ESA Summer School, Frascati, August 2004

Data Assimilation for global CO_2 Inversions

Wolfgang Knorr
Max-Planck Institute for Biogeochemistry, Jena

Programme

- **Minimizing the cost function**
- **Uncertainties of Parameters**
- **Uncertainties of Diagnostics**
- **Global application: The Carbon Cycle Data Assimilation System (CCDAS)**

The Model

model diagnostics

$$f : \vec{m}(\vec{x}) \rightarrow \vec{y}(\vec{x}, t)$$

model parameters

a (non-linear) function from a vector space of (time-independent) parameters to a vector space of (time and space dependent) diagnostics

note: in this example, parameters are varied globally

The Cost Function

model diagnostics

measurements

$$J(\vec{m}) = \frac{1}{2}[\vec{m} - \vec{m}_0]\mathbf{C}^{\mathbf{-1}}_{m0}[\vec{m} - \vec{m}_0]^T + \frac{1}{2}[\vec{y}(\vec{m}) - \vec{y}_0]\mathbf{C}^{\mathbf{-1}}_{y}[\vec{y}(\vec{m}) - \vec{y}_0]^T$$

current values of
model parameters

a priori
parameter values

a priori error covariance matrix
of parameters

error covariance matrix
of measurements

Finding the Minimum

First derivative (Gradient) of *J*(*m*) w.r.t. *m* (model parameters) :

–?*J*(*m*)/?*m*

yields direction of steepest descent

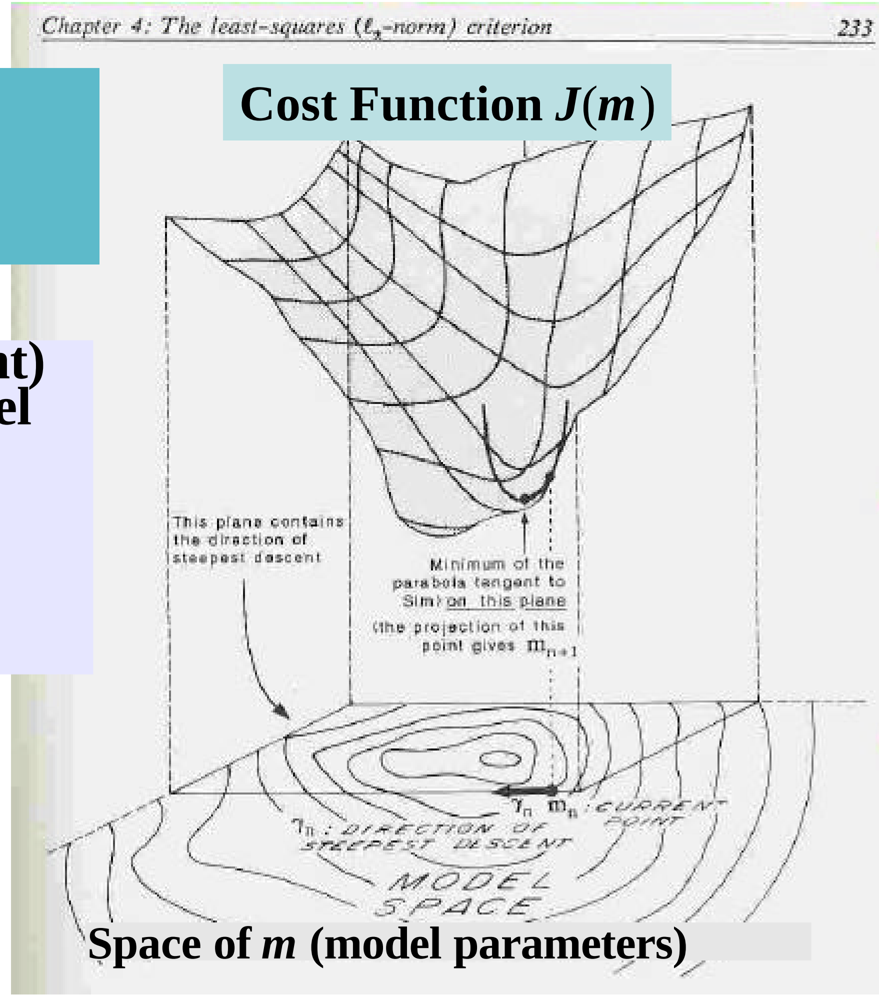


Figure taken from **Tarantola** '87

Programme

- Minimizing the cost function
- **Uncertainties of Parameters**
- Uncertainties of Diagnostics
- Global application: The Carbon Cycle Data Assimilation System (CCDAS)

Uncertainties in Parameters

Cost function:

$$J(\vec{m}) = \frac{1}{2}[\vec{m} - \vec{m}_0]\mathbf{C}^{-1}_{m\,0}[\vec{m} - \vec{m}_0]^T + \frac{1}{2}[\vec{y}(\vec{m}) - \vec{y}_0]\mathbf{C}^{-1}_{y}[\vec{y}(\vec{m}) - \vec{y}_0]^T$$

Taylor expansion around minimum:

$$J(\vec{m}) \approx J(\vec{m}_{opt}) + \underbrace{\left(\frac{\partial J(\vec{m}_{opt})}{\partial \vec{m}}\right)}_{=\,0}[\vec{m} - \vec{m}_{opt}]^T + \frac{1}{2}[\vec{m} - \vec{m}_{opt}]\underbrace{\left(\frac{\partial^2 J(\vec{m}_{opt})}{\partial \vec{m}^2}\right)}[\vec{m} - \vec{m}_{opt}]^T$$

curvature of cost function around optimimum
= inverse of *posterior* error covariance of parameters

Error Covariances of Parameters

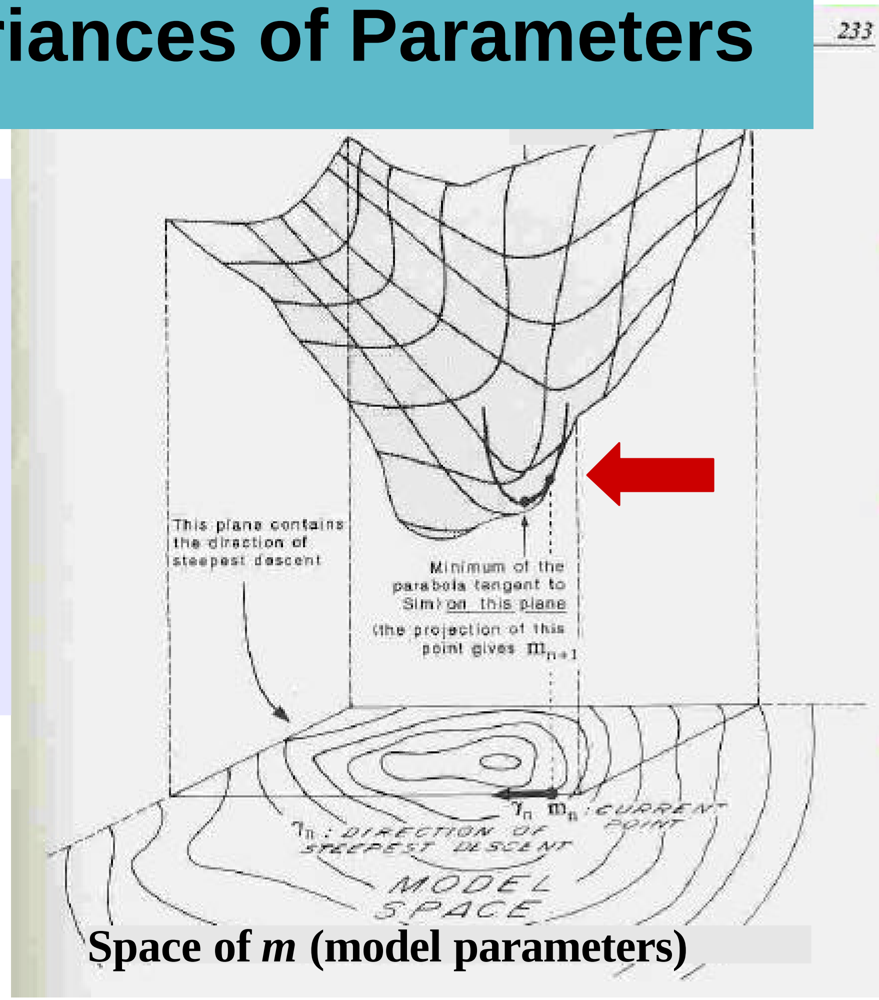
Second Derivative (Hessian) of *J*(*m*):

