Assimilation of atmospheric observations into terrestrial models

T. Kaminski[1], M. Scholze[2], P. Rayner[3], W. Knorr[2], R. Giering[1], H. Widmann[4], and M. Heimann[4]

Copy of presentation at http://CCDAS.org

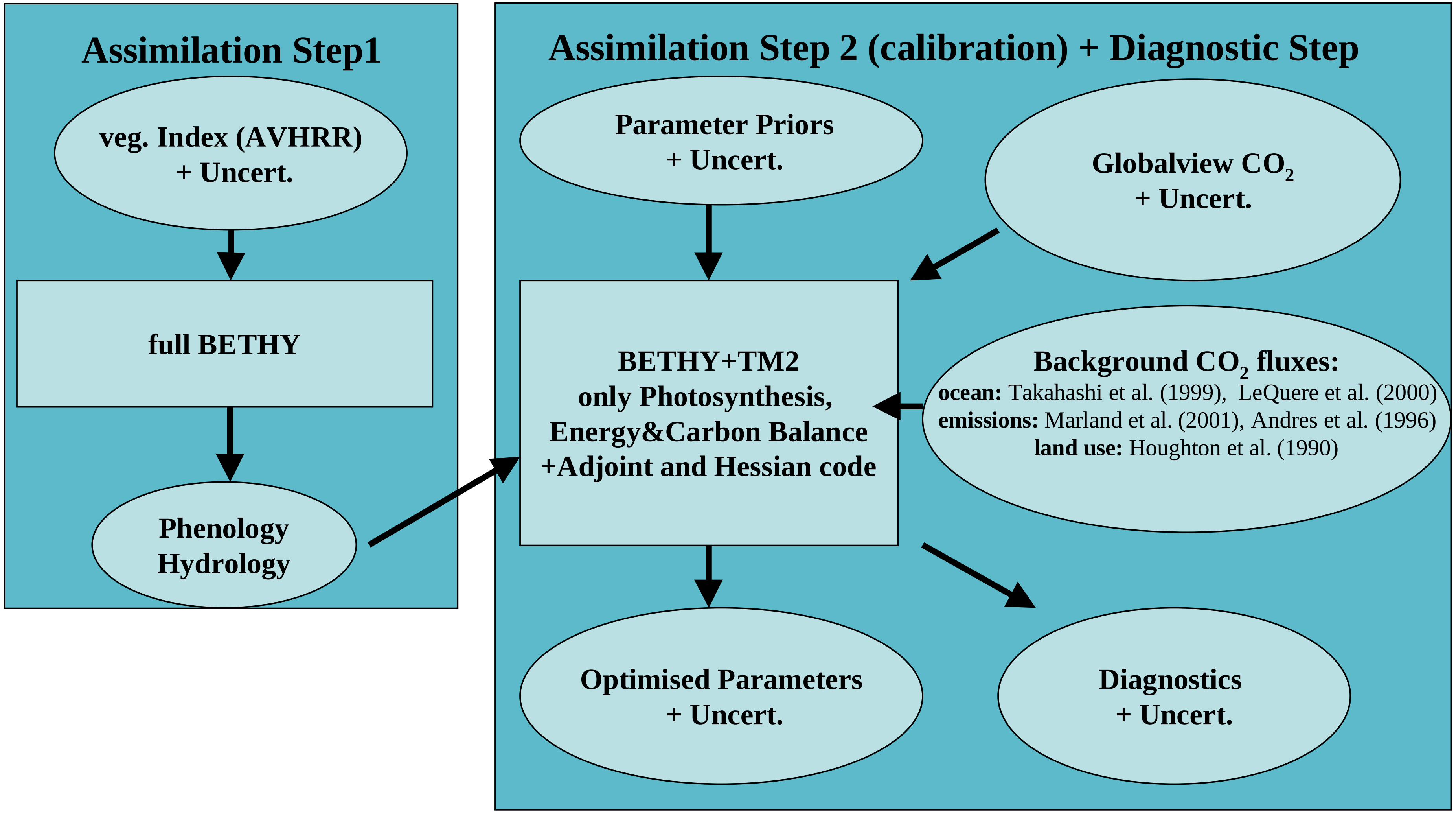
1 FastOpt 2 Quantifying and Understanding the Earth System 3 4 Max-Planck-Institut für Biogeochemie

