Carbon Cycle Modelling: Some new techniques

with application to the Brazilian Proposal

& Earth System Model Development

Ian G. Enting

MASCOS

The University of Melbourne

Acknowledgments

- The Center of Excellence for Mathematics and Statistics of Complex Systems (MASCOS) is funded by the Australian Research Council (ARC).
- My fellowship at MASCOS is supported by CSIRO through a sponsorship agreement.
- The Fortran-90 development is supported by the ARC Earth System Science Network (ARCNESS).
- Collaborators: Cathy Trudinger and YingPing Wang of CSIRO Marine and Atmospheric Research and members of the MATCH working group on the Brazilian Proposal.

Summary

- More than integration:
  Model analysis systems
  - Differentiation and some implementations
- Applications of differentiation in:
  - Analysing the Brazilian proposal
  - Calibrating CASACNP terrestrial carbon model (with N and P)

Model Analysis

Most common model calculation is forward projection by (numerical) integration of DEs.

Much model analysis involves differentiation:

Initialisation Solving for zero rate of change

Sensitivity analysis — derivatives with respect to parameters

Calibration — techniques such as Maximum Likelihood imply optimisations, facilitated by use of derivatives

Data assimilation — real-time model adjustment — dynamic calibration

Data for Carbon Cycle Studies

- Air sampling networks interpreted by inverse modelling;
- Satellite data, for quantities such as leaf-area index and phenology
- Terrestrial biosphere models;
- Convective boundary layer measurements;
- Stand-level flux networks;
- Ecosystem experiments;
- Small cuvettes.

From Canadell et al. 2000.

Also satellite data for CO_2 concentrations.

Key characteristics of statistics

- magnitude;
- degree of correlation between components;
- temporal correlation structure;
- spatial correlation structure;
- distribution;
- mismatches in averaging;
- contribution from model representativeness error.

From Raupach et al. 2005

Algorithmic Differentiation (AD)

Differentiation by successive use of chain rule.

For binary operation $c = f(a, b)$,

$$\frac{\partial c}{\partial \alpha} = \frac{\partial f}{\partial a} * \frac{\partial a}{\partial \alpha} + \frac{\partial f}{\partial b} * \frac{\partial b}{\partial \alpha}$$

e.g.

$$c = a + b \quad \rightarrow \frac{\partial c}{\partial \alpha} = \frac{\partial a}{\partial \alpha} + \frac{\partial b}{\partial \alpha}$$

$$c = a * b \quad \rightarrow \frac{\partial c}{\partial \alpha} = b * \frac{\partial a}{\partial \alpha} + a * \frac{\partial b}{\partial \alpha}$$

Convert program to code for derivatives, one operation at a time.

Using differentiation (by overloading)

Original

```
double F_co2(double c){
double a;
a = log(c/280.0)*5.35;
return a;
};
...
double cc;
...
ff = F_CO2(cc)
```

Transformed

```
Xvar F_co2(Xvar c){
Xvar a;
a = log(c/280.0)*5.35;
return a;
};
...
Xvar cc;
// Derivatives wrt
// initial value of cc
cc.set(280,1);
...
ff = F_CO2(cc)
```

Class Definitions

Fragment of C++ class definition (300 lines) to implement operator overloading:

```
class Xvar{
public :
static const int ns = _NUMDERIVS+1;
double xs[_NUMDERIVS+1];
Xvar operator*(Xvar);
...
};

Xvar Xvar::operator*(Xvar b){ Xvar c;
for (int i=1; i < ns; i++)
   c.xs[i] = xs[i]*b.xs[0]+xs[0]*b.xs[i];
c.xs[0] = xs[0]*b.xs[0];
return c;} ;
...
```

Tangent Linear Model (TLM)

For N DEs: $\quad \frac{d}{dt}x_j = g_j(\{x_k\}, \alpha, t) \qquad$ for $j = 1, N$

we can define sensitivities as

$$y_j = \frac{\partial}{\partial \alpha} x_j \qquad \text{for } j = 1, N \text{ or } \qquad y_{j,p} = \frac{\partial}{\partial \alpha_p} x_j$$

to give 'tangent linear model(s)':

$$\frac{d}{dt} y_m = \frac{\partial}{\partial \alpha} g_m(\{x_k\}, \alpha, t) + \sum_n \frac{\partial}{\partial x_n} g_m(\{x_k\}, \alpha, t)\, y_n$$

$$\frac{d}{dt} y_{m,p} = \frac{\partial}{\partial \alpha_p} g_m(\{x_k\}, \alpha, t) + \sum_n \frac{\partial}{\partial x_n} g_m(\{x_k\}, \alpha, t)\, y_{n,p}$$

