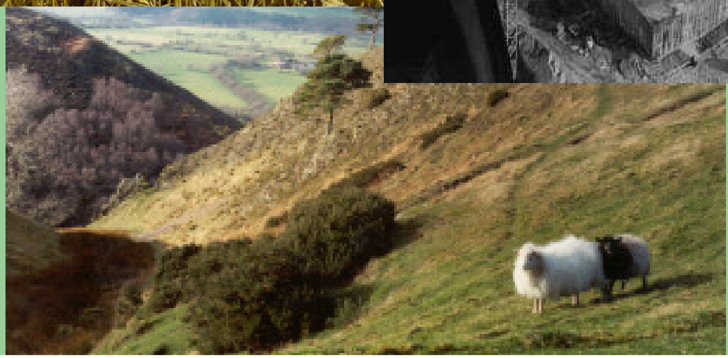
Model reduction and evaluation: methane emissions from soils





Neil Crout

Environmental Science

School of Biosciences

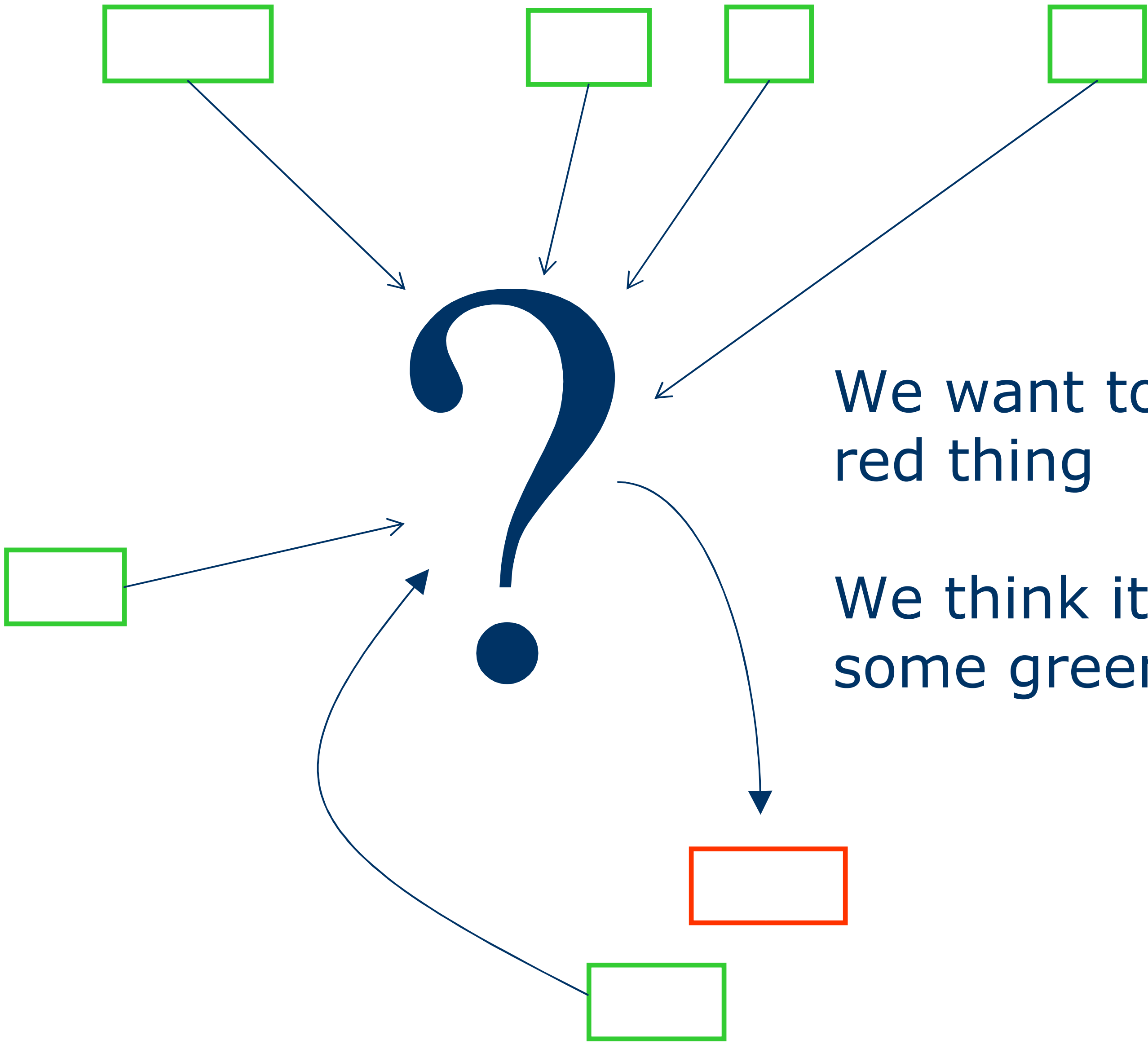
University of Nottingham

Co-workers

- Davide Tarsitano
- Glen Cox
- James Gibbons
- Andy Wood
- Jim Craigon
- Steve Ramsden
- Yan Jiao
- Tim Reid