CCDAS

Outline

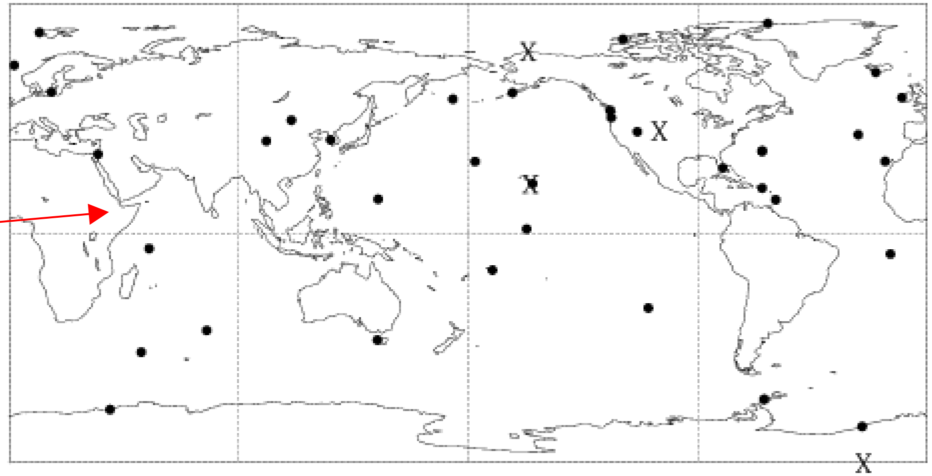
- Methodology, example: CCDAS
- Possible extensions of concept
- Some ongoing activities
- Conclusions

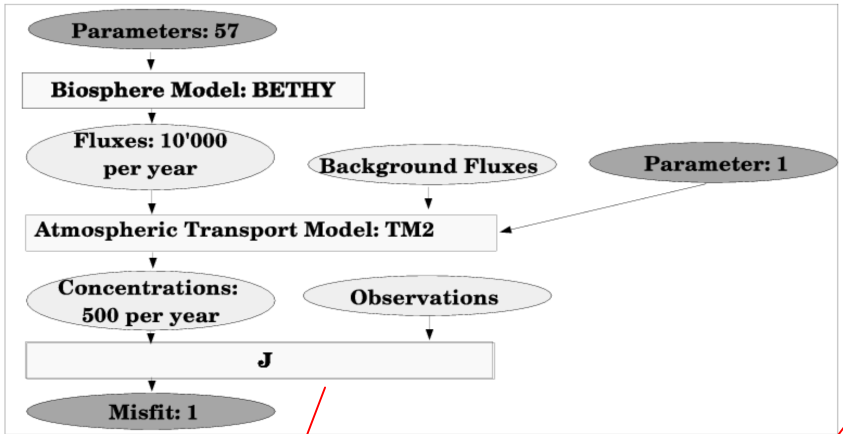
Carbon Cycle Data Assimilation System (CCDAS) current form



CCDAS calibration step

- Terrestrial biosphere model BETHY (Knorr 97) delivers CO2 fluxes to atmosphere
- Uncertainty in process parameters from laboratory measurements
- Global atmospheric network provides additional constraint





covariance of uncertainty in priors for parameters

covariance of uncertainty in measurements + model

priors for parameters

observed concentrations

$$J(m)=\frac{1}{2}(m-m_0)^T C_m^{-1}(m-m_0)+\frac{1}{2}(c(m)-d)^T C_d^{-1}(c(m)-d)$$

FastOpt | Quantifying and Understanding the Earth System | Max-Planck-Institut für Biogeochemie | CCDAS

