Tackling large-scale optimisation problems in Earth Sciences by means of AD

Thomas Kaminski[1], Ralf Giering[1,]*
Marko Scholze[2], Peter Rayner[3], Wolfgang Knorr[2]
Patrick Heimbach[4,],Heinrich Widmann[2]*

Copy of presentation at http://www.FastOpt.com

1 FastOpt   2 Max-Planck-Institut für Biogeochemie   3 CSIRO   4 MIT   *

Outline

- Carbon Cycle Data Assimilation (CCDAS)
- Ocean Data Assimilation (ECCO)
- AD-Tools
- Summary

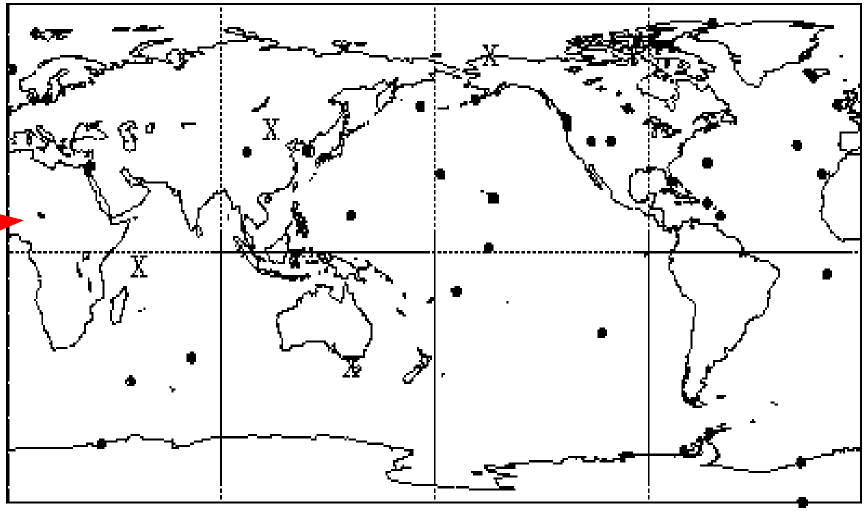
Carbon Cycle Data Assimilation (CCDAS)
Motivation

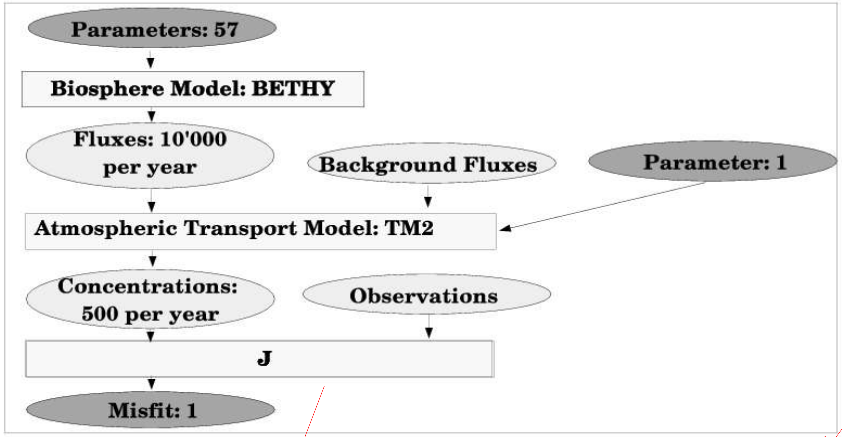
- Currently the terrestrial biosphere absorbes about one half (with large uncertainties) of the antropogenic carbon emissions (IPCC, 2001)
- How will the biosphere respond to climate change?