- with thanks to BBSRC and Leverhulme

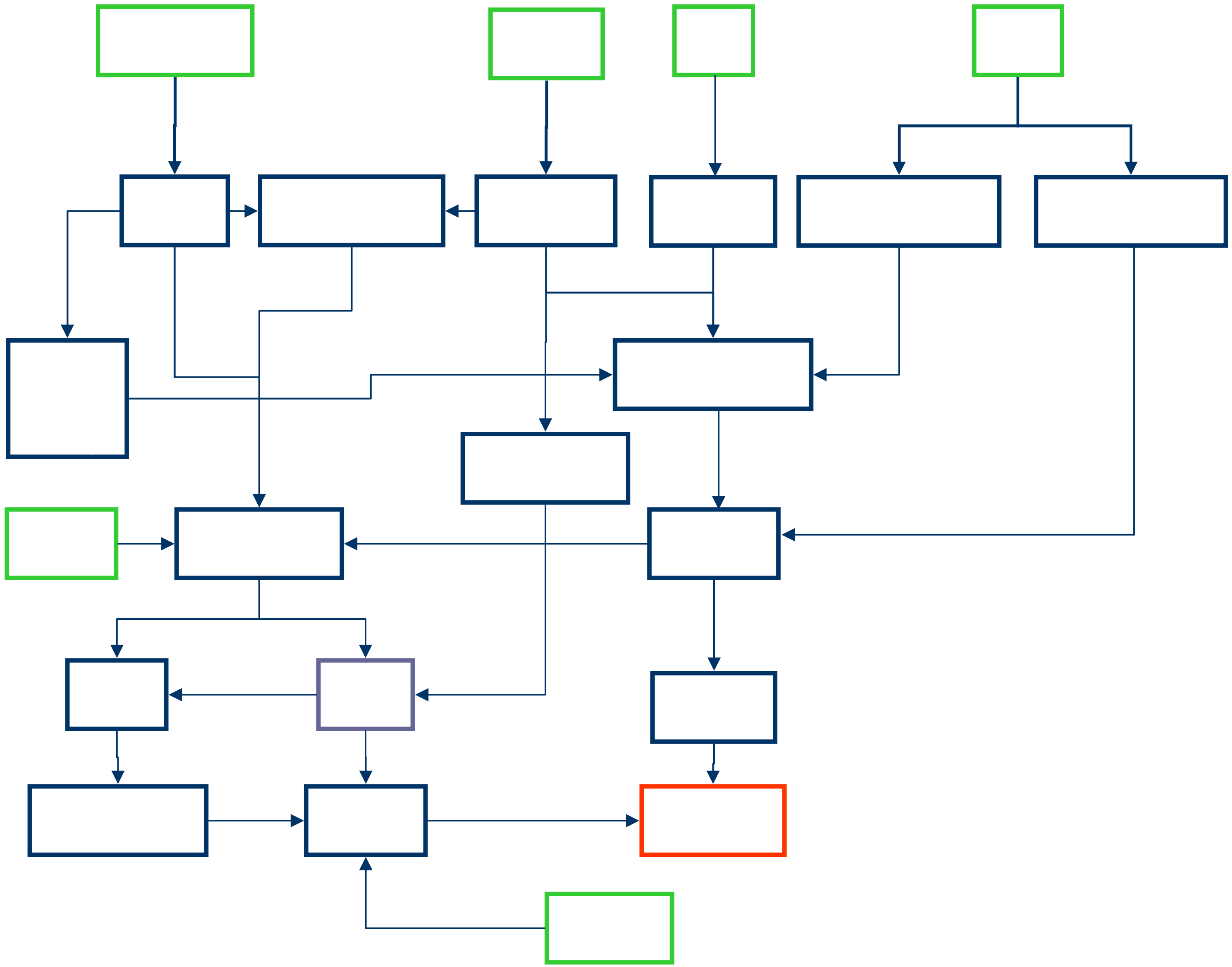
A ‘Typical’ Starting Point



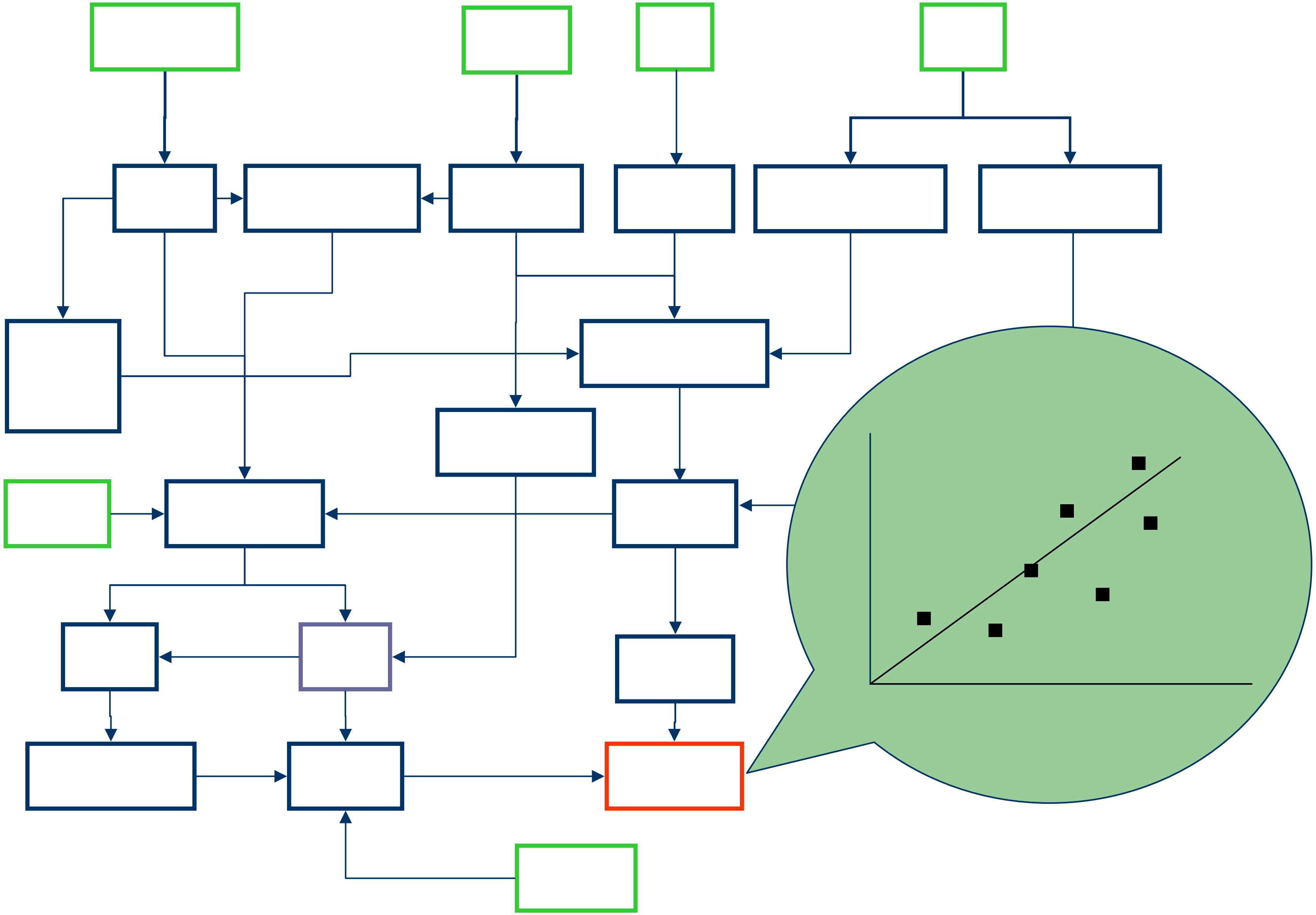
