An Introduction to Forward Modelling

Peter Rayner

June 18, 2015

Outline

- Different meanings of the word “model”;
- Types of model;
- Gaps between theory and model;
- Some important principles;
- There is no “best” model;
- Some important analogies;
- Key fluxes and stores for the terrestrial biosphere;
- The OPTIC model;

What is a Model?

- A quantitative description of a physical system;
- May be embodied as computer code or mathematical equations;
- For us just changing the parameters doesn't generate a new model, statisticians use "model" to refer to a particular parameter set;
- Models don't have to be deterministic, e.g. we often have equations for the variance of a quantity without calculating individual realizations.

Different Classes of Model

- *Prognostic Models* predict future state given current state and forcing. Usually involve a timestepping loop;
- *Diagnostic Models* Calculate some outputs given a different set of inputs;
- Terrestrial biosphere models of both types.

First Principles vs Heuristic Models

- Sometimes we know the fundamental equations of the system exactly (e.g. fluid flow);
- This does *not* translate into perfect knowledge;
- Usually cannot construct model from fundamentals (e.g. biosphere models do not start from quantum Mechanics).
- Proceed either by analogy or by establishing relationships from measurements then generalizing them.

Perfect theory $\neq$ Perfect Model

- Simple model $\frac{\partial z}{\partial t} = a \times b$
- Perfect at a point;
- Can only measure or compute a and b at resolution δ;
- $\frac{\partial \bar{z}}{\partial t} = \overline{a \times b} = \bar{a} \times \bar{b} + \overline{\prime a \times \prime b}$
- overbar = averaging over δ and prime variations $< \delta$;
- Must estimate $\prime a$, $\prime b$ and their relationship;
- Often called "closures" and often semi-empirical.

A few words on Chaos

- The previous slide does *not* describe chaos;
- Chaos for us means large sensitivity of final state to small variations in initial;
- Limits what we can do with the model (e.g. cannot predict weather in one year) but does not imply imperfect model;

Some Basic rules

- Many physical systems conserve quantities (mass, energy, momentum etc);
- Best if models do that too;
- This is surprisingly difficult for numerical approximations;
- Many models trade off accuracy and conservation.

Models for Prediction or Research

Prediction

- Don't need to understand them;
- easy to test against their predictions;
- Forecast performance and computational efficiency most important.

Research

- Numerical laboratories;
- As hard to understand as the system itself but we can test things;
- Usually simpler models to understand complex ones (diagnostics);
- Should be as simple as possible to answer the question, but no simpler.

There is no "best" Model

- Every model has a limited domain;
- For research we must explain the model either directly or with simpler models;

Stores, fluxes and residence times

$$\frac{\partial \text{store}}{\partial t} = \text{flux in} - \text{flux out}$$

or

$$\frac{\partial q}{\partial t} \propto \text{flux in} - \text{flux out}$$

- Comes from conservation of mass;
- Prognostic equation for q;
- Often flux out $= kq$;

Residence time

$$\frac{\partial q}{\partial t} = -kq$$

$$q(t + \delta t) = q(t) - kq(t)\delta t$$

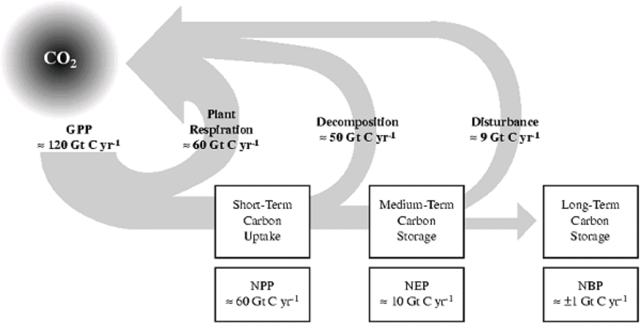
- After time $\delta t = 1/k$ $q(t + \delta t) = 0$;
- $1/k$ called "residence time"
- Beware that residence time defined for processes not quantities.

Michaelis-Menten Kinetics

$$R(x) = R_{\max}\frac{x}{k_m + x}$$

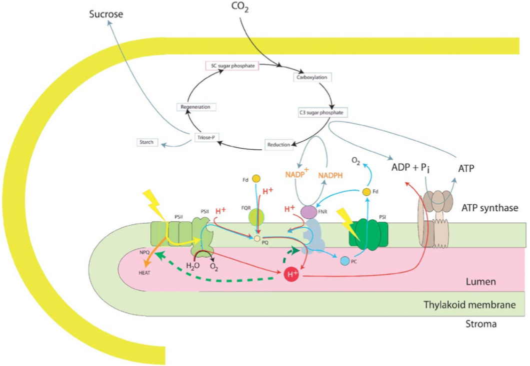
- R reaction rate, x concentration, $R_{\max}$ maximum rate, k_m half-saturation constant.
- Kinetics is the study of how fast things happen;
- Law for enzyme kinetics but widely applied elsewhere;
- when $x = k_m$ $R = \frac{1}{2}R_{\max}$.