- Kyoto Protocol of United Nations Framework Convention on Climate Change (UNFCCC): Annex 1 Parties commit themselves to reducing their greenhous gas emissions by 5% between 1990 and 2010
- National inventories may take specific terrestrial sinks of carbon into account

CCDAS approach

- 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

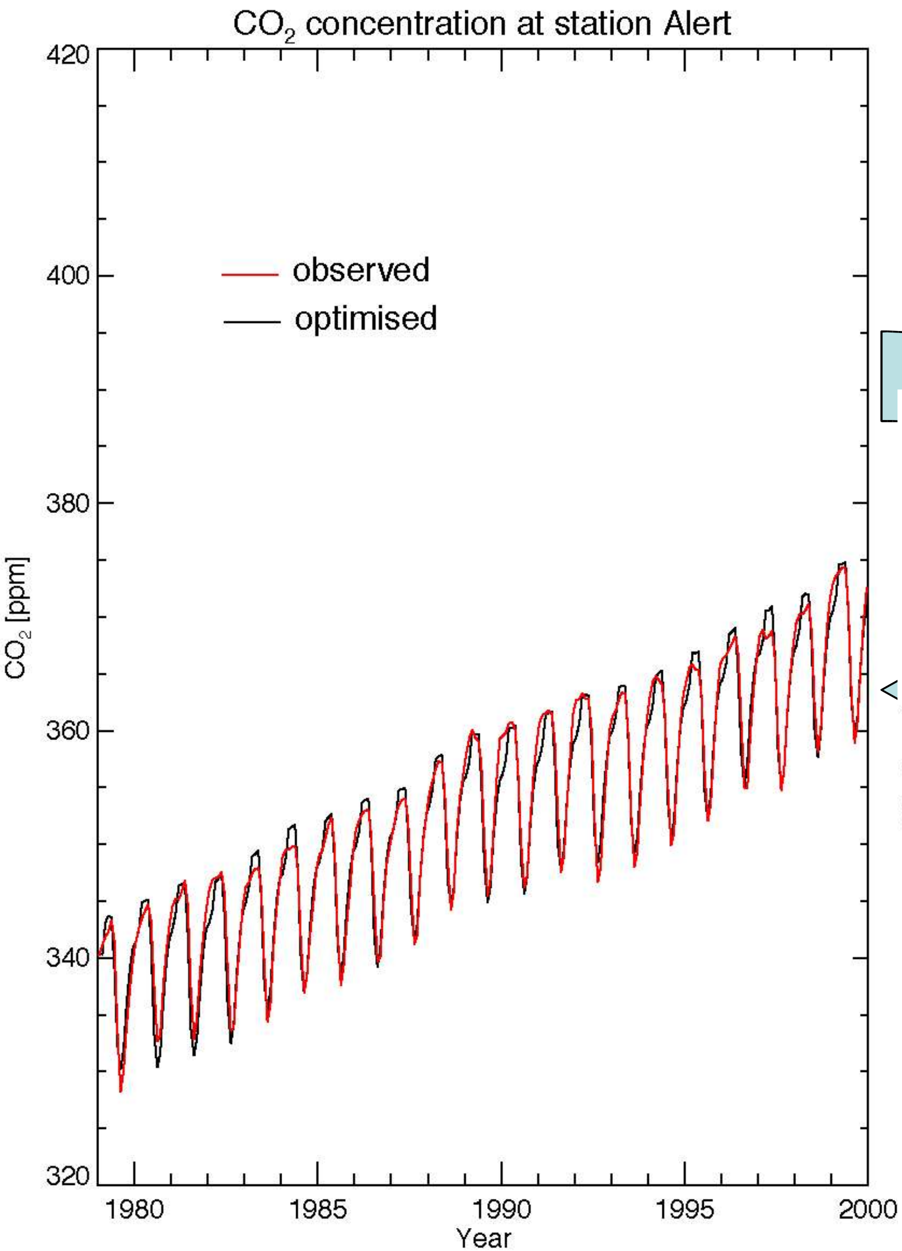
covariance of uncertainty in measurements + model