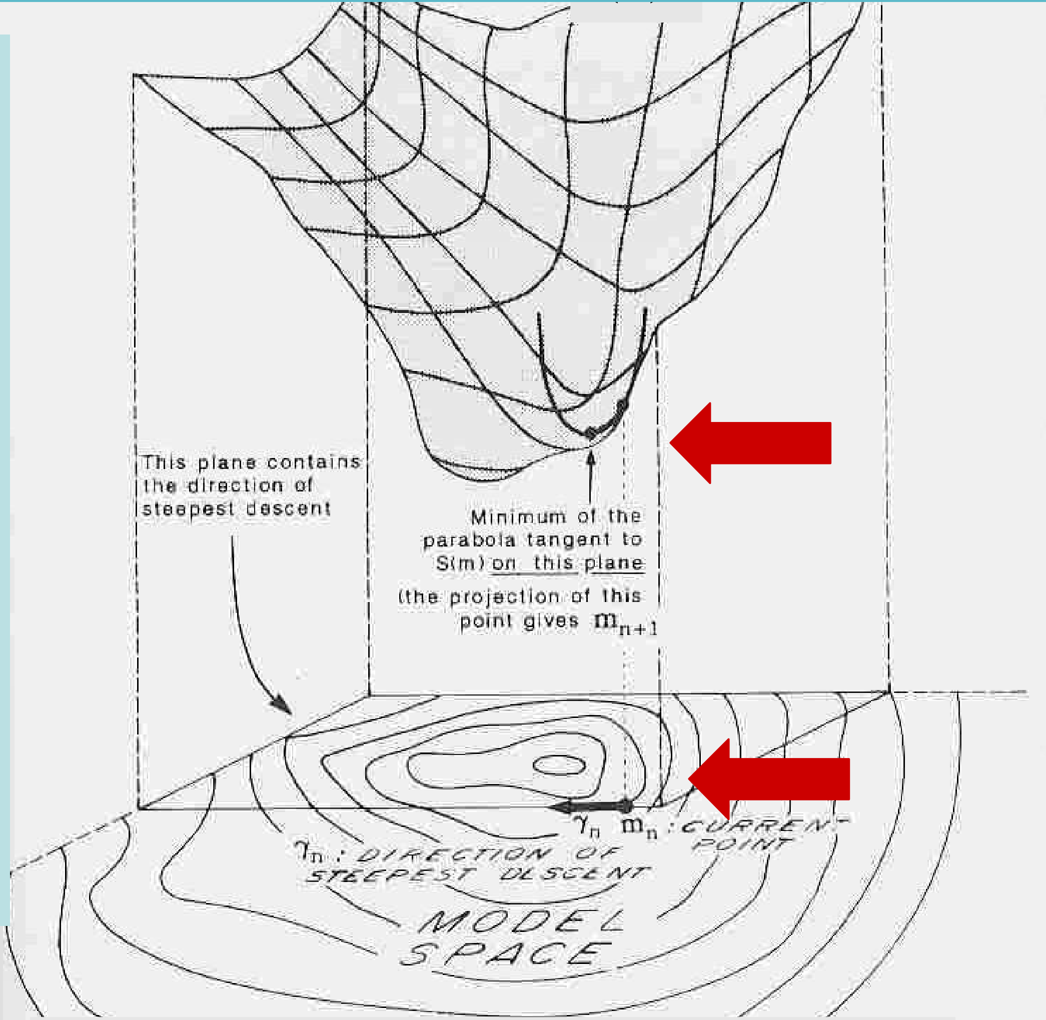
Minimisation and Parameter-Uncertainties

Gradient of *J*(*m*) provides search directions for minimisation.

Second Derivative (Hessian) of *J*(*m*) yields curvature of *J*, provides estimated uncertainty in m_{opt}