?^{2}*J*(*m*)/?*m*^{2}

yields curvature of *J*, provides estimated uncertainty in m_{opt}

$$\mathbf{C}_m = \left\{ \frac{\P^2 J}{\P m_{i,j}^2} \right\}^{-1}$$

= inverse Hessian



Space of *m* (model parameters)

Programme

- Minimizing the cost function
- Uncertainties of Parameters
- **Uncertainties of Diagnostics**
- Global application: The Carbon Cycle Data Assimilation System (CCDAS)

Uncertainties in Diagnostics

linear projection from parameters to diagnostics:

$$\vec{y}(\vec{m}) \approx \vec{y}(\vec{m}_{opt}) + \left(\frac{\partial \vec{y}(\vec{m}_{opt})}{\partial \vec{m}} \right) \left[\vec{m} - \vec{m}_{opt} \right]^T$$

Error Covariances of Diagnostics

Error covariance of diagnostics, *y*, after optimisation:

linearized model

$$\mathbf{C}_y = \left(\frac{\partial \vec{y}(\vec{m}_{opt})}{\partial \vec{m}} \right) \mathbf{C}_m \left(\frac{\partial \vec{y}(\vec{m}_{opt})}{\partial \vec{m}} \right)^T$$

error covariance of parameters

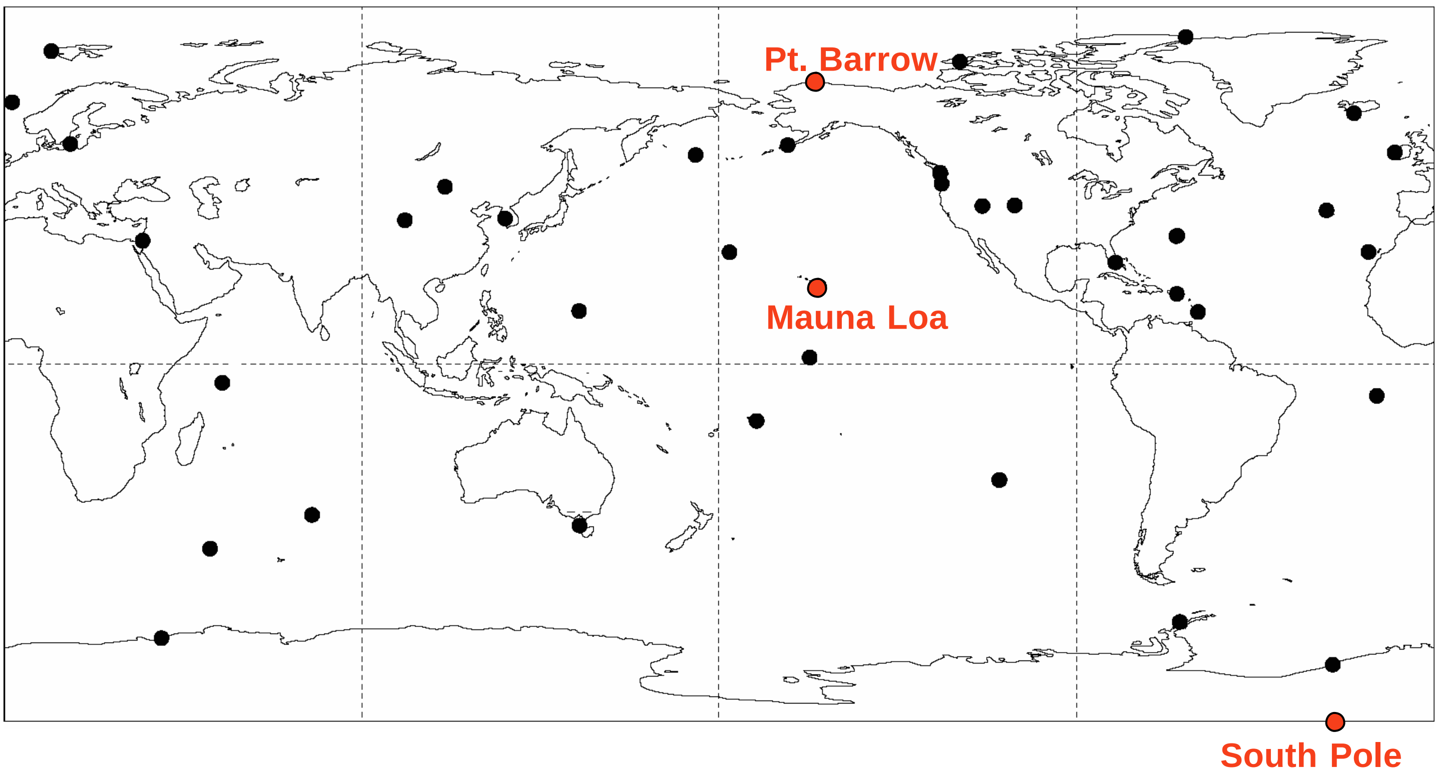
note: think of Alan's second slide above your bed!

End of the Maths Session!

