1006: Ideas in Geography
Environmental Modelling: II

Dr. Mathias (Mat) Disney

UCL Geography

Office: 113 Pearson Building

Tel: 7670 0592

Email: mdisney@ucl.geog.ac.uk

www.geog.ucl.ac.uk/~mdisney

Last week

- Why model?
- **Improve process / system understanding**
  - by attempting to describe important aspects of process/system
- Types of model and examples….

Today

- Other reasons we model
- How do we know if a model is any good?
- Further distinctions between empirical and physical modelling
- Forward and inverse modelling

Why modelling (II)?

Derive / map information through surrogates

e.g.:

REQUIRE	**spatial distribution of biomass**
DATA	spatial observations of reflectance from a satellite
MODEL	**model relating reflectance to biomass**

Crop biomass map ??

Soil moisture map ??

Why Modelling (III)?

Make past / future predictions from current observations (extrapolation)

tend to use ‘physically-based’ models

e.g.:

short term:

weather forecasting, economic models

longer term:

climate modelling, demographics

Interpolation based on limited sample of observations

e.g.:

- vegetation / soil surveys
- Temperatures, rainfall
- political surveys

How useful are models?

- Models are based on a set of assumptions

  ‘As long as assumptions hold’, should be valid

- When developing model
  - <u>Must</u> define & understand assumptions and state these explicitly
- When using model
  - <u>Must</u> understand assumptions (and limitations) & make sure model is relevant

How do we know how ‘good’ a model is?

- Ideally, **‘validate’** over wide range of conditions

For environmental models, typically:

  - characterise / measure system
  - compare model predictions with measurements of ‘outputs’
    - noting error & uncertainty

‘Validation’: essentially - how well does model predict outputs when driven by measurements?

How do we know how ‘good’ a model is?

For environmental models, often difficult to achieve

- can’t make (sufficient) measurements
  - highly variable environmental conditions
  - prohibitive timescale or spatial sampling required
- systems generally ‘open’
  - no control over all interactions with surrounding system
  - E.g. atmosphere, ocean for climate
- use:
  - ‘partial validations’
  - sensitivity analyses

How do we know how 'good' a model is?

'partial validation'

- compare model with other models
- analyse sub-components of system
  - e.g. with lab experiments

sensitivity analyses

- vary each model parameter to see how sensitive output is to variations in input
  - build understanding of:
    - model behaviour
    - response to key parameters
    - parameter coupling i.e. interdependence of parameters

Types of model revisited

- **Statistical / empirical**
  - ‘calibration model’
  - Based on data
  - Not derived using theory or physical laws
- **Physically-based**
  - model <u>physics</u> of interactions
  - in Geography, also used to include many empirical models, if it includes some aspect of physics (c.f. hydrological models)

Types of model revisited

- **deterministic**
  - relationship **a**=f(**b**) is always same
    - no matter when, where calculate it
- **stochastic**
  - exists element of randomness in relationship
    - repeated calculation gives different results
    - E.g. model of coin toss using random numbers
    - Sequence of H, T different in each case but p(H|T) always approaches 0.5

Forward and inverse modelling

<u>Very important distinction</u>

- **forward model**
  - **a**=f(**b**)
  - measure **b**, use model to predict **a**
- **inverse model**
  - **b**=f^{-1}(**a**)
  - measure **a**, use model to predict **b**

- Inversion allows us to derive values for model parameters from observations of system we are modelling
  - **<u>one of main reasons we develop physically-based models</u>**

Forward and inverse modelling: examples

- Forward: reflectance = f(leaf area)
- Inverse: leaf area = f^{-1}(reflectance)
- Forward: pollen count record = f(T_{past})
- Inverse: T_{past} = f^{-1}(pollen count record)
- Forward: Population change = f(Births,Deaths)
- Inverse: B,D = f^{-1}(population change)

Statistical / empirical models

- **Basis:** simple theoretical analysis or empirical investigation gives evidence for relationship between variables
  - Eg a scatter plot with a trend
  - Basis is generally simplistic or unknown, but general trend seems predictable
- Using this, a statistical relationship is proposed