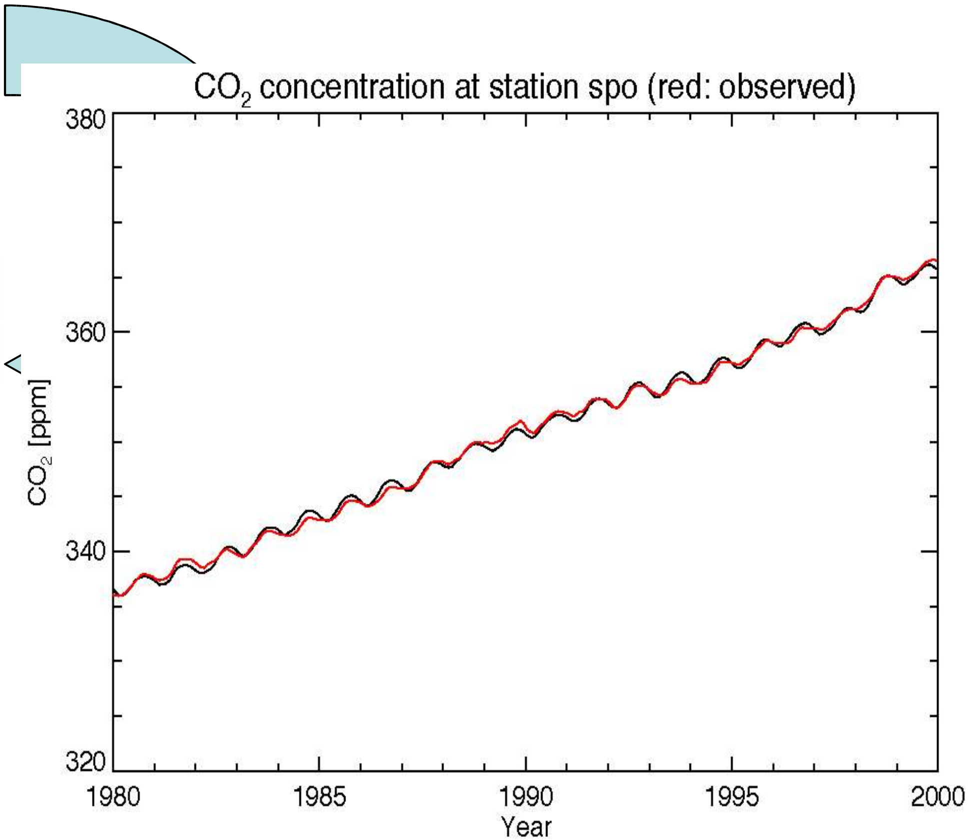
priors for parameters

observed concentrations

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

Optimisation
(BFGS+ adjoint gradient)





Posterior uncertainties

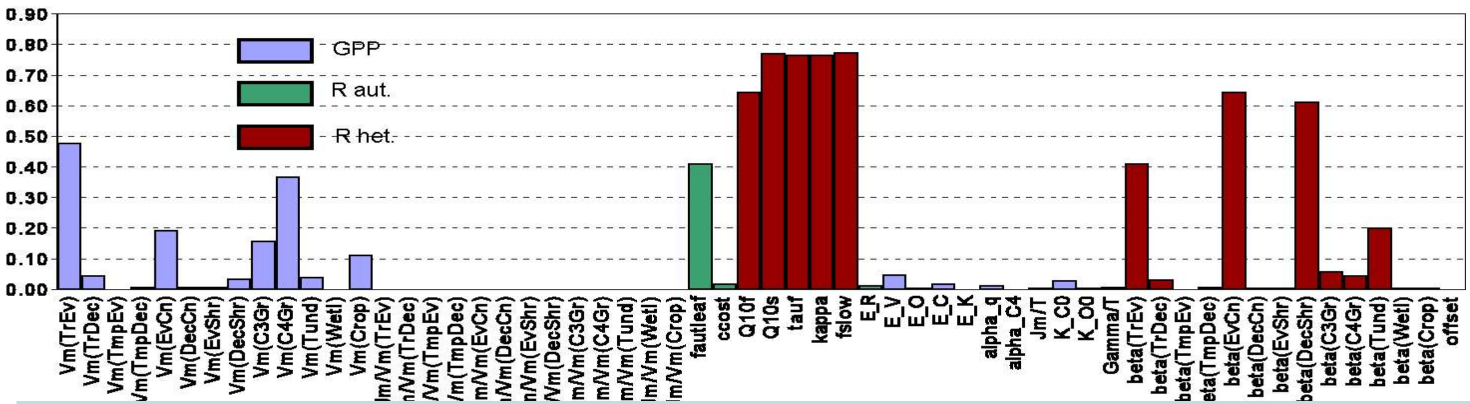
Use inverse Hessian of objective function to approximate posterior uncertainties