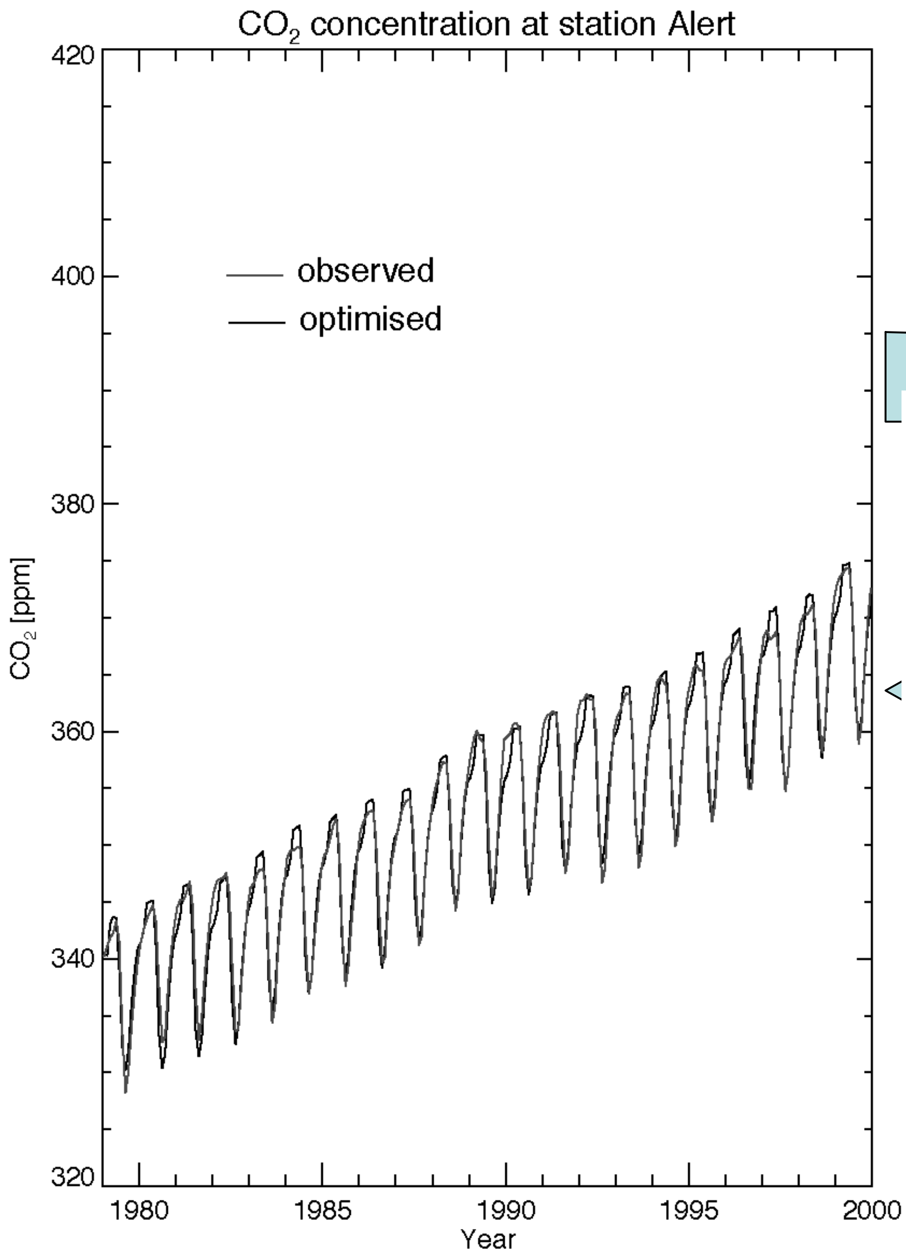
$$\mathbf{C}_m \approx \left\{ \frac{\partial^2 J(m_{opt})}{\partial m_{i,j}^2} \right\}^{-1}$$

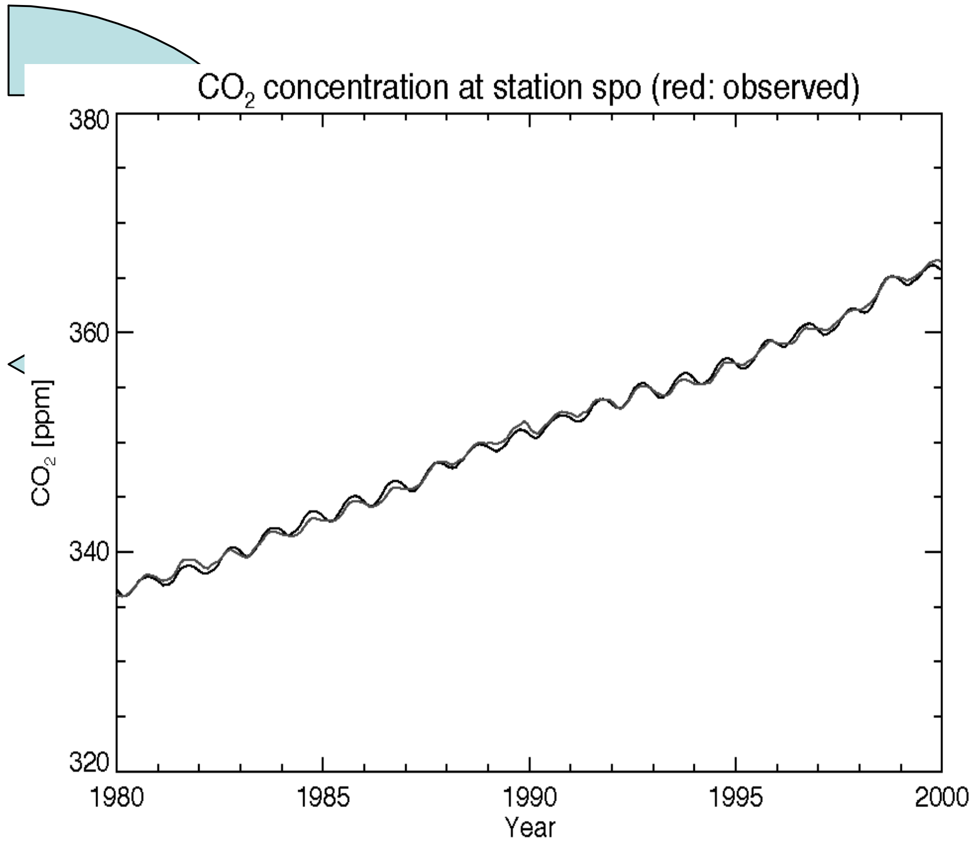
Figure taken from **Tarantola** '87



Space of *m* (model parameters)