Adjoints (of TLM)

$$\frac{d}{dt} y_{m,p} = \frac{\partial}{\partial \alpha_p} g_m(\{x_k\}, \alpha, t) + \sum_n \frac{\partial}{\partial x_n} g_m(\{x_k\}, \alpha, t)\, y_{n,p}$$

$$\frac{d}{dt} y_{m,p} = f_{m,p}(t) + \sum_n g_{m,n}(t)\, y_{n,p}$$

TLM is **linear** and therefore has a (linear)

Green's function operator which has an adjoint.

Jacobian is projection of operator onto matrix.

$J_{k,1}$ gives $\frac{\partial v_k}{\partial u_1}$: carry forward single set of $\frac{\partial x_m}{\partial u_1}$

$J_{1,j}$ gives $\frac{\partial v_1}{\partial u_j}$: carry backward single set of $\frac{\partial v_1}{\partial x_m}$

Brazilian Proposal

Tabled by Brazil during negotiations leading to Kyoto Protocol — Flicked-passed to Subsidiary Body for Scientific and Technical Advice (SBSTA).

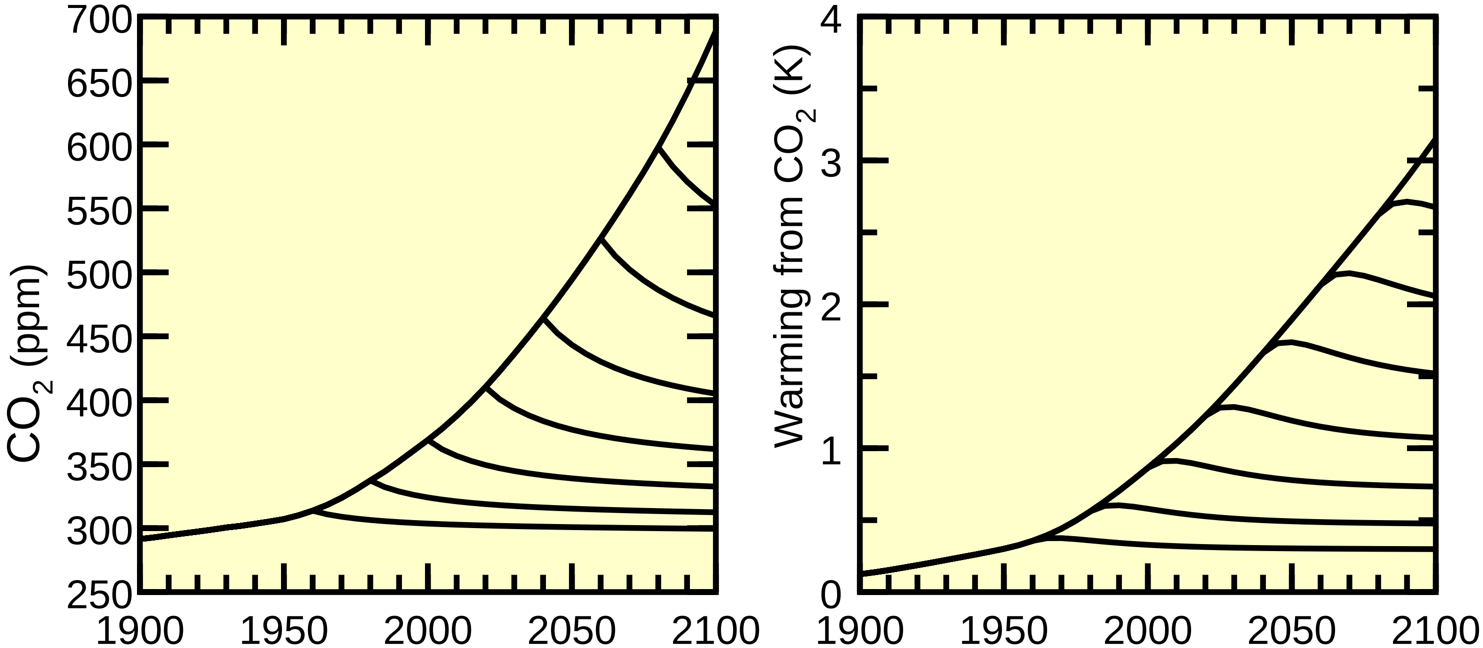
Proposes that emission reduction targets should be proportional to nation's relative responsibility for the greenhouse effect.