We want to know a red thing

We think it depends on some green things

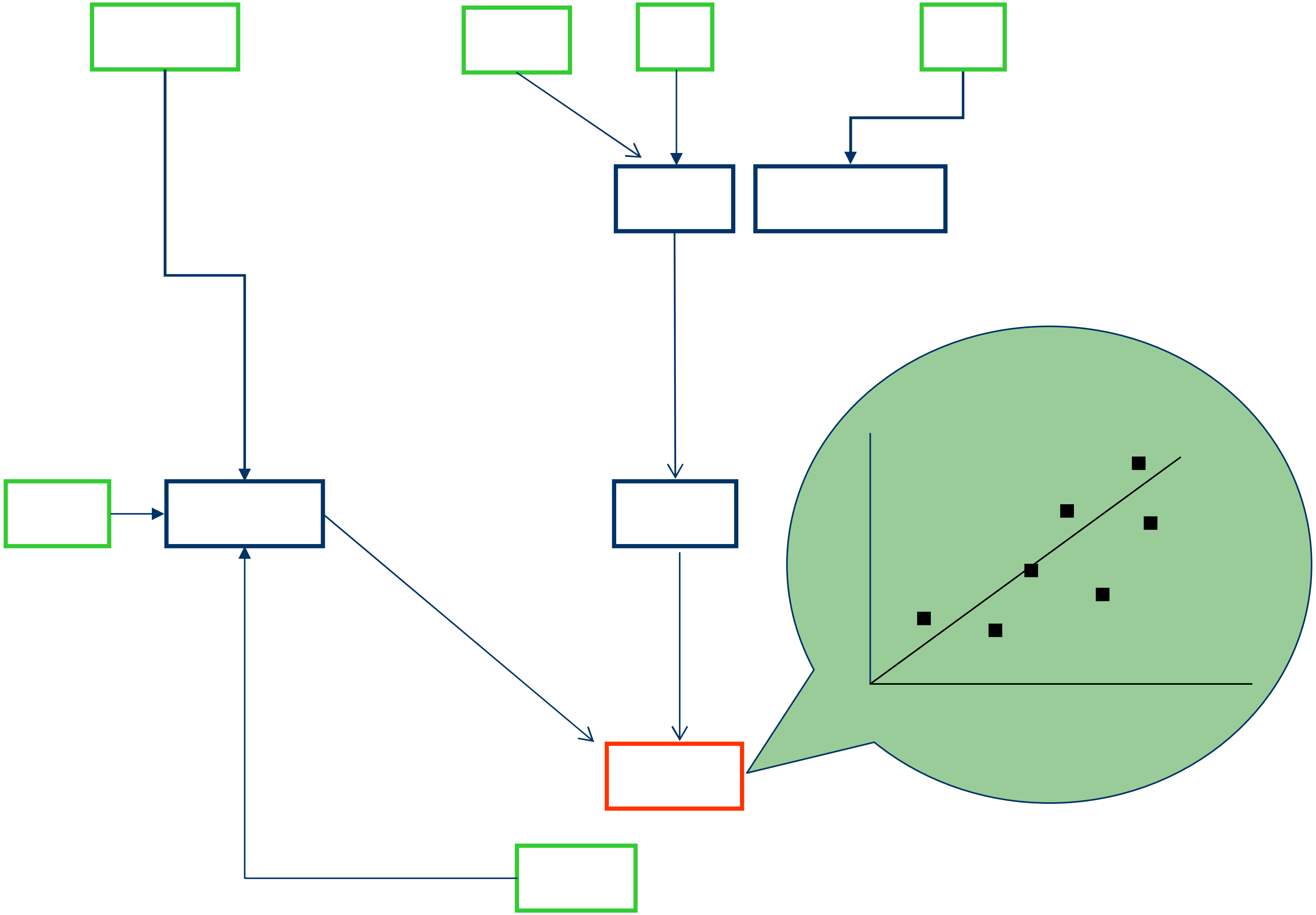
A ‘Typical’ Model



A ‘Typical’ Test?