Optimisation
(BFGS+ adjoint gradient)





Posterior uncertainties on parameters

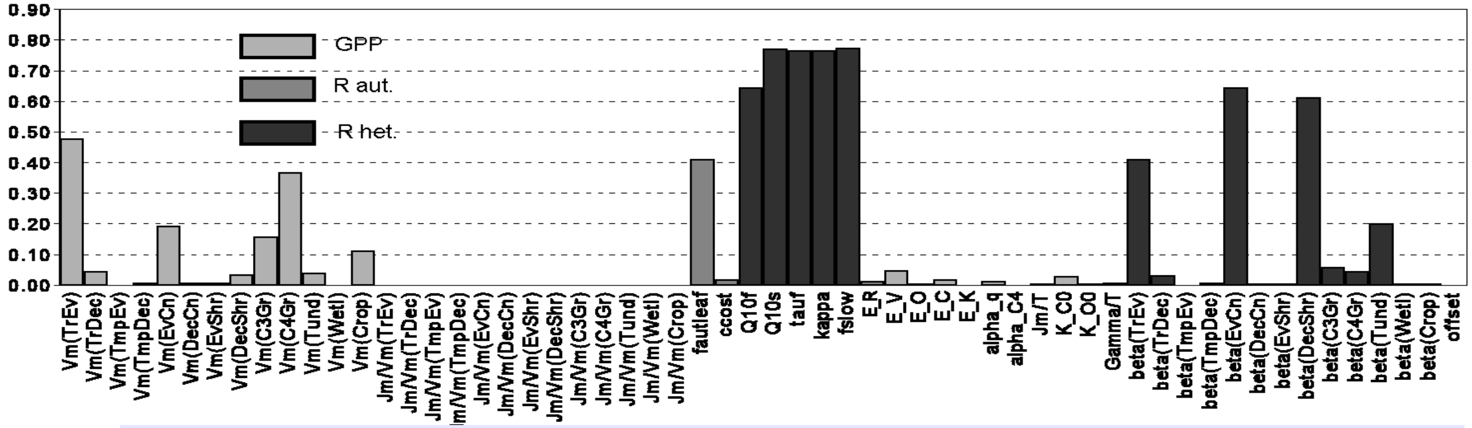
Use inverse Hessian of objective function to approximate posterior uncertainties

$$C_m \approx \left(\frac{\partial^2 J(m_{opt})}{\partial m_{i,j}^2} \right)^{-1}$$

examples:

	first guess	optimized	prior unc.	opt.unc.	Vm(TrEv)	Vm(EvCn)	Vm(C3Gr)	Vm(Crop)
	$\mu mol/m^2 s$	$\mu mol/m^2 s$	%	%	error covariance			
Vm(TrEv)	60.0	43.2	20.0	10.5	**0.28**	0.02	-0.02	0.05
Vm(EvCn)	29.0	32.6	20.0	16.2	0.02	**0.65**	-0.10	0.08
Vm(C3Gr)	42.0	18.0	20.0	16.9	-0.02	-0.10	**0.71**	-0.31
Vm(Crop)	117.0	45.4	20.0	17.8	0.05	0.08	-0.31	**0.80**