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

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**

Relative reduction of uncertainties



Observations resolve about 10- 15 directions in parameter spac

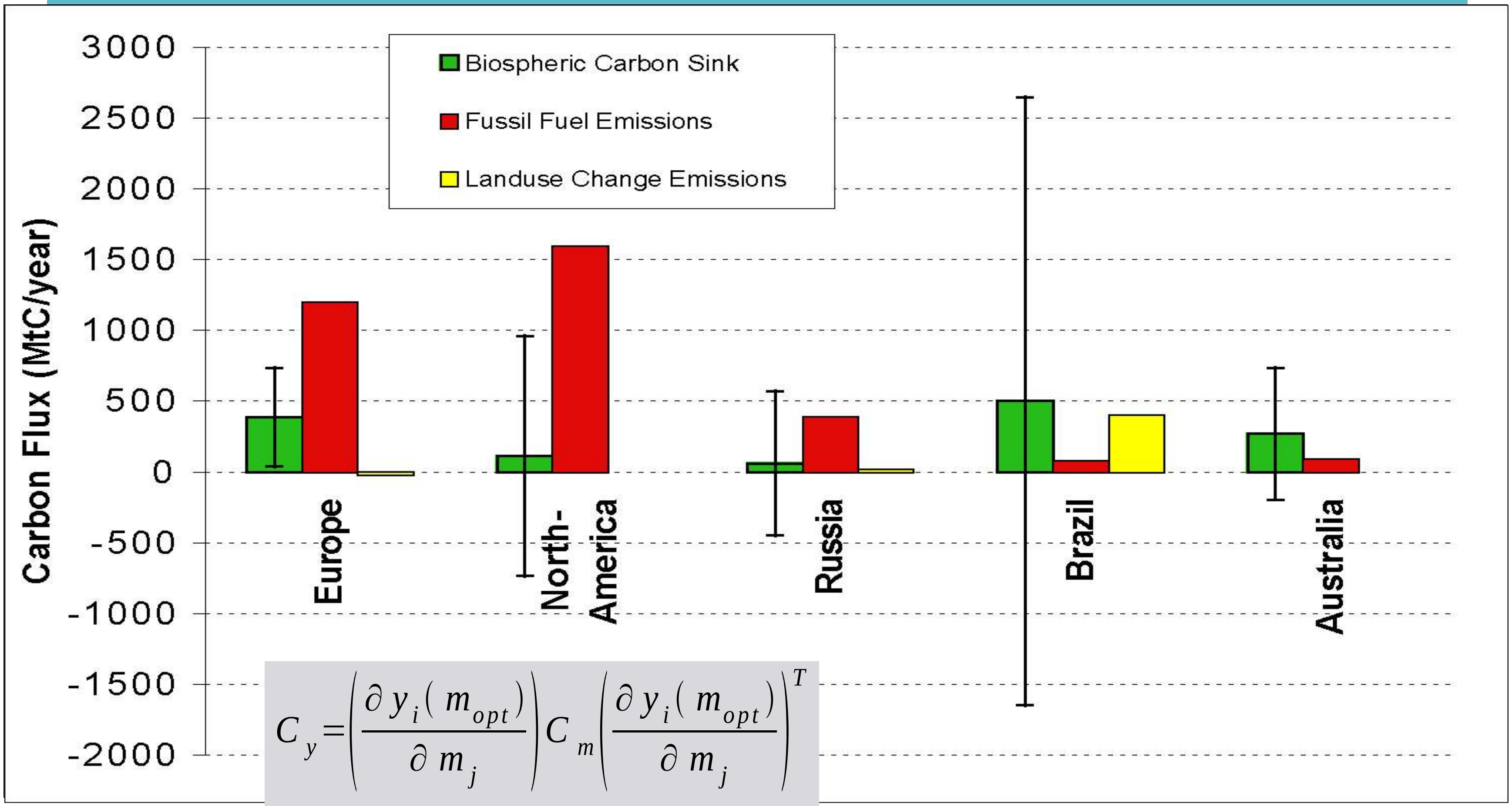
*Fast*Opt

Max-Planck-Institut für Biogeochemie

MPI-BGC Jena

CSIRO

Regional Net Carbon Balance and Uncertainties



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

*Fast*Opt

Max-Planck-Institut für Biogeochemie

MPI-BGC Jena

CSIRO

Automatic Sparsity Detection (ASD) in TAF

- TAF generates code to detect Jacobian sparsity pattern
- Propagates bit-patterns in forward or reverse mode

rows:

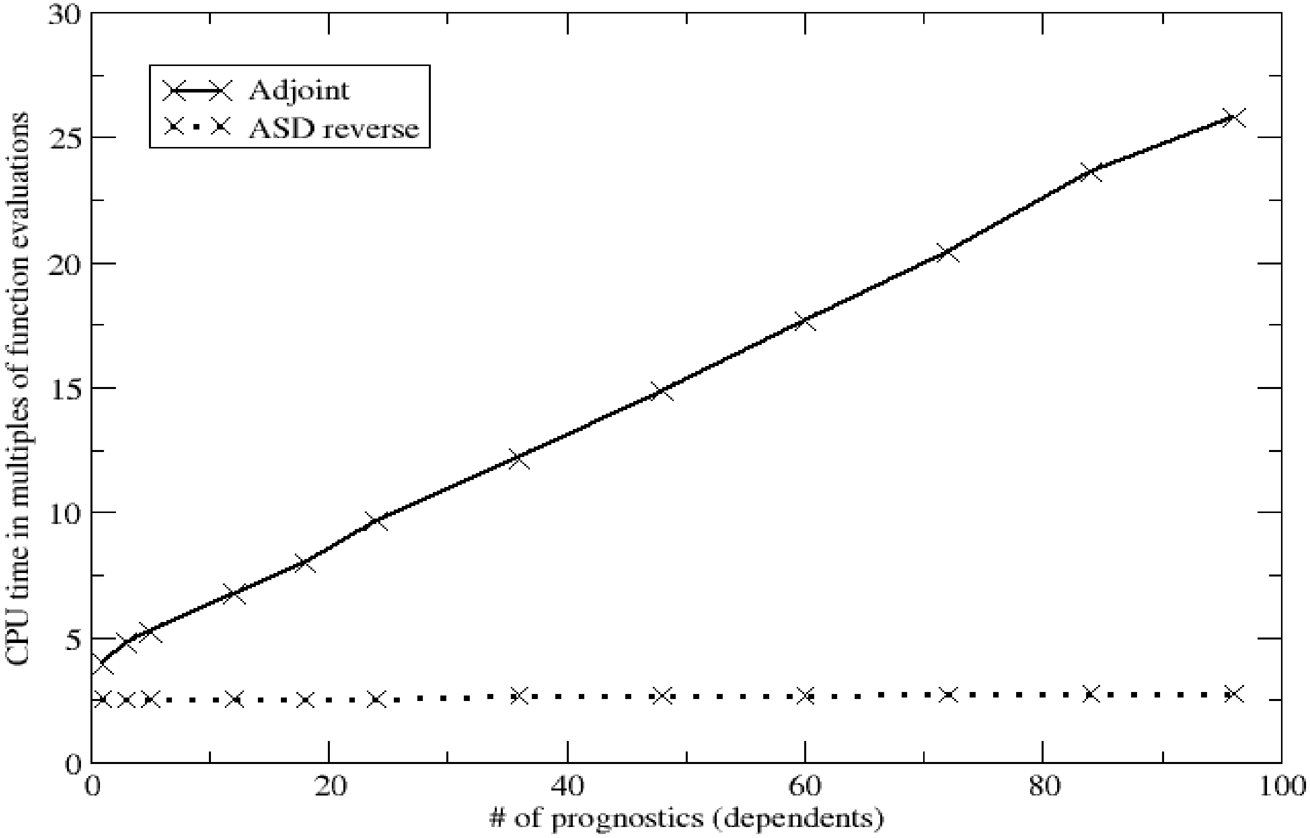
latitudinal bands

```
00000000000000000000000000000000000000000000000000000000000
00000000001000000000000000011111111110000111000000000000000
00000011001100000000110011011111111111010111000000011001100
00011111101100000111111011011111111111010111000111111101100
00011111111110001111111111111111111111111111100011111111110
00011111111110001111111111111111111111111111100011111111110
01111011111110111101111111111111111111111111101111011111110
01111011111110111101111111111111111111111111101111011111110
11111011111111111101111111111111111111111111111111011111110
11011011110111101101111011111111111111111111101101011110110
11010011110111101001111011111111111111111111101001011110110
11010011110111101001111011111111111111111111101001011110110
```