But would this have worked?...



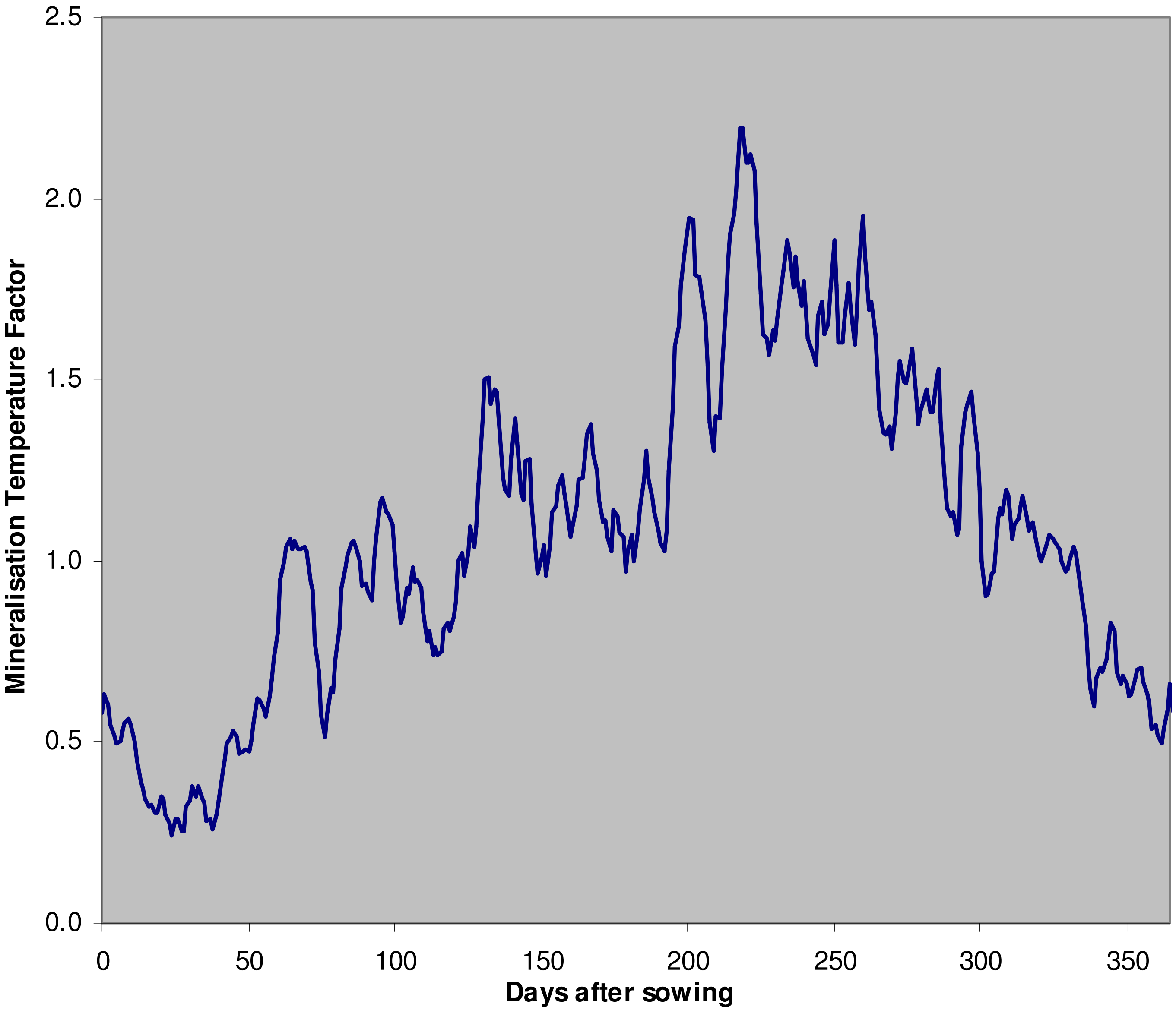
Rationale for Model Reduction

- Want to assess whether a model has the most appropriate level of detail
- We can get some idea on this by comparing models of the same system which have different levels of detail
- But, we don’t have lots of different models
- So we consider reducing the model we have to create alternative model formulations which can be compared
- Comparison is restricted
  - reduced models are drawn from the same source
  - but hopefully better than no comparison at all

