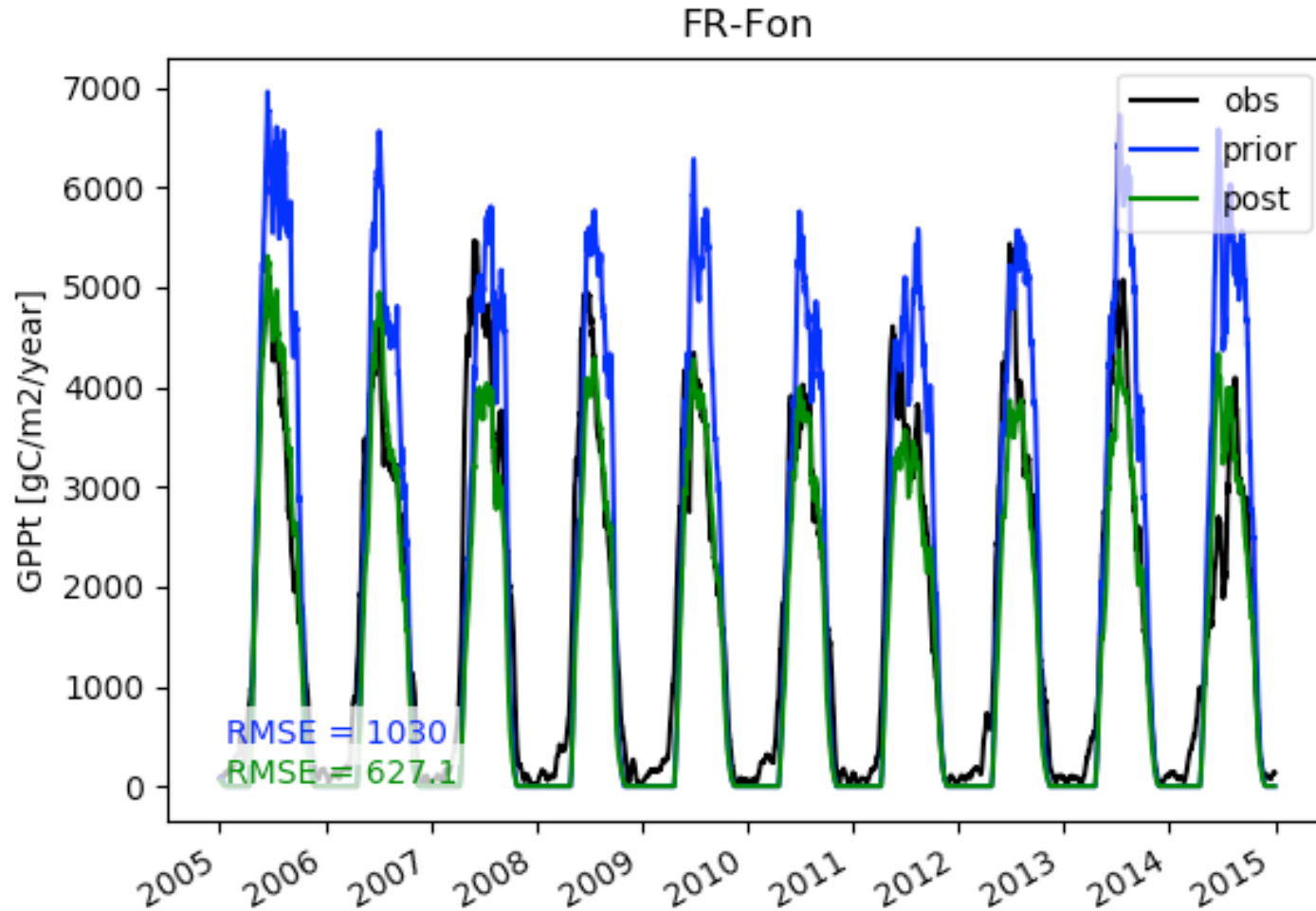


Step 3:
Atmospheric data
≈ 100 params total



Exemple with current Trunk (CN)

- Optimisation against GPP / TER at Deciduous Broadleaf sites
- Around 20 parameters are optimised (2 cases: w/wo NUE)



Exemple with current Trunk (CN)