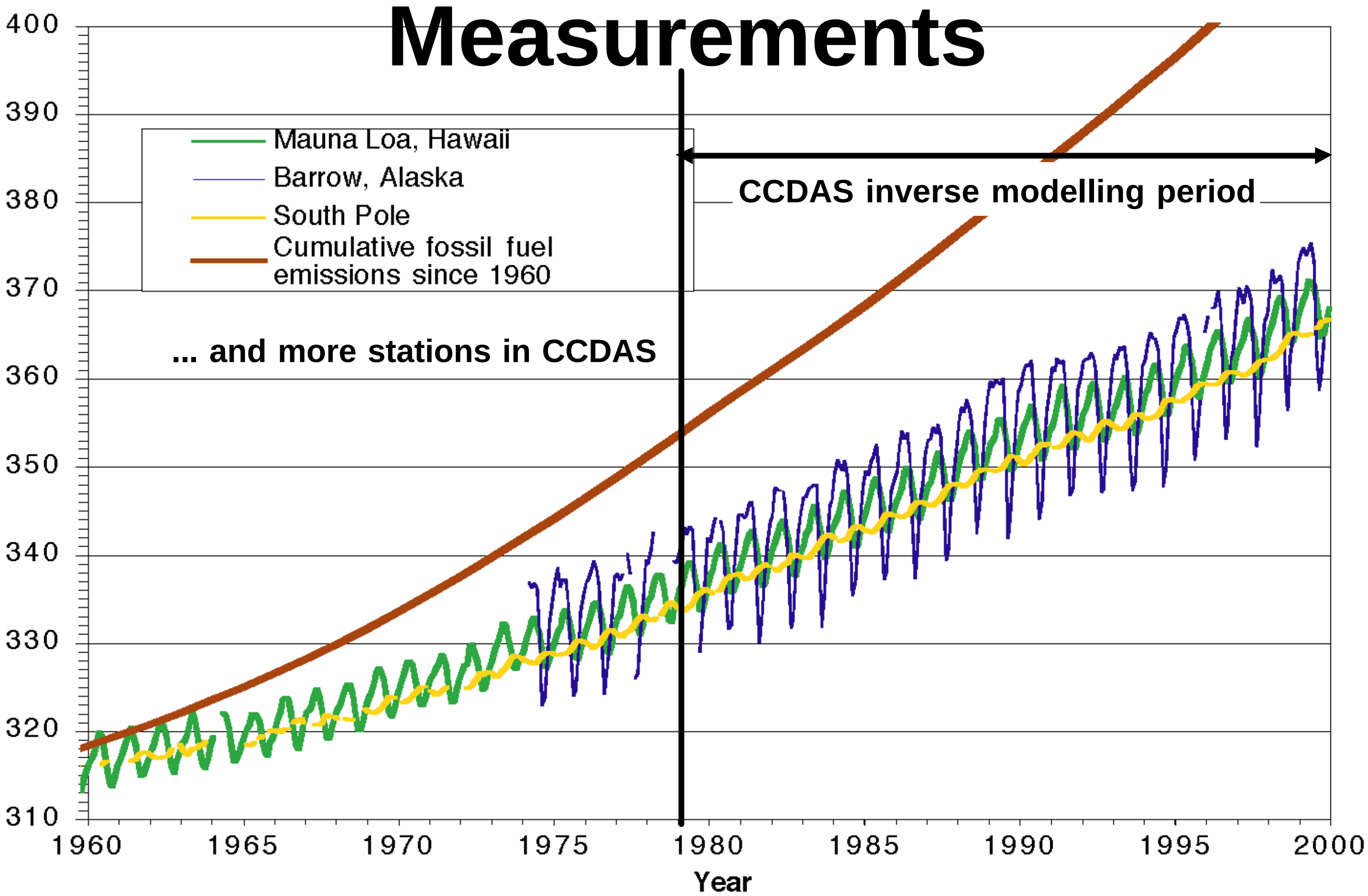
Programme

- Minimizing the cost function
- Uncertainties of Parameters
- Uncertainties of Diagnostics
- **Global application: The Carbon Cycle Data Assimilation System (CCDAS)**