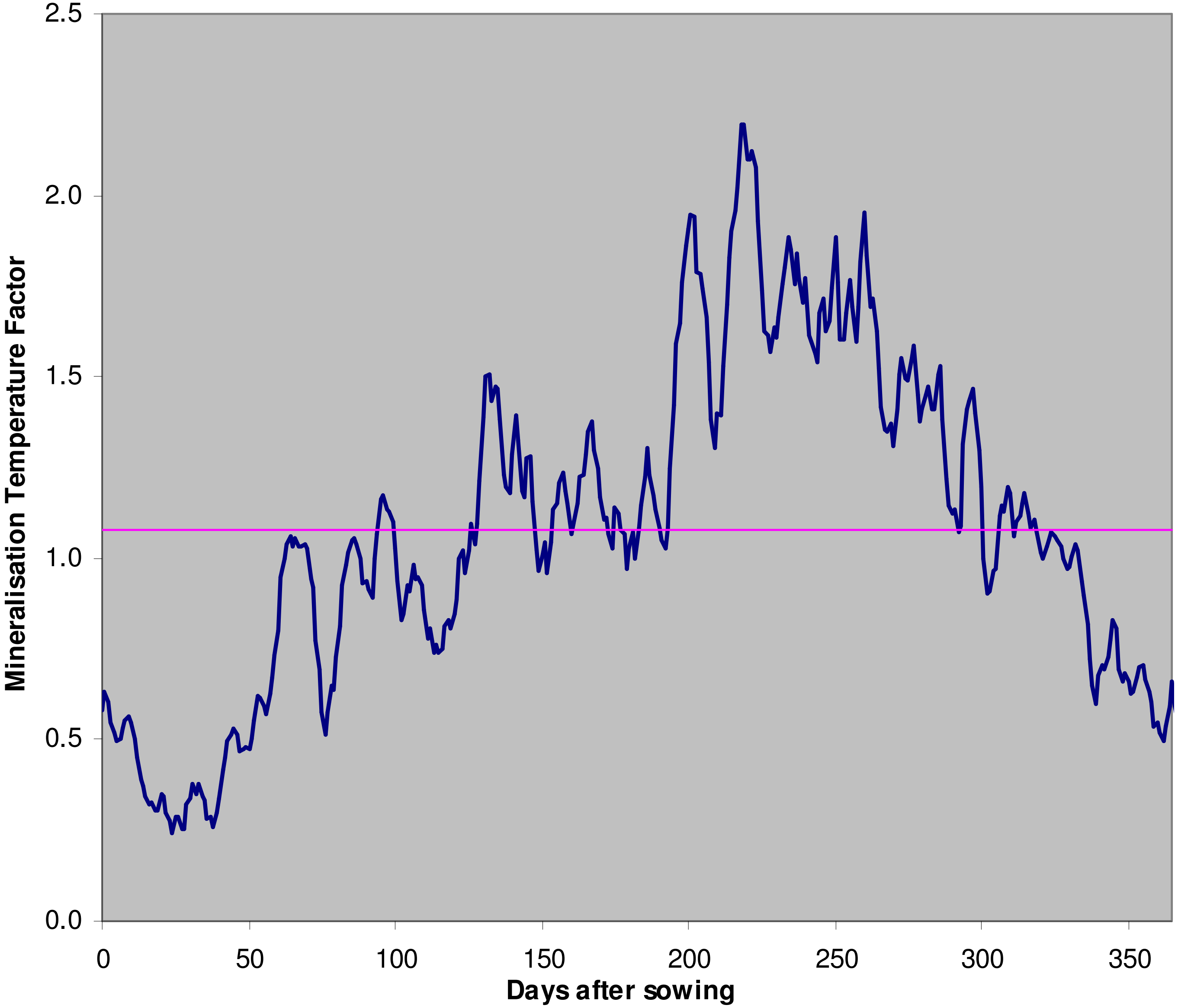
Model Reduction – How?

- Start with a ‘full’ model of a system
- Reduce it (systematically, automatically)
- Producing a ‘set’ of alternative model formulations for the same system
- Assess model ‘performance’ by comparing to observation
- Reduction?
  - Replace a model variable by zero
    - e.g. ignore diffusion
  - But often inter-connected nature of typical models means can’t simply leave things out
  - Replace a model variable by an alternative formulation
    - e.g. Michalis-Menten becomes linear
  - Replace variables with a constant
    - Simple case the mean or median that the variable attains in the full model
    - More sophisticated ‘integrate’ over range of values