Optimisation including Nitrogen use efficiency param

FR-Fon

Optimisation without Nitrogen use efficiency param

PARAMETER	PRIOR	MIN	MAX	POST	sigma post	sigma post / post
CN_LEAF_INIT__06	29	22	35	33.33	4.89	0.1467
CN_LEAF_MAX__06	45	36	54	48.04	7.17	0.1492
CN_LEAF_MIN__06	16	11	23	16.18	2.29	0.145
EXT_COEFF__06	0.5	0.3	1	0.8882	0.0814	0.092
EXT_COEFF_N__06	0.15	0.13	0.18	0.1586	0.0194	0.122
FCN_ROOT__06	1	0.6	1.2	0.6464	0.19	0.294
FCN_WOOD__06	0.087	0.06	0.12	0.06719	0.0213	0.317
FRAC_GROWTHRESP__06	0.28	0.2	0.36	0.2085	0.0194	0.093
K_LATOSA_MAX__06	5000	4000	6000	4101	538	0.131
K_ROOT__06	4e-07	3e-07	5e-07	3.886e-07	7.61-08	0.196
K_SAP__06	0.003	0.002	0.004	0.002086	6.59e-04	0.316
LEAFAGECRIT__06	180	90	240	118.8	4.3	0.036
LEAFFALL__06	10	8	12	11.89	1.55	0.130
NUE_OPT__06	33	23	43	24.04	1.66	0,069
RATIO_K_LATOSA__06	0.8	0.7	1	0.8506	0.119	0,1399
RECYCLE_LEAF__06	0.5	0.4	0.6	0.4964	0.0795	0,154
RECYCLE_ROOT__06	0.2	0.1	0.3	0.2386	0.0785	0,329
SLAINIT__06	0.03	0.02	0.04	0.03604	2.04e-04	0,056
SOIL_Q10	0.69	0	1.1	1.093	0.0334	0,031
VMAX_UPTAKE__1	3	2	4	3.541	0.792	0,224
VMAX_UPTAKE__2	3	2	4	2.262	0.757	0,335
COST	7304			2498		(0.3419)
COST_GPpt	3652			1355		(0.3709)
COST_rap	3652			1126		(0.3084)

PARAMETER	PRIOR	MIN	MAX	POST	sigma post	sigma post / post
CN_LEAF_INIT__06	29	22	35	33.05	4.44	0.134
CN_LEAF_MAX__06	45	36	54	45.17	7.17	0.159
CN_LEAF_MIN__06	16	11	23	22.12	2.02	0.0913
EXT_COEFF__06	0.5	0.3	1	0.3878	0.0253	0.065
EXT_COEFF_N__06	0.15	0.13	0.18	0.1686	0.0194	0.115
FCN_ROOT__06	1	0.6	1.2	0.8022	0.203	0.253
FCN_WOOD__06	0.087	0.06	0.12	0.09283	0.0234	0.252
FRAC_GROWTHRESP__06	0.28	0.2	0.36	0.2074	0.0156	0.075
K_LATOSA_MAX__06	5000	4000	6000	4050	392	0.097
K_ROOT__06	4e-07	3e-07	5e-07	4.476e-07	7.03e-08	0.157
K_SAP__06	0.003	0.002	0.004	0.002651	6.76e-04	0.255
LEAFAGECRIT__06	180	90	240	124.4	5.85	0.045
LEAFFALL__06	10	8	12	10.57	1.49	0.141
RATIO_K_LATOSA__06	0.8	0.7	1	0.8717	0.119	0.137
RECYCLE_LEAF__06	0.5	0.4	0.6	0.597	0.0798	0.134
RECYCLE_ROOT__06	0.2	0.1	0.3	0.2362	0.0794	0.336
SLAINIT__06	0.03	0.02	0.04	0.02617	2.28e-03	0.087
SOIL_Q10	0.69	0	1.1	0.9013	0.0318	0.035
VMAX_UPTAKE__1	3	2	4	3.623	0.8	0.221
VMAX_UPTAKE__2	3	2	4	2.357	0.791	0.336
COST	7304			2239		(0.3066)
COST_GPpt	3652			1140		(0.3121)
COST_rap	3652			1088		(0.2978)

10
2

Some take home messages

- Model development should include to a certain extent parameter calibration (DA)
- Definition of prior parameter physical range / uncertainties is crucial
- Be careful about equifinality / overfitting / ...
(see next presentation by Natasha on multi-constraints)
- Adjoint model of ORCHIDEE will help doing DA