Modeller's view of the Biosphere



GPP	Gross primary productivity (photosynthesis)
NPP	Net primary productivity (plant growth)
NEP	Net ecosystem productivity (undisturbed carbon storage) Some models start with NPP.
NBP	Net biome productivity (carbon sorage)

Mechanistic Photosynthesis Models



- Photons absorbed by chlorophyll, electrons transported and enzyme-catalysed fixation.
- Processes mutually limiting;
- CO_2 in leaves must diffuse from out*side through stomates;
- Also risks losing water so trade-off for plants.

Autotrophic Respiration

- Autotrophic means from itself;
- Staying alive and growing both cost plants energy;
- Energy generated from internal carbon;
- Usually modelled as fractions of GPP and fractions of growth;
- $\mathrm{NPP} = \mathrm{GPP} - \mathrm{r_A}$

Efficiency or Monteith Models

- Estimate NPP or GPP as proportional to some other flux, e.g. absorbed photons or transpired water
- Light-use efficiency model

$$\mathrm{NPP} = \epsilon_L \mathrm{APAR}$$

  where APAR absorbed photosynthetically active radiation and ϵ_L is light-use efficiency;
- Water-use efficiency model

$$\mathrm{NPP} = \epsilon_W L_E$$

  where L_E is latent heat flux (evaporation through leaves) and ϵ_W is water-use efficiency;
- Strictly holds for GPP so asumes NPP fixed fraction of GPP

Allocation of Growth

- Models may carry one term for live biomass or several, e.g. leaves, roots and stems;
- If they carry several then there must be a scheme to allocate carbon to each box;
- Usually fixed fractions that vary with type of plant.

The Fate of Biomass

- Leaves usually have a finite life so removal modelled like a residence time;
- Distinctly northern hemisphere bias to this formulation
- Roots treated similarly;
- Stems treated with statistical mortality term
- All fluxes go to soil carbon box or boxes.

Soil Carbon

- Often divided into “fast” and “slow” compartments;
- Heterotrophic respiration usually written as function of soil moisture and temperature;
- Functional forms and slopes matters of intense debate.

Summary Statistics

Name	Value Gt C/y			
	prior	1980–2000	1980–1990	1990–2000
GPP	135.70	134.80	134.30	135.30
Growth	23.50	22.35	22.31	22.39
Maint	44.04	72.70	72.13	73.28
NPP	68.18	40.55	40.63	40.46
Fast Resp.	53.83	27.40	27.60	27.21
Slow resp.	14.46	10.69	10.71	10.67
NEP	-0.11	2.453	2.318	2.587

Taken from BETHY model with parameters adjusted to fit observations (doi:10.1029/2004GB002254)

Summary

- Unsatisfyingly complicated but maybe nature is like that;
- Lots of Empirical parameters;
- Few fundamental constraints apart from conservation of mass and energy.

Motivation for OPTIC

- A test model for data assimilation that looked like biogeochemical models;
- Computationally simple enough to be very fast and easy to translate to many languages;
- Nonlinear enough to be interesting.
- trudinger et al. (2007,2008) (doi:10.1002/env.910, doi:10.1029/2006JG000367);

The OPTIC Model

$$\frac{\partial x_1}{\partial t} = f_i A_0 \frac{x_1}{A_1 + x_1} \frac{x_2}{A_2 + x_2} - A_3 x_1 + A_4$$
$$\frac{\partial x_2}{\partial t} = A_3 x_1 - A_5 x_2$$

- x_1 leaf biomass, x_2 other biomass;
- $A_1 \ldots A_5$ parameters;
- F forcing (light, rainfall etc).

First Equation

$$\frac{\partial x_1}{\partial t} = F_i A_0 \frac{x_1}{A_1 + x_1} \frac{x_2}{A_2 + x_2} A_3 x_1 + A_4$$

- First term product of two Michaelis-Menten rates (e.g. growth increases with amount of leaves and of substrate);
- 2nd term is linear decay;

second Equation

$$\frac{\partial x_2}{\partial t} = A_3 x_1 - A_5 x_2$$

- Inflow to x_2 is outflow from x_1 (imagine leaf litter);
- Outflow from x_2 first-order decay (imagine soil respiration).

Summary

- "Model" is a word with many meanings;
- They are usually imperfect even if underlying theory is complete;
- There are many common forms because principles like mass conservation and thermodynamics are everywhere;
- Interesting models can be constructed from fairly simple combinations;