columns: parameters

FastOpt | Max-Planck-Institut für Biogeochemie | MPI-BGC Jena | CSIRO

Jacobian/ASD Performance



Model development within CCDAS

- System can test a given combination observational data + model formulation with uncertain parameters, and deliver optimal parameters, prognostics, and their a posteriori uncertainties
- Model is developed further within system, Adoint, Jacobian, and Hessian codes updated with TAF
- Model development benefits from sensitivity information and comparison with data (often tough!)
- Work is ongoing, numbers are from model formulation we are not yet happy with

Estimation of Circulation and Climate of the Ocean (ECCO)

- ECCO consortium is a large data assimilation effort by MIT, Scripps, and NASA-JPL plus international collaborators
- The ECCO data assimilation system is built around the MITgcm, a global 3 dimensional NS solver
- Estimation of an ocean state that is consistent with the MITgcm and various streams of data
- About 10^8 control variables:
  Initial temperature and salinity plus time dependent surface forcing (boundary condition)
- Adjoint code generated and maintained by TAF

The ECCO costfunction – observational elements

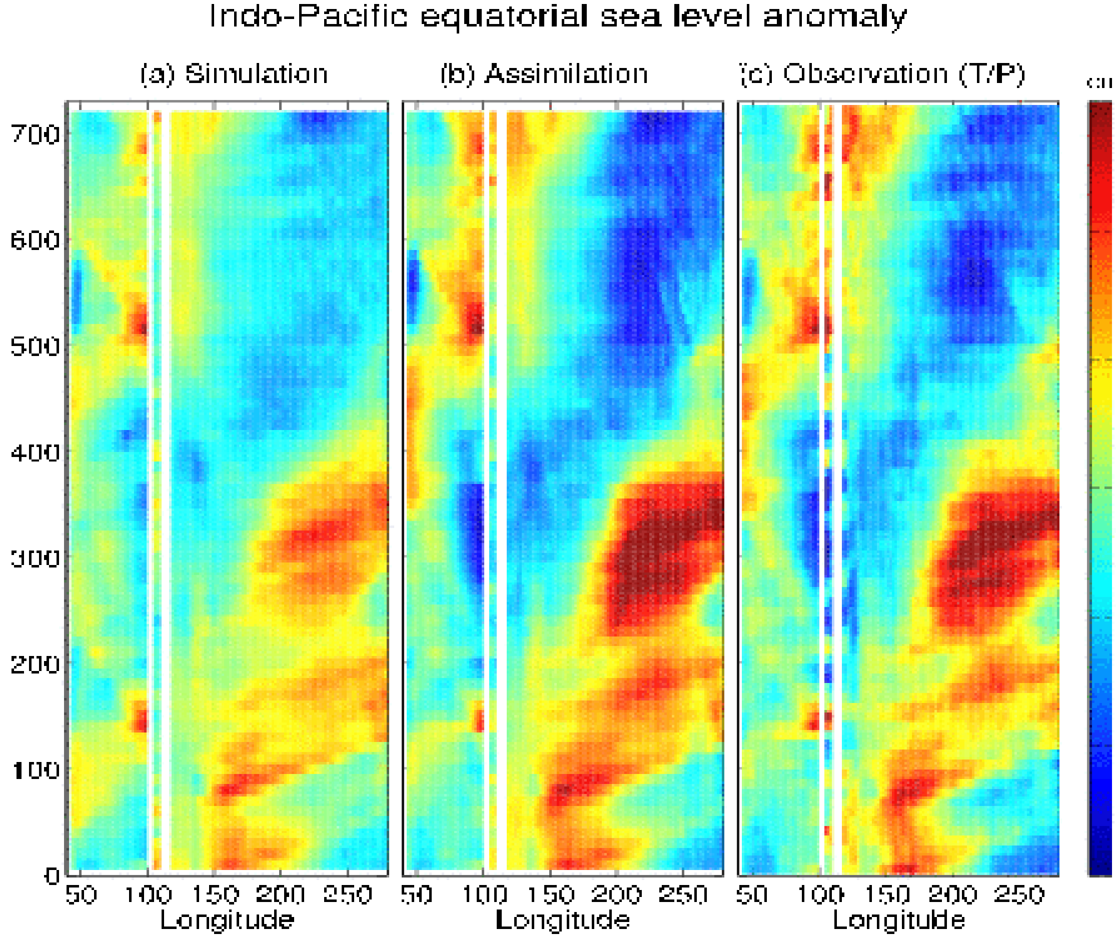
$\mathcal{J} = (\bar{\eta} - \bar{\eta}_{TP})^t \, \mathbf{W}_{\mathbf{geoid}} \, (\bar{\eta} - \bar{\eta}_{TP})$	TOPEX absolute SSH
$- \left(\eta - \eta'_{TP}\right)^t \mathbf{W}_{\mathbf{TP}} \left(\eta - \eta'_{TP}\right)$	TOPEX SSH anomalies
$- \left(\eta - \eta'_{ERS}\right)^t \mathbf{W}_{\mathbf{ERS}} \left(\eta - \eta'_{ERS}\right)$	ERS SSH anomalies
$- \left(\bar{T}_{surf} - \bar{T}_{Reyn}\right)^t \mathbf{W}_{\mathbf{SST}} \left(\bar{T}_{surf} - \bar{T}_{Reyn}\right)$	Reynolds SST
$\left(\bar{T} - \bar{T}_{Lev}\right)^t \mathbf{W}_{\mathbf{T_{Lev}}} \left(\bar{T} - \bar{T}_{Lev}\right)$	Levitus clim.
$\left(\bar{S} - \bar{S}_{Lev}\right)^t \mathbf{W}_{\mathbf{S_{Lev}}} \left(\bar{S} - \bar{S}_{Lev}\right)$	Levitus clim.
$- \left(\tau_x - \tau_{x,NCEP}\right)^t \mathbf{W}_{\tau_x} \left(\tau_x - \tau_{x,NCEP}\right)$	zonal wind stress
$- \left(\tau_y - \tau_{y,NCEP}\right)^t \mathbf{W}_{\tau_y} \left(\tau_y - \tau_{y,NCEP}\right)$	merid. wind stress
$- \left(H_Q - H_{Q,NCEP}\right)^t \mathbf{W}_{\mathbf{H_Q}} \left(H_Q - H_{Q,NCEP}\right)$	NCEP heat flux
$\left(H_F - H_{F,NCEP}\right)^t \mathbf{W}_{\mathbf{H_F}} \left(H_F - H_{F,NCEP}\right)$	NCEP freshwater flux