Reduction by variable replacement



Reduction by variable replacement

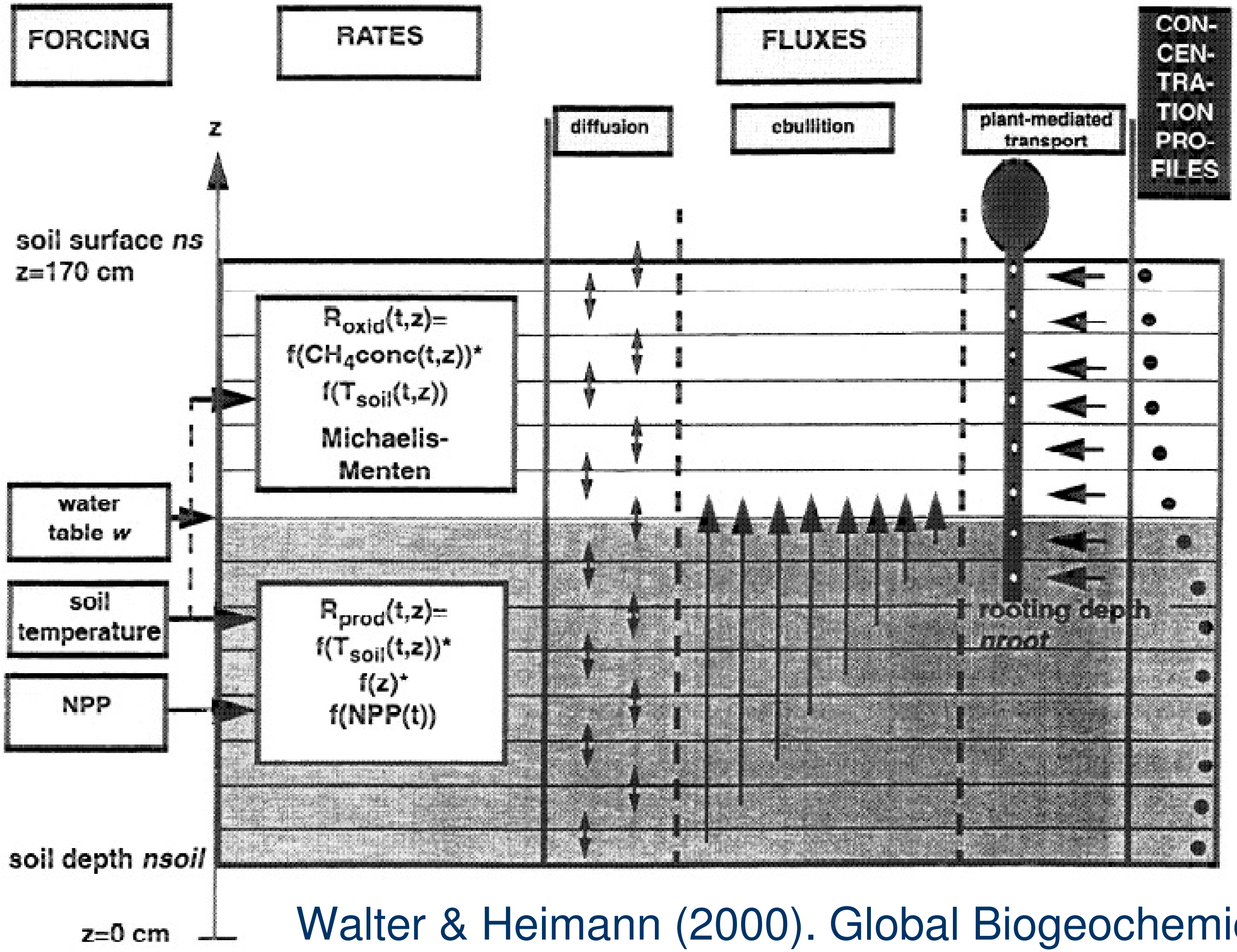


Search the ‘replacement space’

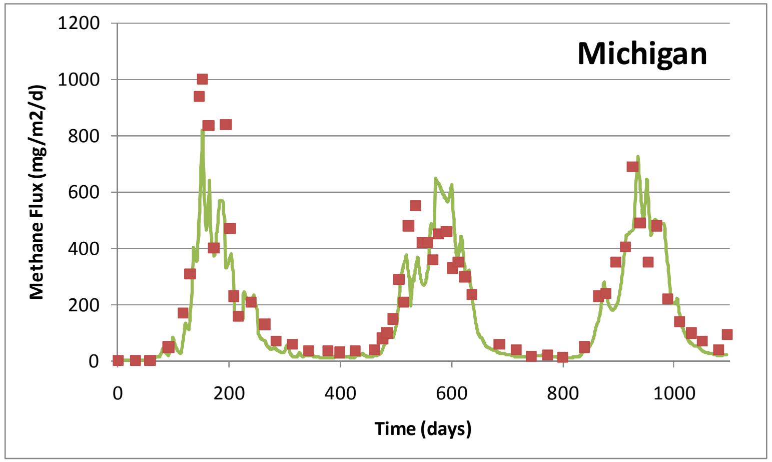
- Search the possible combinations of replacement
- Example case: model with 3 variables to be replaced
- For each one, test the ‘performance’ of the model (compare to observations)
- Various methods for searching, such as
  - Exhaustively search all combinations
  - MCMC walk through the combinations of possible models discrete space, can be quite efficient
  - Screening step may be useful

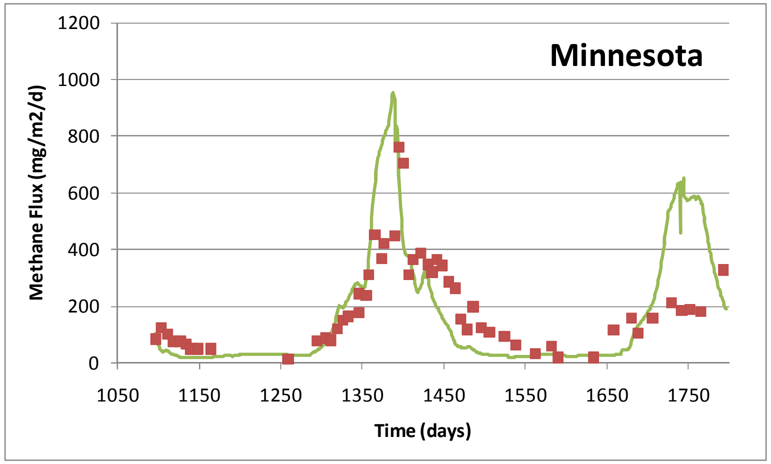
	V1	V2	V3
Full Model	☑	☑	☑
1	☒	☑	☑
2	☑	☒	☑
3	☑	☑	☒
4	☒	☒	☑
5	☒	☑	☒
6	☑	☒	☒
7	☒	☒	☒