Issues:

- Indicator? What quantity is used as a measure of the greenhouse effect?
- For what period of emissions is responsibility attributed?
- How are non-linear responses attributed?

Timescales

CO_2 concentrations and consequent warming, partitioned according to time of emission.



Lowest bands are from pre-1960 emissions, next from 1960 to 1980 emissions, etc. Increase in contribution to warming after time of emissions from ‘committed warming’ effect.

Brazilian Proposal as Derivatives

As example, use indicator $T^* = T_{\mathrm{CO2}}(2100) =$ warming in 2100 from CO_2 emissions.

T^* is to be attributed to emissions $E_j(t)$ from country j with $E(t) = \sum_j E_j(t)$.

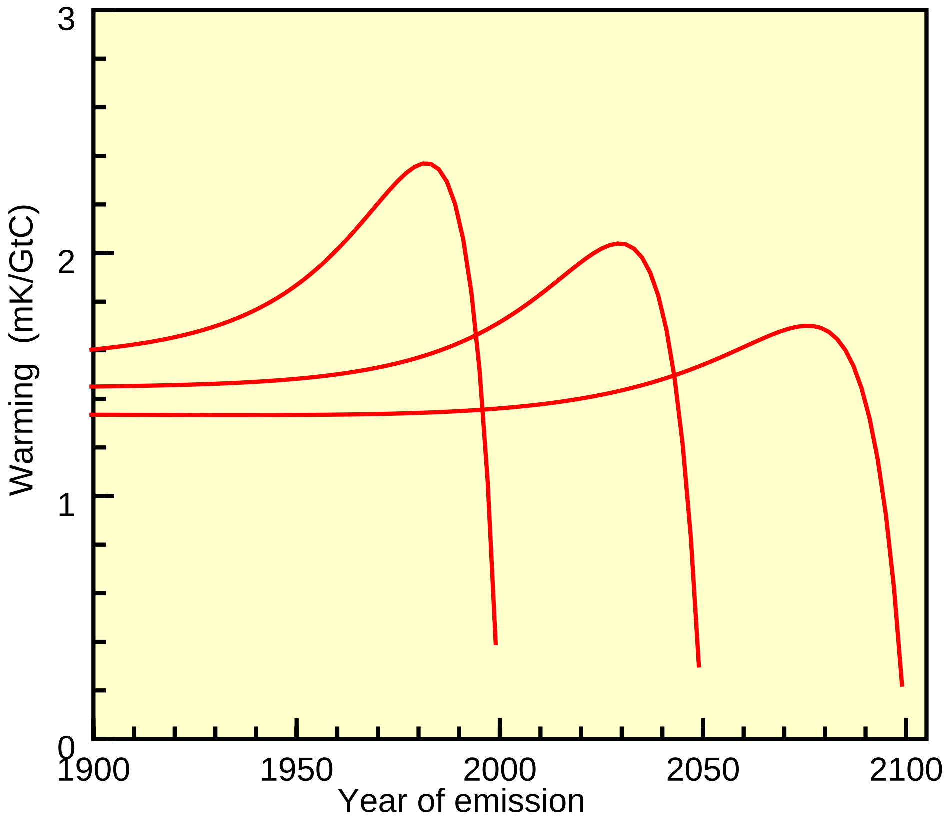
Differential attribution to country j of emissions at time t is

$$\frac{\partial T^*}{\partial E_j(t)} E_j(t) = \frac{\partial T^*}{\partial E(t)} E_j(t) = S(t) E_j(t)$$

where $S(t)$ is a Frechet derivative.

Cumulated attribution: $T_j^* = \int S(t) E_j(t)\, dt$

Results: Frechet Derivatives



Assumes IS92a emissions. Represents temperature by response function. Linear responses for ocean and biotic carbon, coupled non-linearly to atmospheric CO_2 (as in CSIRO study).

$$\frac{\partial}{\partial E(t)} T(\tau) \text{ for } \tau = 2000,\ 2050,\ 2100.$$

Decrease as $t \rightarrow \tau$ shows 'committed warming'.
At any time, warming from the most recent releases is yet to happen.