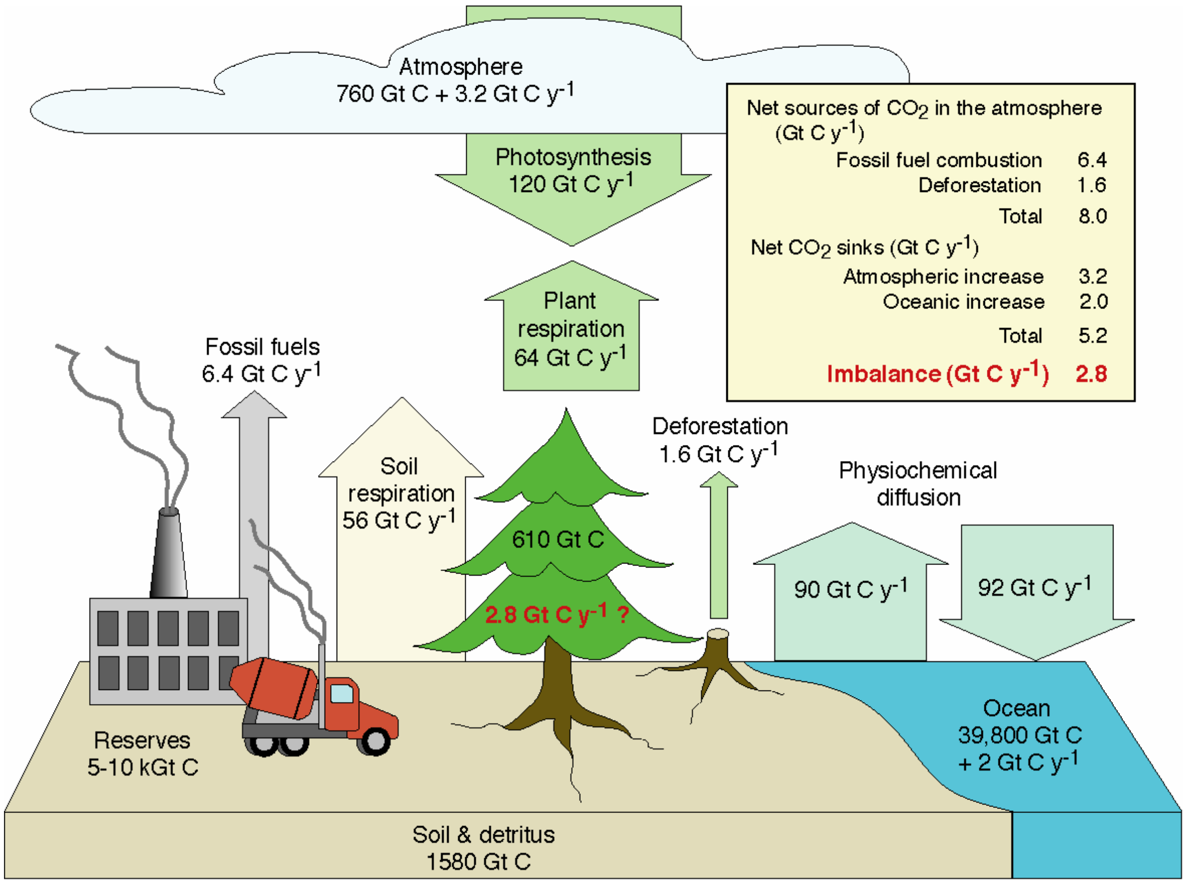
The Station Network



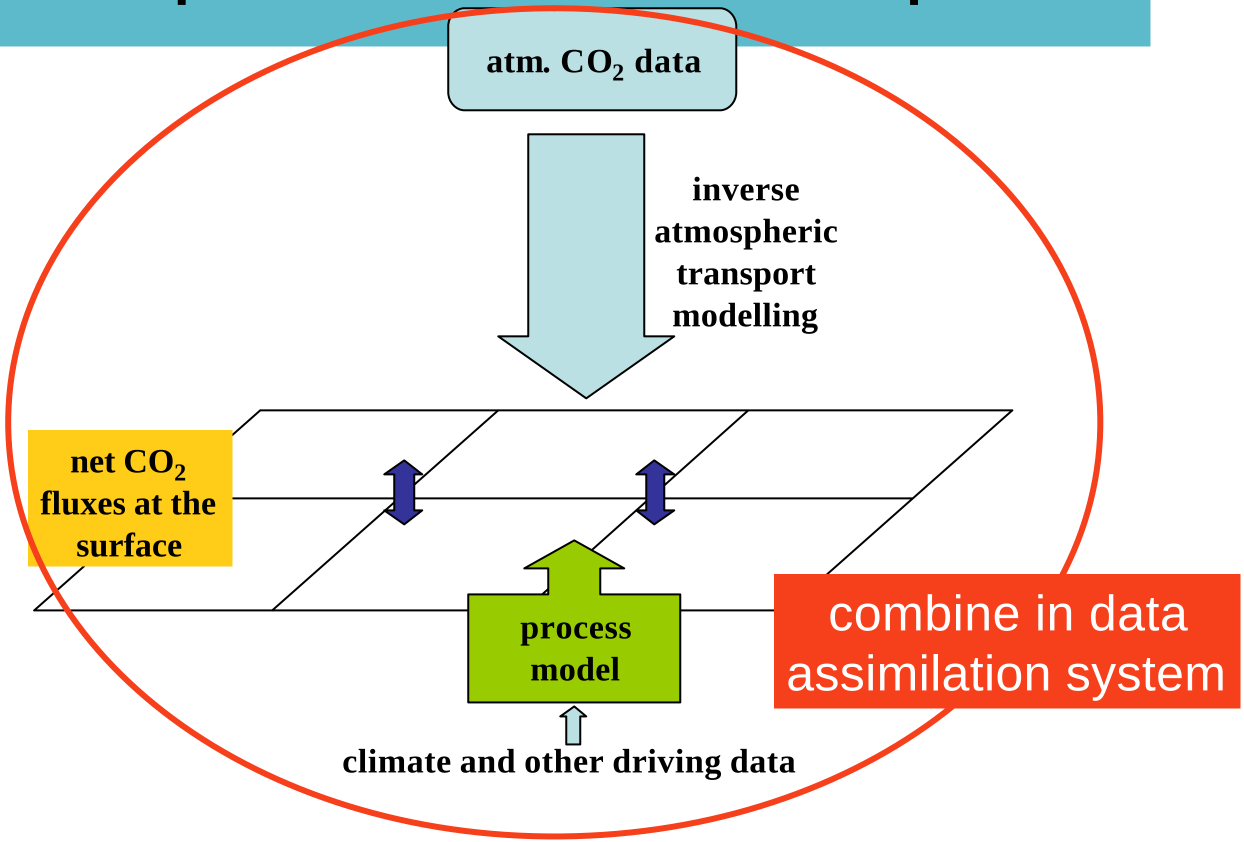
Atmospheric CO_2 Measurements



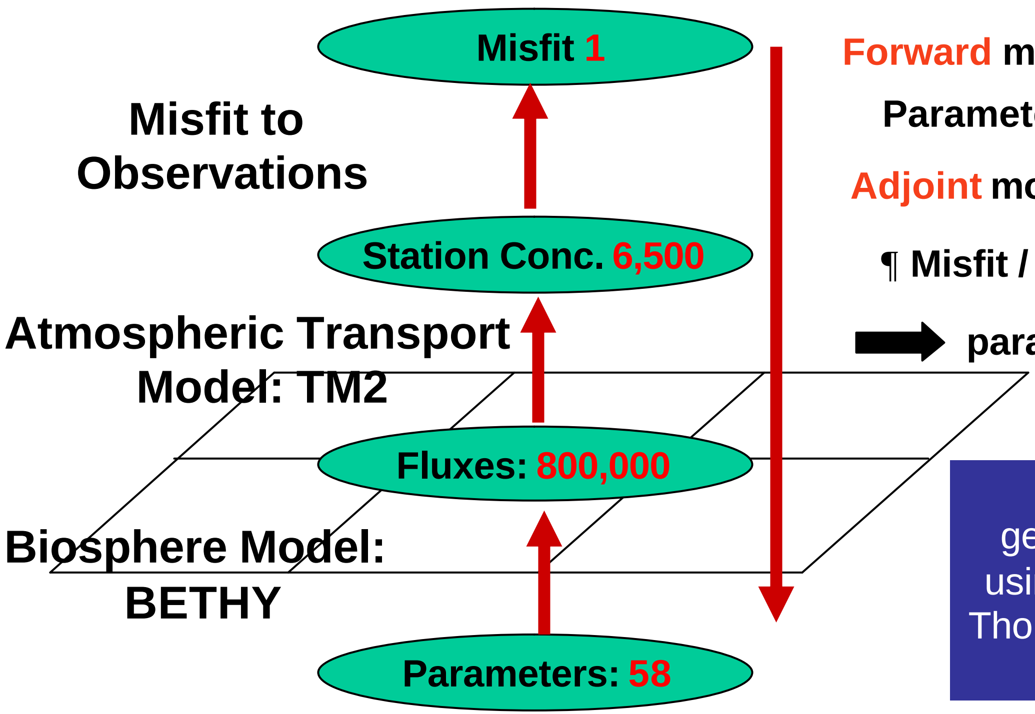
The Global Carbon Cycle



Top down / Bottom up



Carbon Cycle Data Assimilation Systems (CCDAS)



Forward modelling:

Parameters –> Misfit

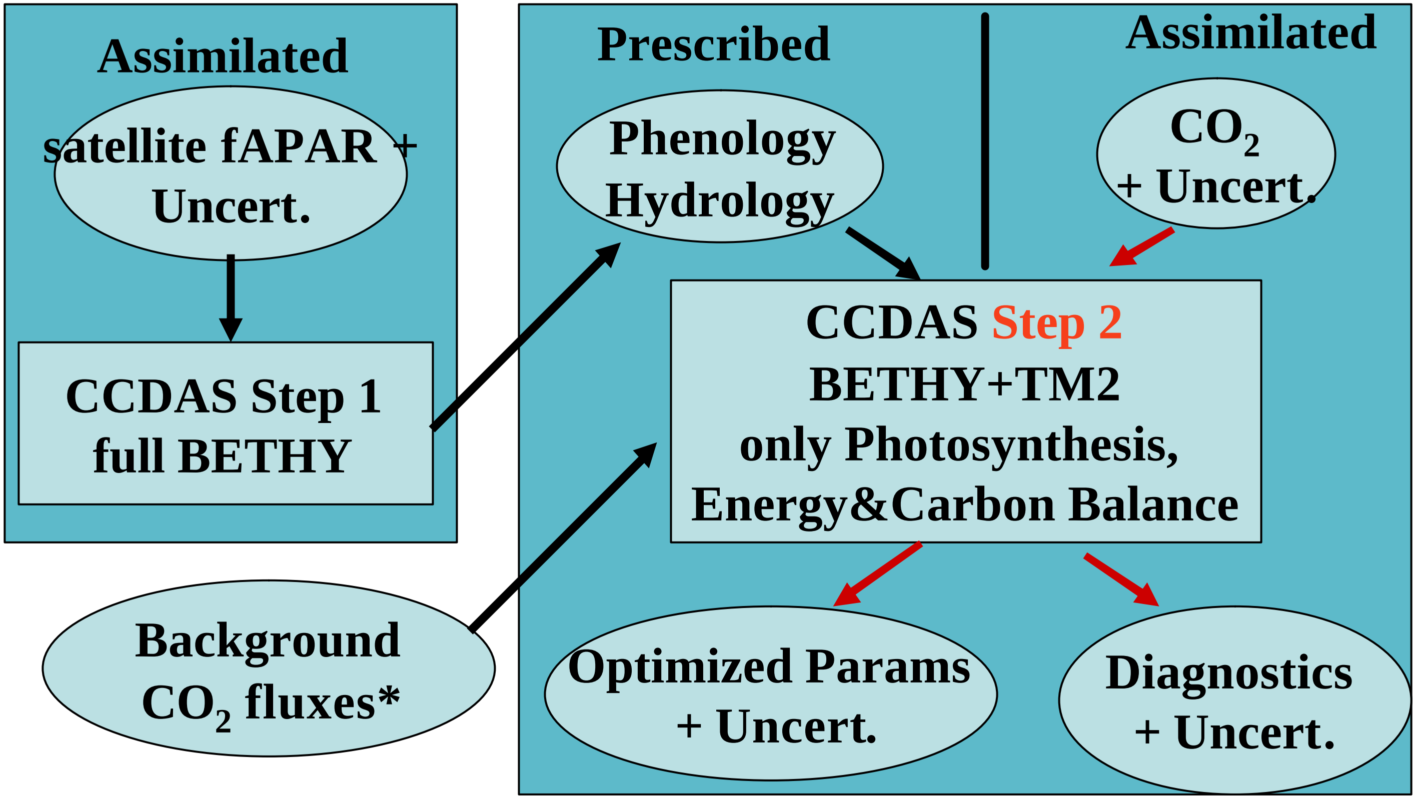
Adjoint model:

¶ Misfit / ? Parameters

➡ parameter optimization

adjoint of BETHY generated automatically using TAF (Ralf Giering & Thomas Kaminski, FastOpt, Hamburg)

Carbon Cycle Data Assimilation System (CCDAS)



* **ocean:** Takahashi et al. (1999), LeQuere et al. (2000); **emissions:** Marland et al. (2001), Andres et al. (1996); **land use:** Houghton et al. (1990)

Gradient Method

First derivative (Gradient) of $J(m)$ w.r.t. m (model parameters) :

–?$J(m)$/?m

yields direction of steepest descent

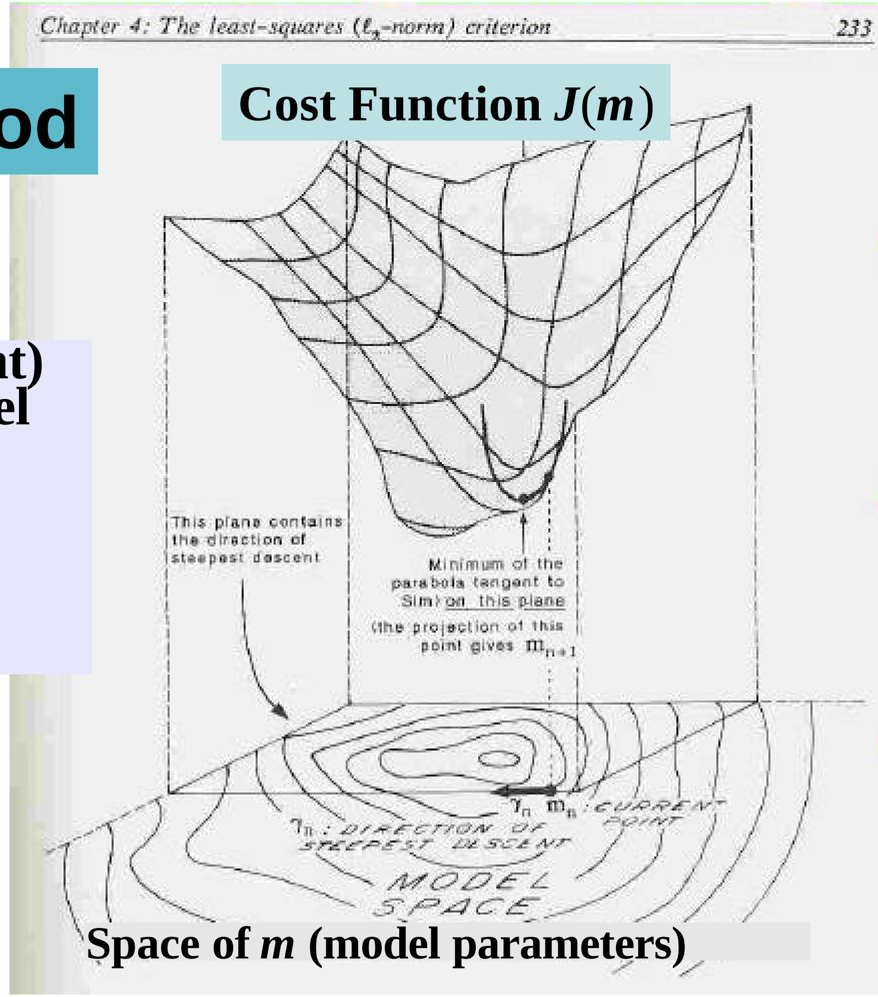
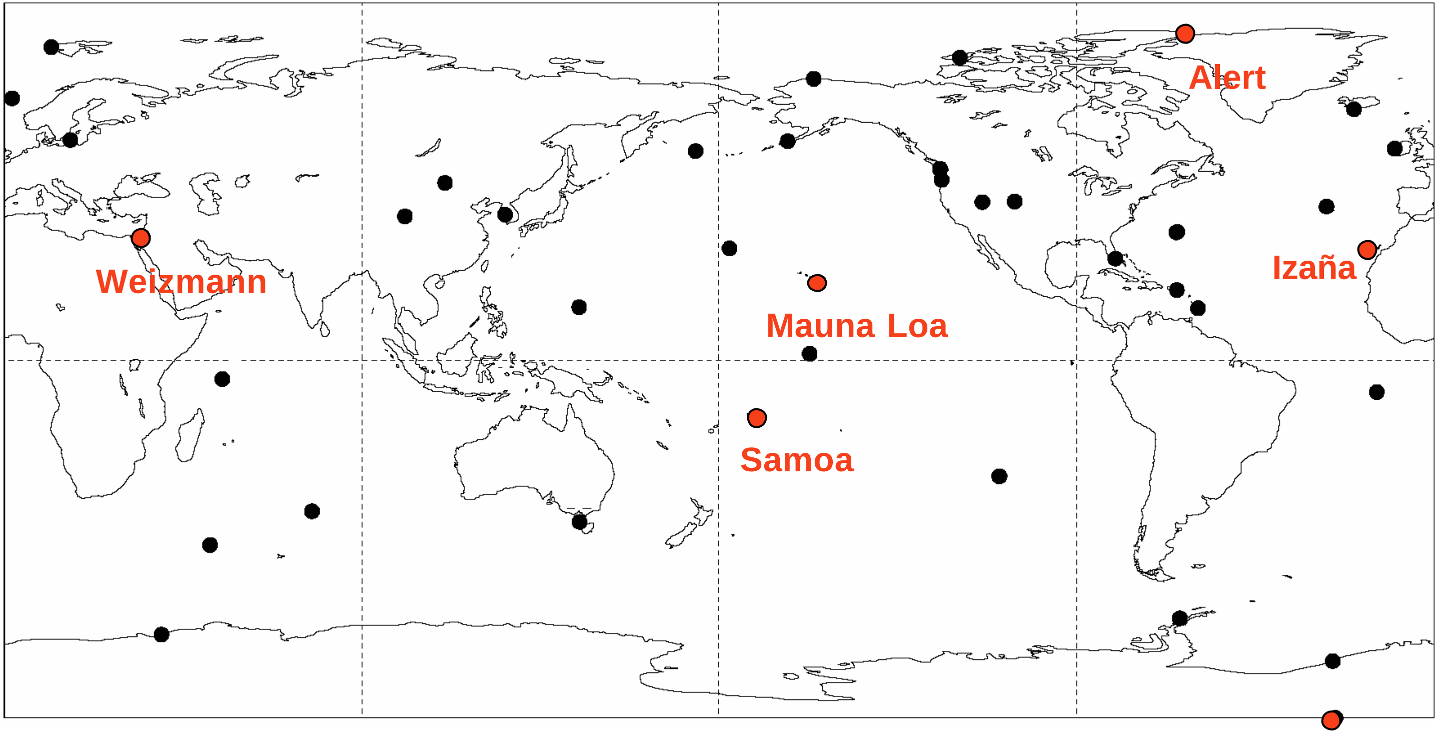