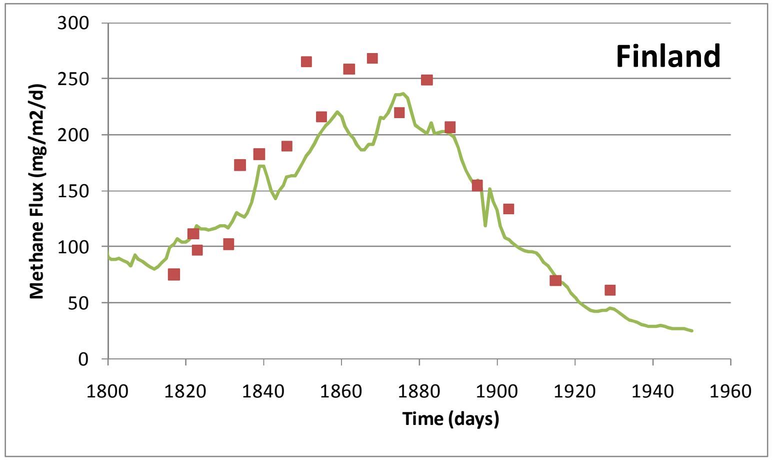
Methane emission from wetlands

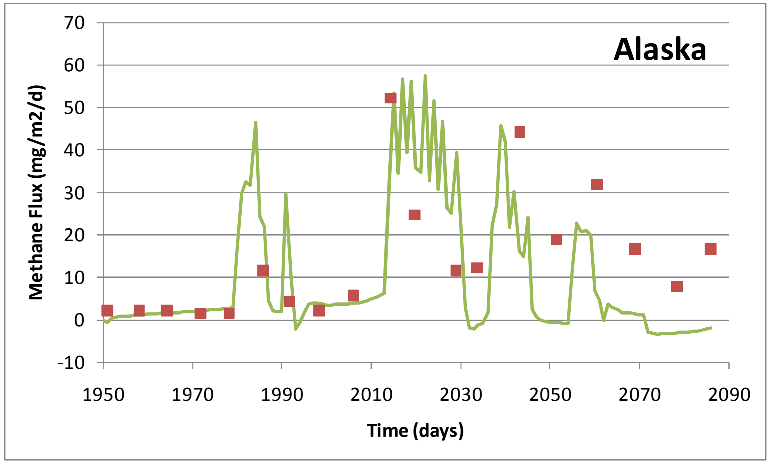


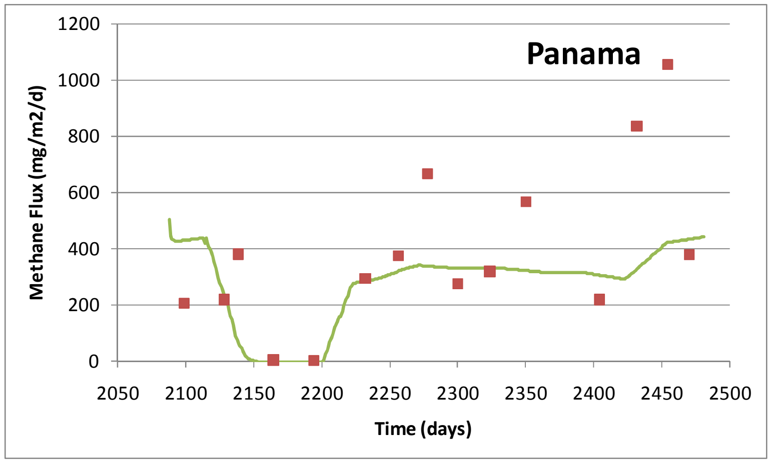
Walter & Heimann (2000). Global Biogeochemical Cycles, 14: 745-765









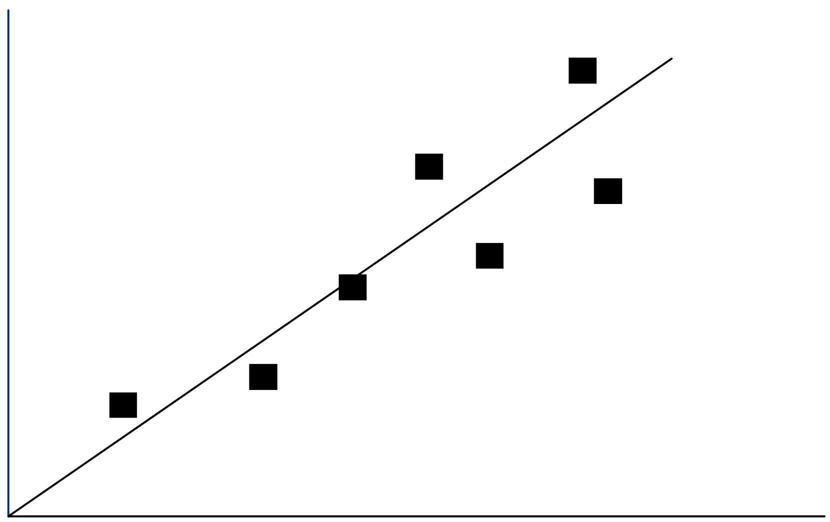


‘Nottingham implementation’ of the model compared to the 5 sites used by Walter and Heimann (2000)