Implications

- For a given indicator, T^*, calculation of $S(t)$ allows attribution to any nation.
- $S(t)$ most efficiently calculated from adjoint model, but for multiple indicator times, tangent linear model not too inefficient.
- Sensitivity of T_j^* to model uncertainties can be obtained as second derivatives.
- Sensitivity of T_j^* to uncertainties in emissions can be obtained as

$$\text{Var}[T_j^*] = \int\int S(t)\text{Cov}[E_j(t), E_j(t')]S(t')\,dt'\,dt$$

CASACNP AD Project

- Develop Fortran-90/95 implementation of AD by operator overloading;
- Use CASACNP as test case of AD in existing model;
- Explore use of CASACNP derivatives for calibration;
- Document procedure for ARC Network for Earth System Science.

Development funding obtained from ARCNESS for CASACNP and extension to CABLE.

Earth system Modelling in Australia

ACCESS: Australian Community Climate and Earth System Simulator.
Bureau of Meteorology, CSIRO and universities.

- UKMO atmosphere (and 4DVAR for NWP)
- Existing ocean used by CSIRO and BMRC (BoM Research Centre)
- Land surface component:
  - existing CSIRO (CABLE)
  - CASA for carbon pools (adding N and P)
  - LPJ for dynamic vegetation

CASACNP

Extend CASA terrestrial carbon model to include nitrogen and phosphorus. (Wang, Field and Houlton).

- Initially single region, 8 C pools, 9 N pools, 12 P pools.
- Models response to addition of P, N and competition between N-fixers and non-fixers.
- Validation transect (of soil age) in Hawaii.
- Initial conversion to use AD successful: May 2006 — development on-going.

Putting AD in CASACNP

- Added *max* and *min* to initial AD definitions
- Redeclare types of variables
  — global edit — not efficient code
- Convert input to write to value – initialise derivatives to zero using compiler option
- Identify variables for differentiation
- Explicit re-writes of .GT. tests using values
- Explicit loops for *array* = *real* — explicit *1/delt*

Progress on CASACNP

- C++ AD project on Brazilian Proposal: Presented at MODSIM 2005. Proof-of concept for AD by operator overloading.

- Successful proof-of concept of AD by operator overloading in Fortran-95.

- Successful runs using 'brute-force' conversion of CASACNP (May 2006)

- Designing 'vector' version (Spring 2006).

- *Development on-going, supported by ARCNESS*

Parameter Sensitivity: 1

Which parameters most influence the results in the desired application areas?

Parameter Sensitivity: 2

Which parameters influence observable quantities?
(thereby allowing calibration).

Does this set of parameters include all those you need?

In Bayesian terms, compare s.d. of observations $\times \frac{\partial m}{\partial \alpha}$
with prior s.d. of parameters.

Some general considerations

- A carefully designed tool will usually be more effective than a generic tool (Griewank).
  (c.f. Unix philosophy: tools that do one thing well, and are combined as required by user)

- The local nature of land-surface processes (carbon and water) suggest that tangent linear model may be as effective as adjoint model.

Future directions

- Incorporate AD into calibration tools.
- Second derivatives in Fortran.
- Explore overloading for Fortran vector operations.
- Develop calibration strategy for CASACNP and other carbon components of Australian Community Climate and Earth System Simulator (ACCESS).

Conclusions

Algorithmic differentiation —

Operator overloading is a straightforward way of developing tangent linear models (and obtaining higher derivatives if needed).

Brazilian Proposal —

Attribution in terms of derivatives is readily calculated using algorithmic differentiation. Higher derivatives give sensitivities.

Global change —

Potential should extend to other analyses of uncertainties in global change.

Further Information

Andreas Griewank, 2000, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, (SIAM, Philadelphia).

MATCH website (Brazilian Proposal):
http://www/match-info.net

I.G. Enting, 2005, *Automatic differentiation in the analysis of strategies for mitigation of global change*, International Congress on Modelling and Simulation, Melbourne, 2005. Ed. A. Zerger and R. M. Argent, 7pp
http://www.mssanz.org.au/modsim05/papers/enting.pdf