Relative reduction of uncertainties

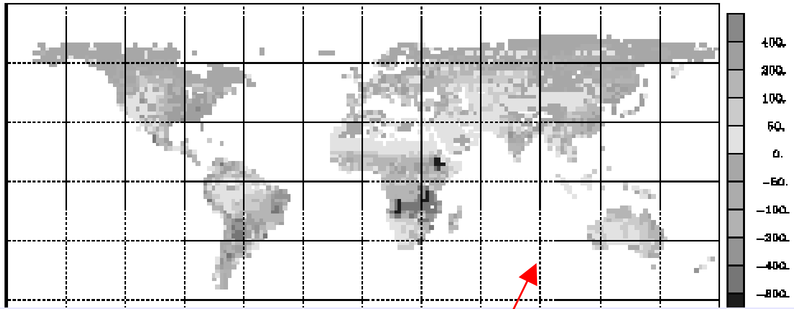


Observations resolve about 10-15 directions in parameter space

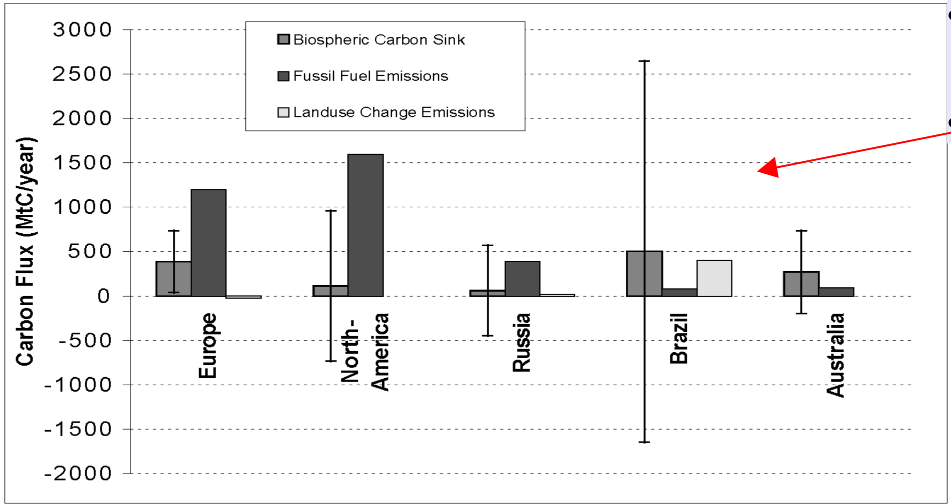
CCDAS diagnostic step

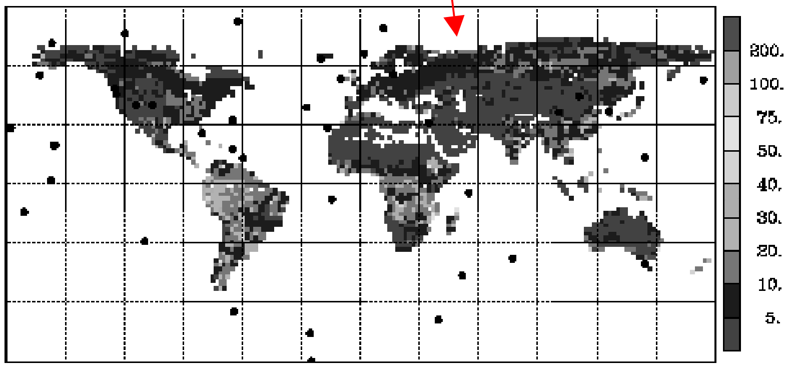
Global fluxes and their uncertainties

$$C_y = \left(\frac{\partial y_i(m_{opt})}{\partial m_j} \right) C_m \left(\frac{\partial y_i(m_{opt})}{\partial m_j} \right)^T$$



- Examples for diagnostics:
- Long term mean fluxes to atmosphere (gC/m2/year) and uncertainties
- Regional means





FastOpt | QUEST Quantifying and Understanding the Earth System | Max-Planck-Institut für Biogeochemie | CCDAS | CAMELS

CCDAS
methodological aspects

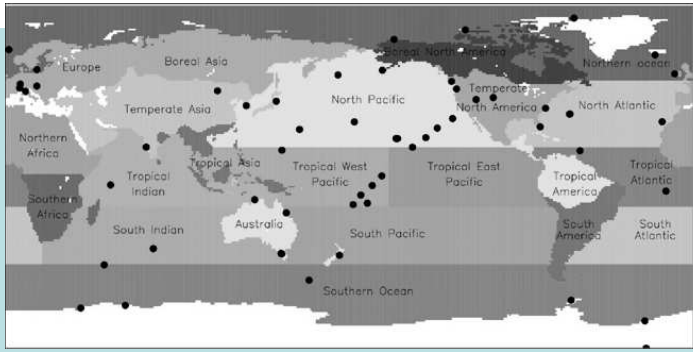
- remarks:
  - CCDAS tests a given combination of observational data plus model formulation with uncertain parameters
  - CCDAS delivers optimal parameters, diagnostics/prognostics, and their a posteriori uncertainties
  - all derivative code (adjoint, Hessian, Jacobian) generated automatically from model code by compiler tool TAF: quick updates of CCDAS after change of model formulation
  - derivative code is highly efficient
  - CCDAS posterior flux field consistent with trajectory of process model rather than linear combination of prescribed flux patterns (as transport inversion)
  - CCDAS includes a prognostic mode (unlike transport inversion)
- some of the difficulties/problems:
  - Prognostic uncertainty (error bars) only reflect parameter uncertainty What about uncertainty in model formulation, driving fields…?
  - Uncertainty propagation only for means and covariances (specific PDFs), and only with a linearised model
  - Result depends on a priori information on parameters
  - Result depends on a single model
  - Two step assimilation procedure sub optimal
  - lots of other technical issues (bounds on parameters, driving data, Eigenvalues of Hessian ...)