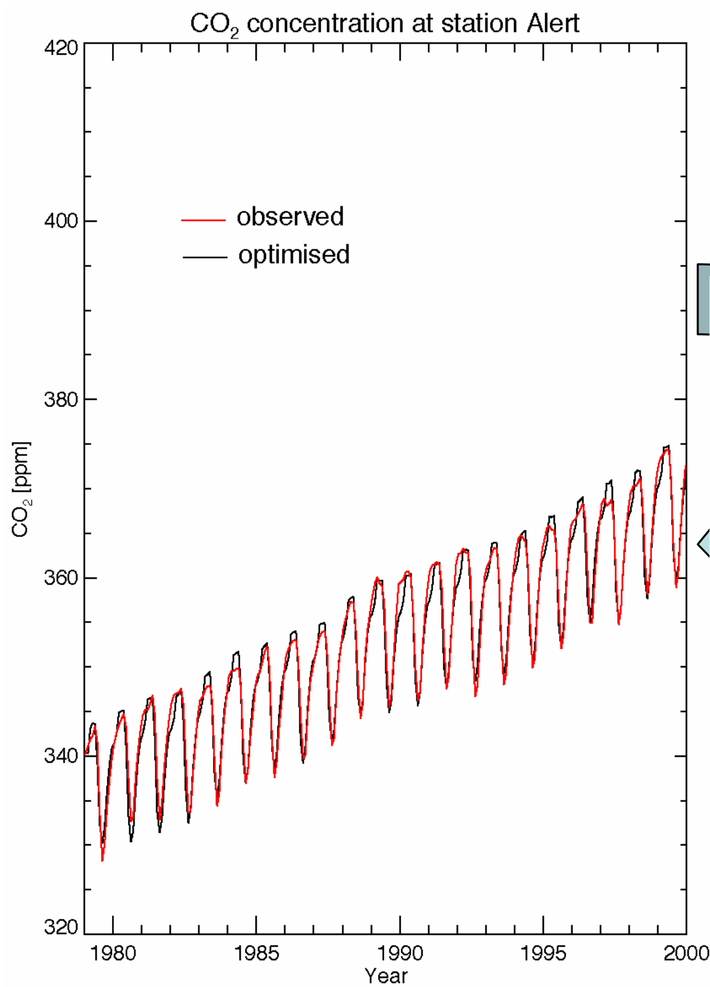
Figure taken from **Tarantola** '87

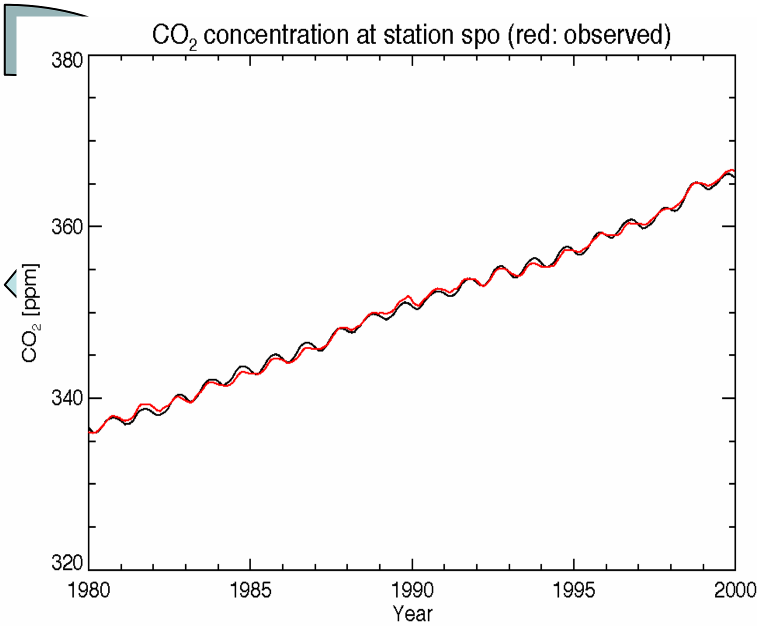
The Station Network



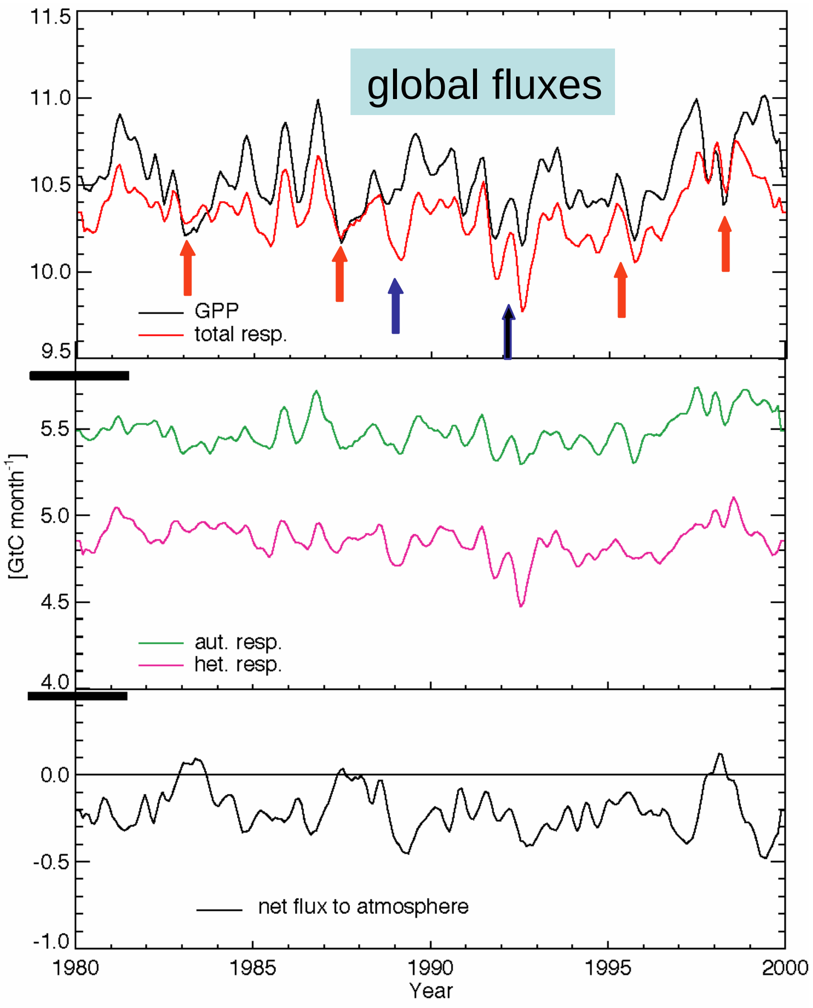
examples shown...

Optimisation





Optimised fluxes (1)



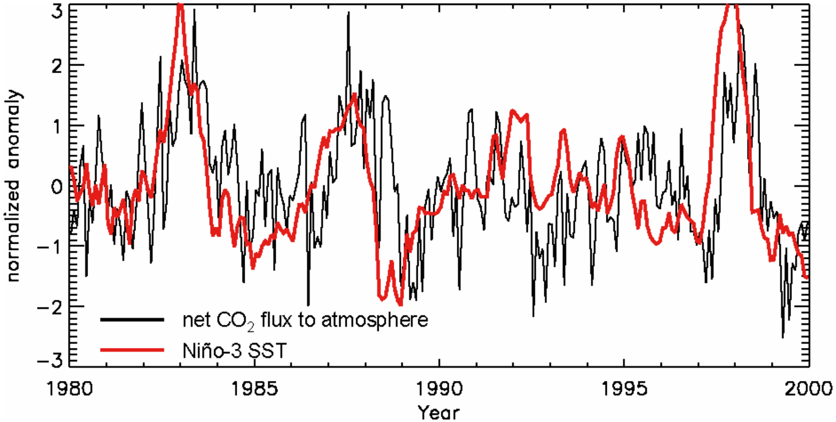
Major El Niño events

Major La Niña event

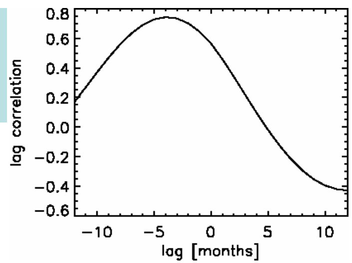
Post Pinatubo Period

Optimised fluxes (2)