Currently added:

- Jason-1 altimetry (sea surface height)
- WOCE hydrography, XBT, TAO buoys
- PALACE/ARGO tracer profiles and drift velocities
- surface drifter velocities
- NSCAT/QuickScat surface wind stress fields
- TRMM/TMI tropical surface temperature fields

Optimisation: BFGS using Adjoint – 48 h per iteration on O3k

Results: Tropical Pacific State Estimation and El Niño



Sea surface height (SSH) evolution between 1997/98 (T. Lee et al., JPL/NASA)

More results on http://ecco-group.org

TAF
Transformation of Algorithms in Fortran

- Source-to-source translator for Fortran-77/95
- Commercial successor of TAMC
- Forward and reverse mode (1^{st} derivatives):
  Tangent linear and adjoint models
- Scalar and vector mode
- Efficient Hessian (2^{nd} derivative) code
  by applying TAF twice (e.g. forward over reverse)
- Command line program with many options
- TAF-Directives are Fortran comments
- Extensive and complex code analyses
  (similar to optimising compilers)
- Generated code is structured and well readable

TAF
More features

- Generation of flexible store/read scheme for required values triggered by TAF init and store directives
- Generation of simple checkpointing scheme (Griewank, 1992) triggered by combination of TAF init and store directives
- Generation of efficient adjoint (Christianson, 1996, 1998) for converging iterations triggered by TAF loop directive
- TAF flow directives for black-box routines,
  or to include user provided derivative code
  (exploit self-adjointness, MPI wrappers, etc...)
- Automatic Sparsity Detection
- Basic support for MPI and OpenMP

some larger TAF Derivatives

Model (Who)	Lines	Lang	TLM	ADM	Ckp	HES
NASA/GMAO (w. Todling & Lin)	87'000	F90	2.7	6.8	2 lev	-
MOM3 (Galanti & Tziperman)	50'000	F77	Yes	4.6	2 lev	-
MITGCM (ECCO Consortium)	100'000	F77	1.8	5.5	3 lev	11.0/1
BETHY (w. Knorr, Rayner, Scholze)	5'400	F90	1.5	3.6	2 lev	12.5/5
Nav.-Stokes-Solver (Hinze, Slawig)	450	F77	-	2.0	steady	-
NSC2KE (w Slawig)	2'500	F77	2.4	3.4	steady	9.8/1
HB_AIRFOIL (Thomas & Hall)	8'000	F90	-	3.0		-

- Lines: total number of Fortran lines without comments
- Numbers for TLM and ADM give CPU time for (function + gradient) relative to forward model
- HES format: CPU time for Hessian * n vectors rel. t. forw. model/ n
- 2 (3) level checkpointing costs 1 (2) additional model run(s)

TAC++

Transformation of Algorithms in C++

- Transfer well-proved TAF concepts to TAC++
- Design application driven
- Completed basic feasibility study (Vossbeck et al., 2004)
  - Generated adjoint of Roe solver
  - from EULSOLDO by Müller (1991)
  - 129 statements in C code (by f2c from Fortran-77)
  - C code contains simple constructs
  - Performance comparable to TAF generated code
- Functionality will be extended "application by application“"

Summary

- AD is a key technology in CCDAS and ECCO:
  Allows optimisation and propagation of uncertainties
- Concept can be generalised to other modelling systems
- AD helps to reduce the delay from model development to optimisation

More Info:

- http://CCDAS.org
- http://ecco-group.org
- http://FastOpt.com