Extensions of concept

1. More processes/components

- **Have tested a version extended by an extremely simplified form of an ocean model:**

  flux(x,t) = Σ coefficient(i) * pattern(i,x,t)



- **Optimising coefficients for biosphere patterns would allow the optimisation to compensate for errors (e.g. missing processes) in biosphere model (weak constraint 4DVar, see ,e.g., Zupanski (1993))**
- **But it is always preferable to include a process model, e.g for fire, marine biogeochemistry**
- **Can also extend to weak constraint formulation for state of biosphere model: include state as unknown with prior uncertainty estimated from model error**

Extension of concept

2. Adding more observations

Atmospheric Concentrations (could also be column integrated)

$$J(m) = \tfrac{1}{2}\,(m-m_0)^T\, C_m^{-1}(m-m_0)$$

$$+\ \tfrac{1}{2}\,(c_{mod}(m)-c_{obs})^T\, C_c^{-1}(c_{mod}(m)-c_{obs})$$

$$+\ \tfrac{1}{2}\,(f_{mod}(m)-f_{obs})^T\, C_f^{-1}(f_{mod}(m)-f_{obs})$$

$$+\ \tfrac{1}{2}\,(I_{mod}(m)-I_{obs})^T\, C_I^{-1}(I_{mod}(m)-I_{obs})$$

$$+\ \tfrac{1}{2}\,(R_{mod}(m)-R_{obs})^T\, C_R^{-1}(R_{mod}(m)-R_{obs})$$

$$+\ \text{etc} \ldots$$

Flux Data

Inventories

Atmospheric Isotope Ratios

- Can add further constraints on any quantity that can be extracted from the model (possibly after extensions/modifications of model)
- Covariance matrices are crucial: Determine relative weights!
- Uses Gaussian assumption; can also use logarithm of quantity (lognormal distribution), ...