normalized CO_2 flux and ENSO



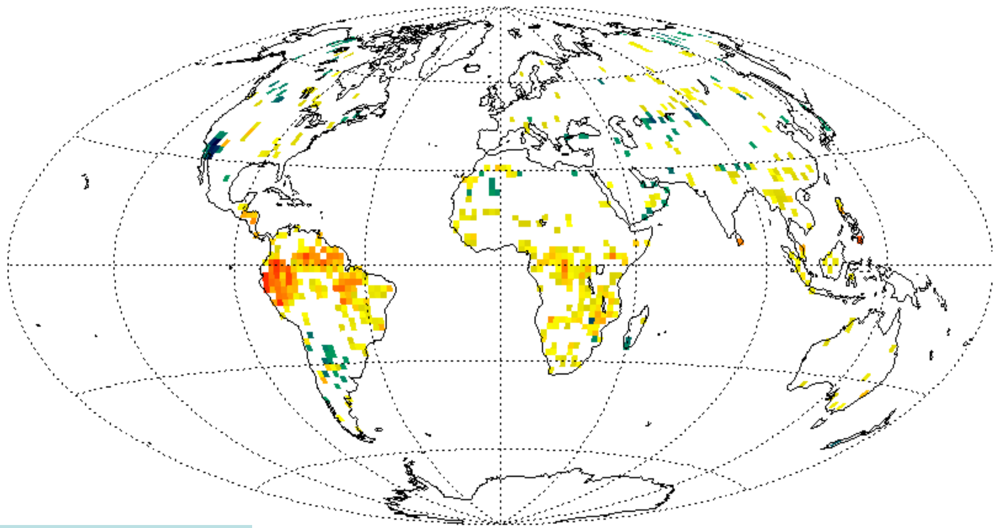
lag correlation
(low-pass filtered)

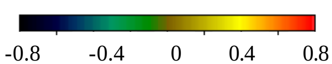


ENSO and terr. biosph. CO_2: correlation seems strong

correlation between Niño-3 SST anomaly and net CO_2 flux shows maximum at 4 months lag, for both El Niño and La Niña states

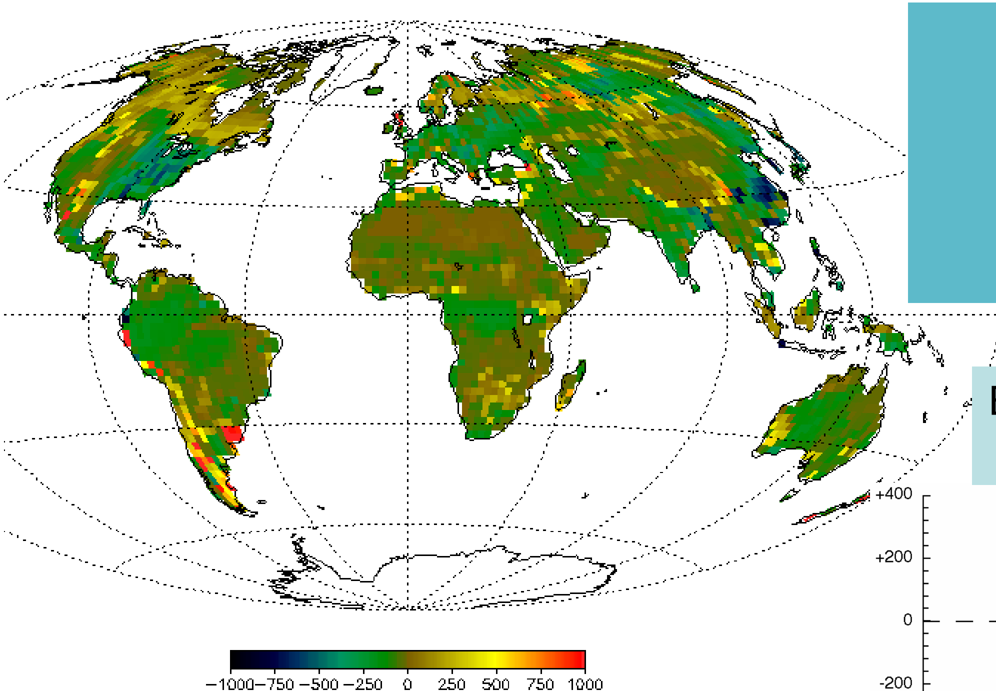
Optimised fluxes (3)





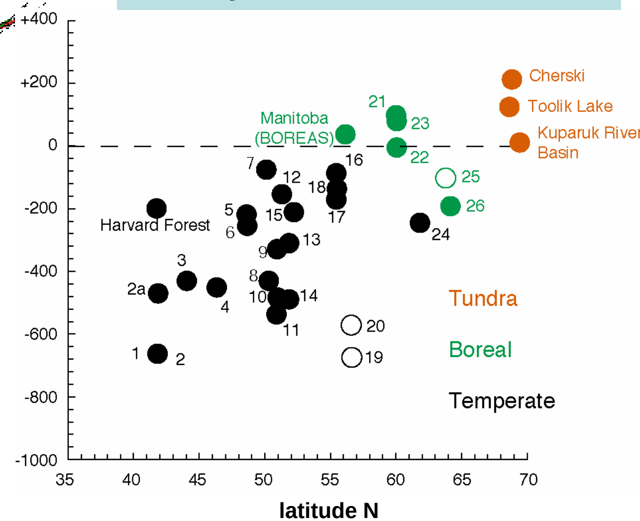
lagged correlation
at 99% significance

Carbon Balance



net carbon flux 1980-2000
gC / (m^2 year)

Euroflux (1-26) and other eddy covariance sites*



*from Valentini et al. (2000) and others

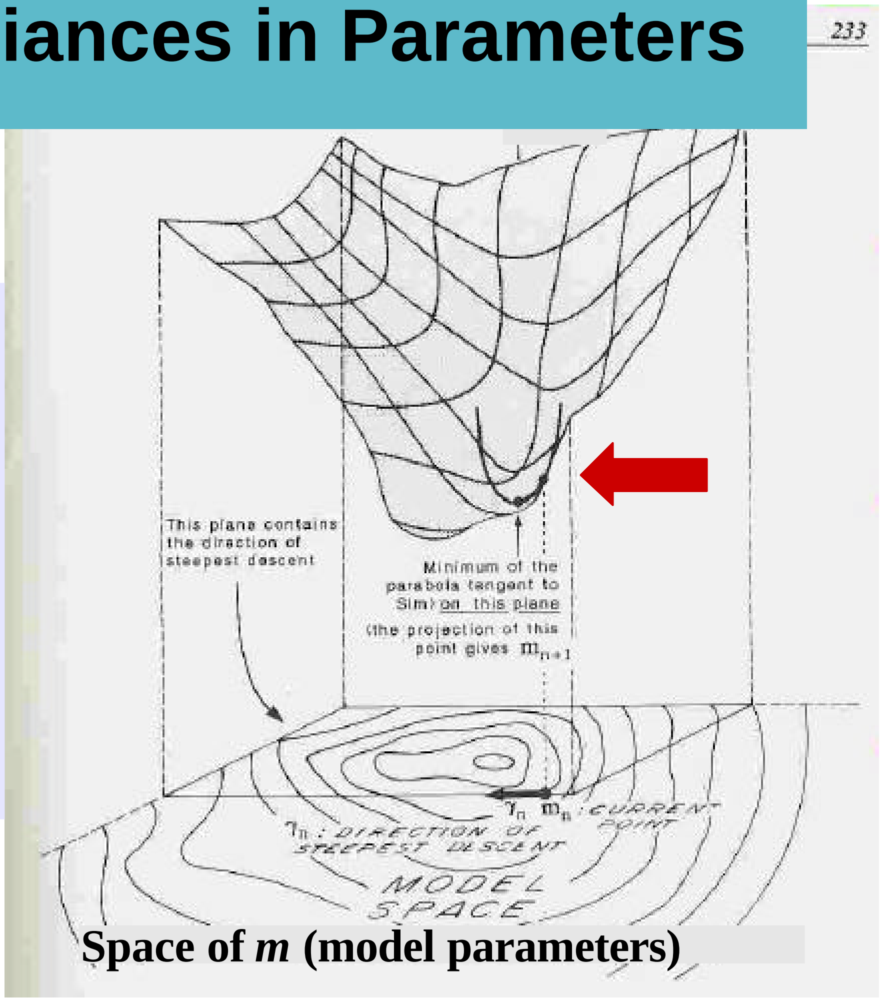
Error Covariances in Parameters

Second Derivative (Hessian) of *J*(*m*):