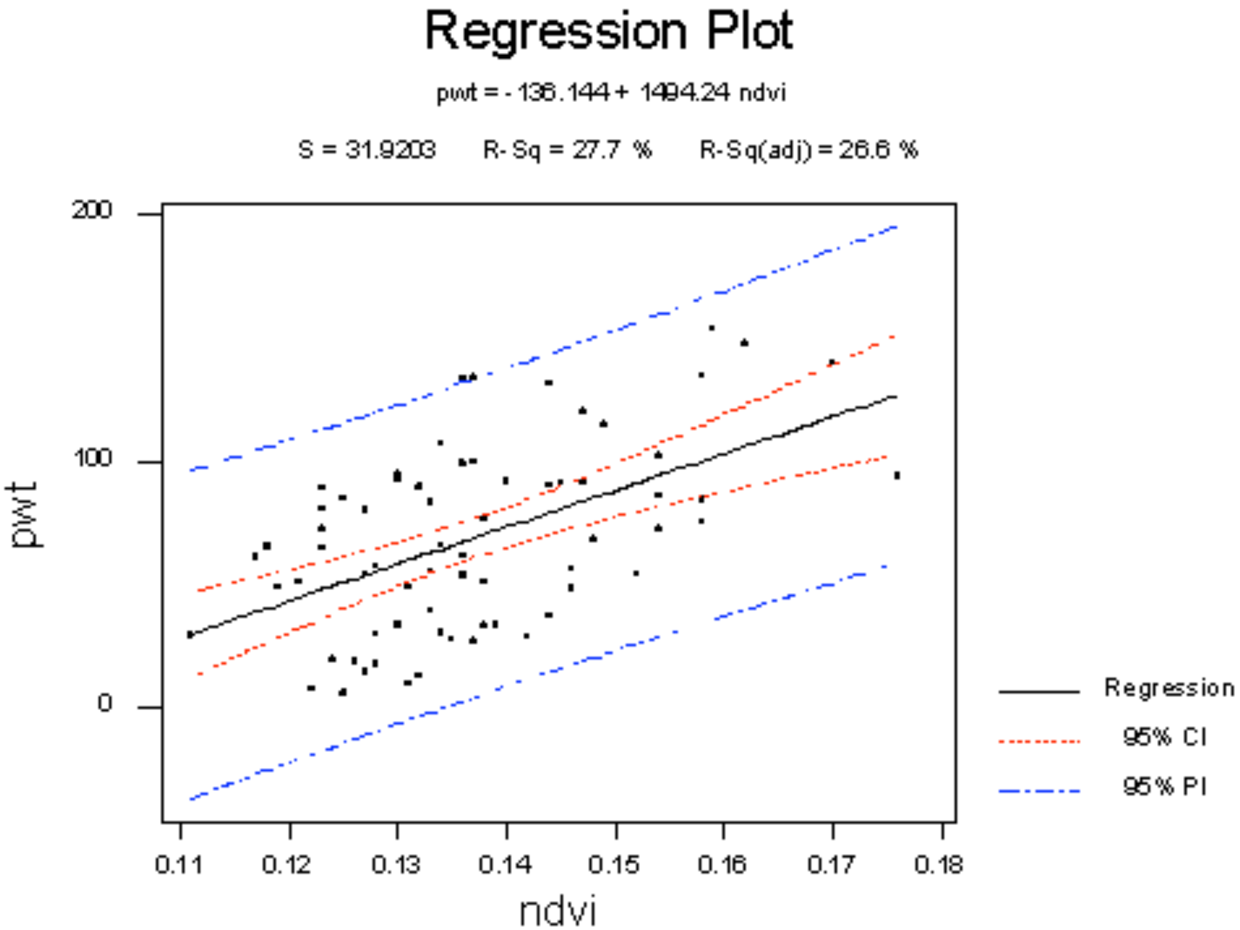
Statistical / empirical models

E.g.:

- From observation & basic theory, we observe:
  - Satellite images in red and near infra-red seem to correlate with the amount of vegetation
  - Calculate NDVI (normalised difference vegetation index)