Extensions of concept

3. Observation-System Design

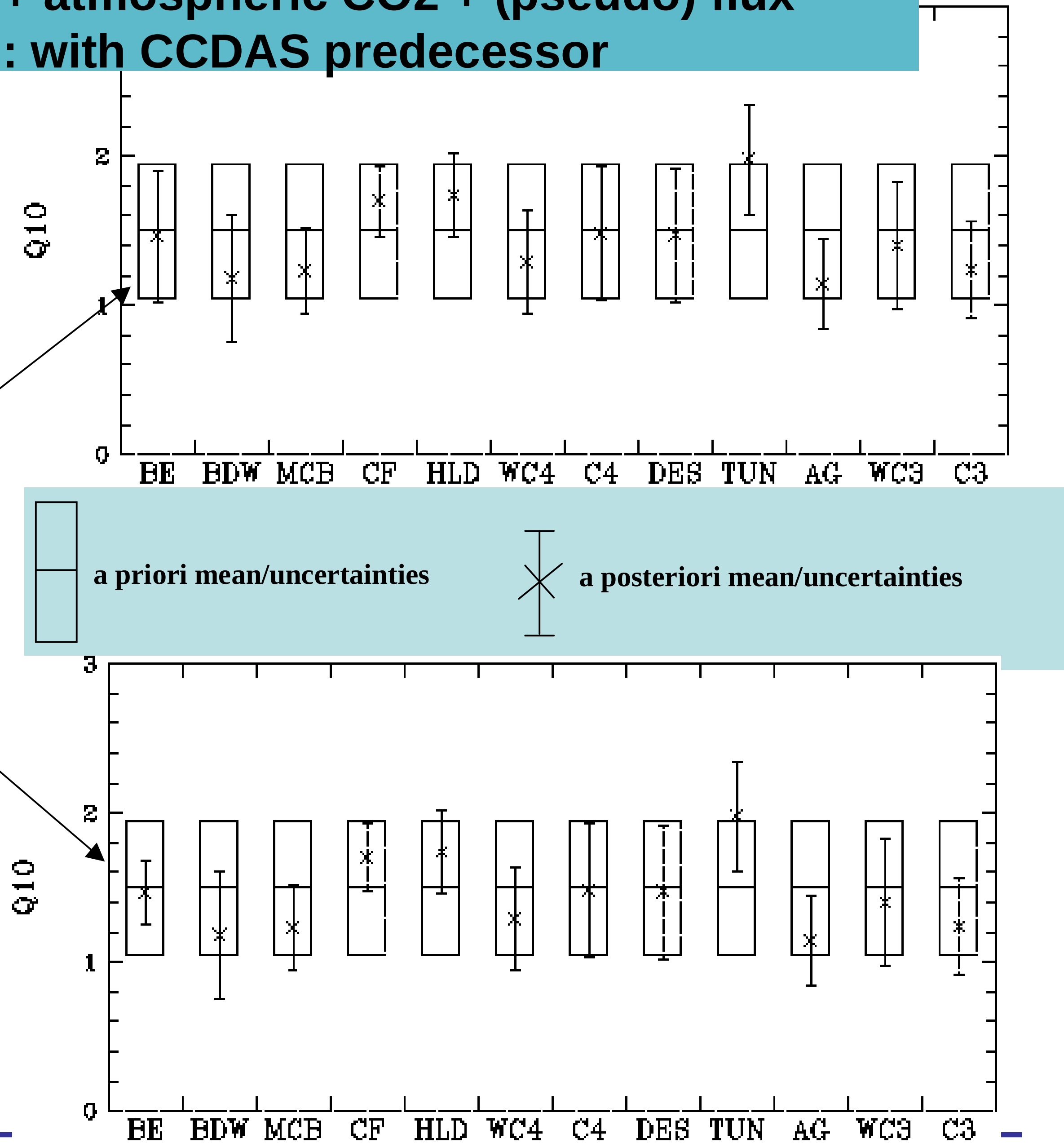
- **Background: For a linear model, only uncertainties of potential observations are required to infer posterior uncertainties on unknowns**
- **Used by *Rayner et al.* (1996) for design of atmospheric sampling network**
- **Same strategy can be used in CCDAS**
  - **Define target quantity (e.g. European balance for 1990 or 2010)**
  - **Repeat following procedure for each candidate observing system:**
    - **Prescribe uncertainty of potential observations**
    - **Prescribe prior parameter uncertainties**
    - **Backpropagate uncertainties to parameter uncertainties (like pseudo flux measurement for SDBM)**
    - **Propagate parameter uncertainties forward to target quantity**
- **Best candidate (lowest posterior uncertainty on target quantity) depends on:**
  - **Target quantity**
  - **Process models**
  - **prescribed uncertainties**
- **Candidate observing systems may be generated automatically by an optimisation algorithm, as in *Rayner et al.* (1996)**

Example: A priori info + atmospheric CO2 + (pseudo) flux measurement : with CCDAS predecessor

Comparison shows impact of a (pseudo) flux measurement in the broadleaf evergreen biome on Q10 estimated by an inversion of SDBM:

***Upper panel*: only concentration data**

***Lower panel*: concentration data + pseudo flux measurement (mean: as predicted sigma: 10gC/m^2/year)**



Extensions of concept

4. Different scale

- Can be applied to scales other than global, e.g. continental.
  May need to alter/improve process representation
  (e.g. atmospheric transport)
  Need to handle domain boundaries somehow

- Can run for longer time scales.
  For efficient handling of slow pool spin-up
  see Kaminski et al., LNCS, 2005

Some ongoing Activities

- CCDAS sister at CSIRO DAR (Pak, Wang + coworkers)
  - built around CABLE model
  - currently runs at site scale and assimilates eddy flux data
  - extension by transport model planned

- CCDAS
  - Prognose biospheric uptake in 2050 (Rayner et al. Boulder abstract)
  - Include fire model (Scholze et al. Boulder abstract)
    Burnt area model (METEOSAT Albedo product, Goverts et al. 2002) +
    Emissions model
  - Derive prior uncertainties from inversions against leaf and site levels (J. Kattge)
  - Support design of ocean observation system (FastOpt activity for CarboOcean, C. Heinze)
    Deliver (second order adjoint) sensitivity maps of European budget w.r.t. ocean flux.
  - IMECC proposal (co-ordinated by P. Rayner) includes construction of network design tool

Conclusions

- **Concept demonstrated by CCDAS and its sister in Australia**
- **Concept can be extended**
  - **to assimilate additional types of observations, e.g column CO2, burnt area …**
  - **to quantify their impact on interesting quantities,
e.g. past of future continental scale budgets**
  - **to support observation system design**
- **Concept looks well-suited as a tool for integration**
- **More info, papers, etc:
http://CCDAS.org, http://FastOpt.com**