$$?^2 J(m)/?m^2$$

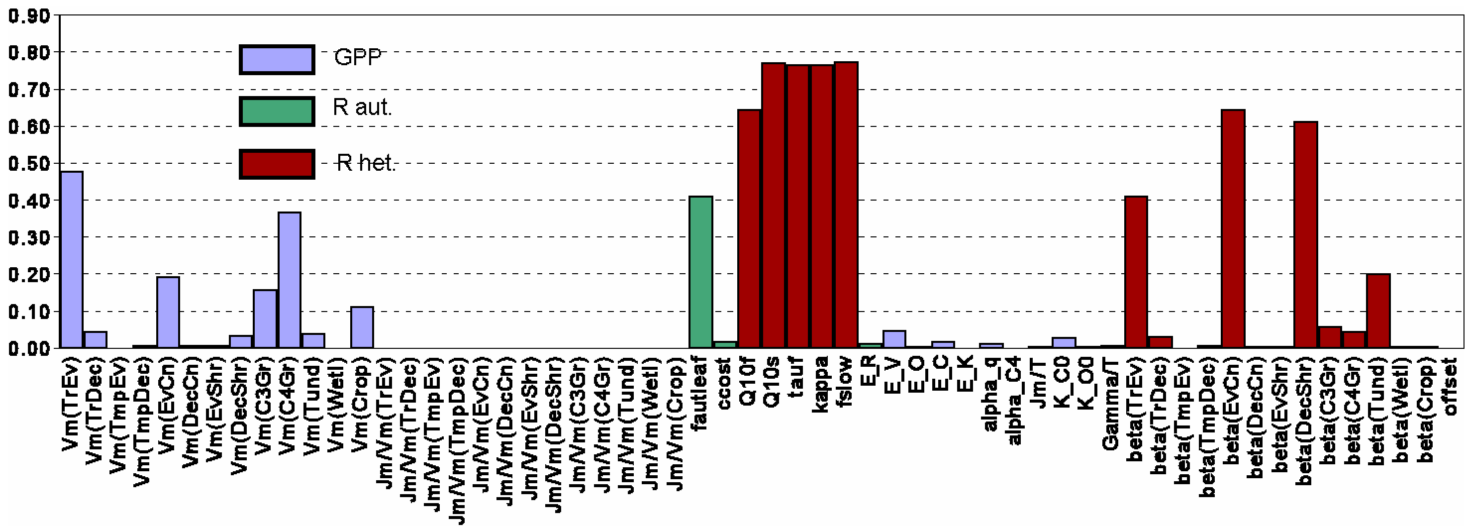
yields curvature of *J*, provides estimated uncertainty in m_{opt}

Figure taken from **Tarantola** '87



Space of *m* (model parameters)

Relative Error Reduction



$1 - s_{opt}/s_{prior}$

Error Covariances in Parameters

Error covariance of parameters after optimisation:

$$\mathbf{C}_m = \left\{ \frac{\partial^2 J}{\partial m_{i,j}^2} \right\}^{-1}$$ ***= inverse Hessian***

examples:

	first guess	optimized	prior unc.	opt.unc.	Vm(TrEv)	Vm(EvCn)	Vm(C3Gr)	Vm(Crop)
	μmol/m^{2}s	μmol/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**

Error Covariances in Diagnostics

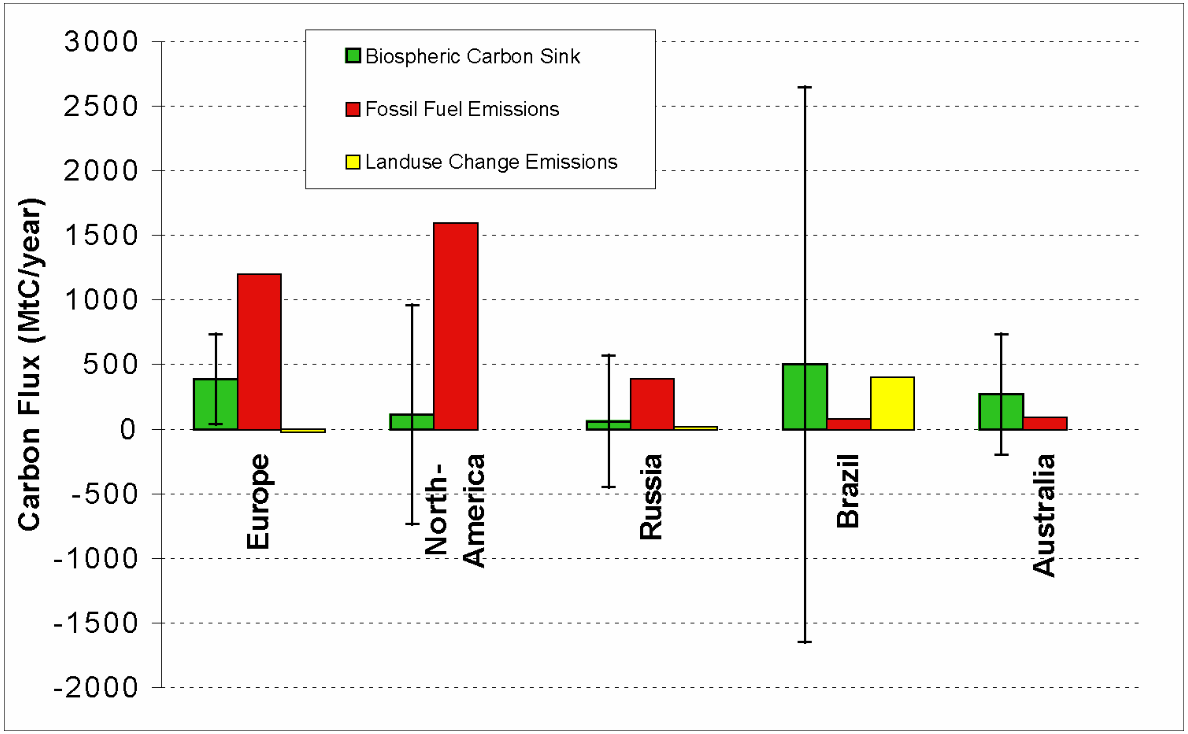
Error covariance of diagnostics, *y*, after optimisation (e.g. CO_2 fluxes):

adjoint or tangent linear model

$$\mathbf{C}_y(\vec{m}_{opt}) = \left(\frac{\partial y_i(\vec{m}_{opt})}{\partial m_j}\right) \mathbf{C}_m \left(\frac{\partial y_i(\vec{m}_{opt})}{\partial m_j}\right)^T$$

error covariance of parameters

Regional Net Carbon Balance and Uncertainties



Conclusions

- CCDAS with 58 parameters can already fit 20 years of CO_2 concentration data
- Sizeable reduction of uncertainty for ~13 parameters
- terr. biosphere response to climate fluctuations dominated by ENSO
- System can test model with uncertain parameters, and deliver a posteriori uncertainties of parameters, and of fluxes