"NDVI"

Make measurements to investigate the trend further:-



http://sevilleta.unm.edu/research/local/plant/tmsvinpp/documents/sevsymp2001/sevsymp2001_files/v3_document.htm

Statistical / empirical models

- Propose **linear** relationship between vegetation amount (biomass) and NDVI
- Calibrate model coefficients (slope, intercept)

- **Biomass (g/m^2) = -136.14 + 1494.2*NDVI**

  - This is a very simple model
  - Model fit apparently "reasonable" (R^2 = .27)

http://sevilleta.unm.edu/research/local/plant/tmsvinpp/documents/sevsymp2001/sevsymp2001_files/v3_document.htm

Statistical / empirical models

Dangers:

- changing environmental conditions (or location)
  - lack of generality
  - Model developed using a single specific dataset, single location, single time, etc…
  - Limited regime of validity (small range of variables)
- Have not accounted for all important variables
  - tend to treat as ‘uncertainty’ (spread away from linear fit)
- Does not contain any “understanding” so doesn’t allow us to understand system

‘Physically-based’ e.g.: population growth

- Developed using laws or understanding
- e.g. Simple population growth
- Require:
  – model of population Q over time t
- Theory:
  – in a ‘closed’ community, population change given by:
    - increase due to births (the birth rate)
    - decrease due to deaths (the death rate)
  – over some given time period δt

‘Physically-based’ models

- We can state this problem:-

$$Q(t+\delta t) = Q(t) + births(t) - deaths(t)$$

- If model parameters B, D (birth and death rate) constant and ignoring age/sex distribution and environmental factors e.g. food, disease etc. then…..

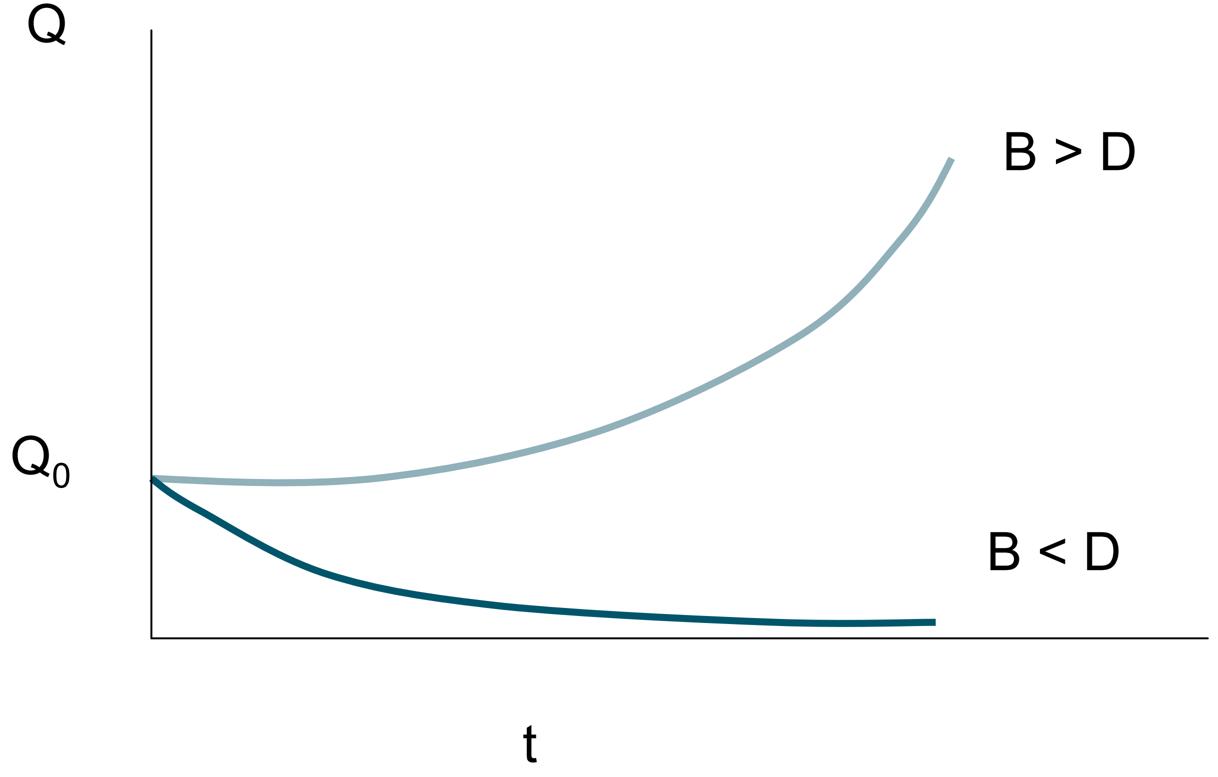
‘Physically-based’ models

$$Q(t) = Q_0 e^{(B-D)t}$$

- Note all our assumptions
  - Only births and deaths matter (nothing else!)
  - B, D constant
- Can make this model much more complicated:-
  - B, D depend on time, and on place
  - Stochastic effects (war, disease, migration etc)

Model predictions?

$$Q(t) = Q_0 e^{(B-D)t}$$



Other model type distinctions: **Analytical / Numerical**

- Analytical
  - Can actually “write down” equations governing models (formulae, paramters etc.)
  - e.g. biomass = a + b*NDVI
  - e.g.

$$\frac{dQ}{dt} = aQ \longrightarrow Q = Q_0 e^{at}$$

Practically, and especially in environmental modelling, always need to consider:

- **uncertainty**
  - in measured inputs
  - in model
  - and so **should have distribution of outputs**
    - Eg climate prediction from models
    - Try to attach an uncertainty to give a feel for how “trustworthy” the prediction is (eg increase in temperature of 5C ±0.1 or 5C ± 30?!)
- **scale**
  - different relationships over different scales
    - principally consider over **time / space**

Summary: which type of model to use?

- Statistical/empirical
  - **advantages**
    - Simple, quick to calculate
    - require little / no knowledge of underlying (e.g. physical) principles
    - (often) easy to invert as have simple analytical formulation
  - **disadvantages**
    - only appropriate to limited range of parameters & under limited observation conditions
    - validity in extrapolation difficult to justify
    - does not improve general understanding of process

Which type of model to use?

- Physical/Theoretical/Mechanistic
  - **advantages**
    - applicable to wider range of conditions
    - use of numerical solutions (& fast computers) allow great flexibility in modelling complexity
    - may help to understand process
      - e.g. examine role of different assumptions
  - **disadvantages**
    - more complex models require more time to calculate
      - get a faster computer!
    - Need to know/include all important processes and variables
    - often difficult to obtain analytical solution hence hard to invert

Reading

Basic texts

- Barnsley, M. J., 2007, *Environmental Modelling: A Practical Introduction*, (Routledge). Excellent, practical introduction with many examples, and code using freely-available software.
- Kirkby, M.J., Naden, P.S., Burt, T.P. and Butcher, D.P. 1993 *Computer Simulation in Physical Geography*, (Chichester: John Wiley and Sons). Excellent introduction to EM illustrated by some straightforward, easy-to-run BASIC computer programs of environmental models.
- *Computerised Environmental Modelling: A practical introduction using Excel*, J. Hardisty et al., 1993, John Wiley and Sons.
- Haynes-Young, R. and Petch, J. 1986 *Physical Geography: its nature and methods*, (London: Harper Row). Very good on general philosophical issues associated with models and modelling.
- Goodchild, M.F., Parks, B.O. and Steyaert, L.T. 1993 *Environmental Modelling with GIS*, Oxford: Oxford University Press.
- Casti, John L., 1997 *Would-be Worlds* (New York: Wiley and Sons). A nice easy-to-read introduction to the concepts of modelling the natural world. Excellent examples, and well-written. A good investment.

Advanced texts

- Gershenfeld, N. , 2002, *The Nature of Mathematical Modelling*,, CUP.
- Boeker, E. and van Grondelle, R., *Environmental Science, Physical Principles and Applications*, Wiley.
- Monteith, J. L. and Unsworth, M. H., *Principles of Environmental Physics*, Edward Arnold.