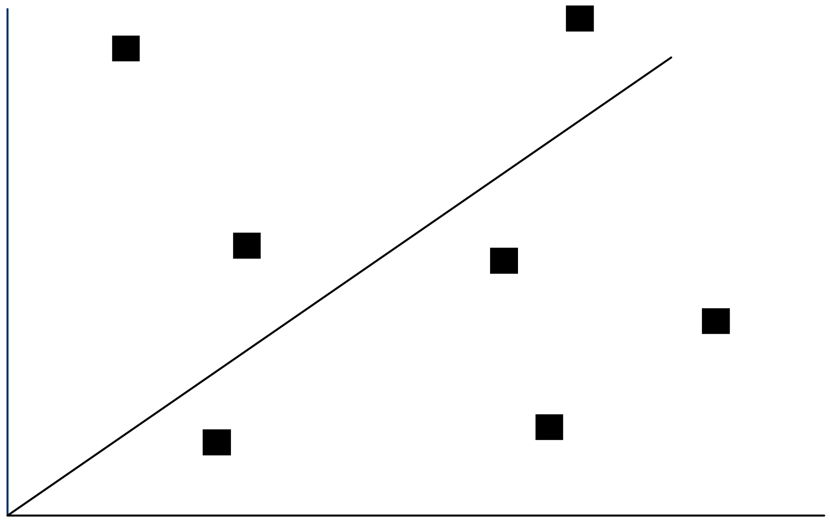
Candidate reduction ‘variables’

Forcing	
Water table	reduced to a constant
Temperature	constant
Rates	
Oxidation	First order with methane concentration
	Linear with temperature
	Independent of temperature
Production	Linear with temperature
	Independent of temperature
	Constant vertical distribution of organic matter
	Constant factor for seasonal variation in NPP
Fluxes	
Bubbles	Ignored completely
Diffusion	Ignored completely
Plant-mediated	Ignored Completely
transport	Constant temperature dependent growth factor
	Constant vertical root density factor

- Compare model to observations
- ‘Probability’ for each replacement combination



Relatively high probability



Relatively low probability

Interpretation: Reduction 'Probabilities'

	V1	V2	V3	Model 'Probability'
Full Model	☑	☑	☑	0.10
1	☒	☑	☑	0.05
2	☑	☒	☑	0.40
3	☑	☑	☒	0.00
4	☒	☒	☑	0.45
5	☒	☑	☒	0.00
6	☑	☒	☒	0.00
7	☒	☒	☒	0.00
Reduction 'Probability'	0.50	0.85	0.00	

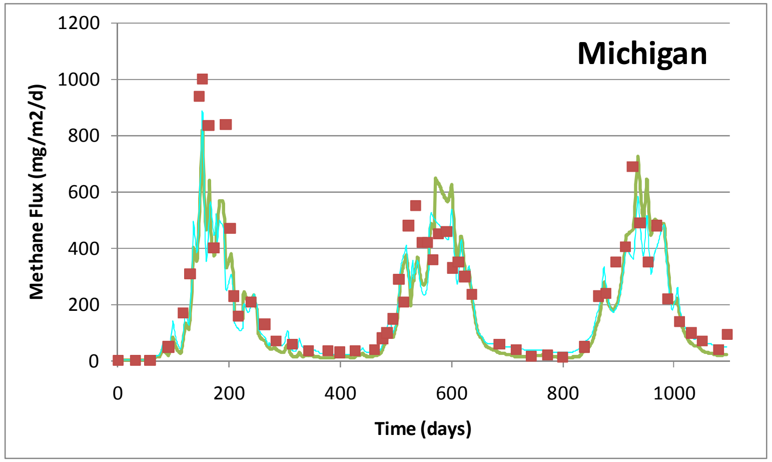
Redundant 'Noisy' Required

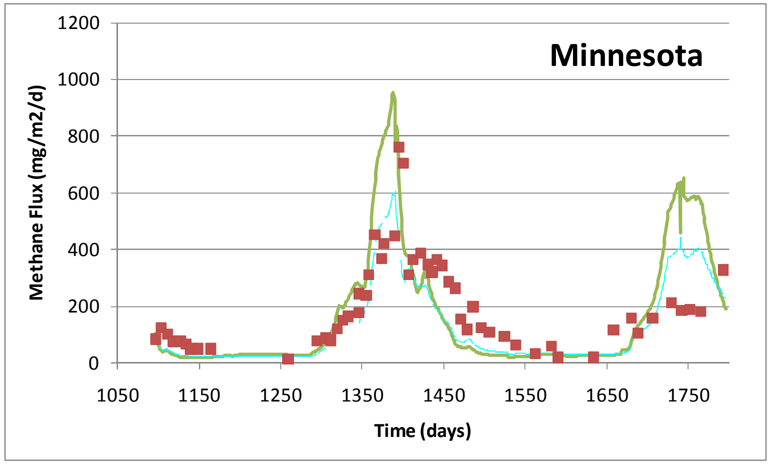
Forcing	
Water table	reduced to a constant
Temperature	constant
Rates	
Oxidation	First order with methane concentration
	Linear with temperature
	Independent of temperature
Production	Linear with temperature
	Independent of temperature
	Constant vertical distribution of organic matter
	Constant factor for seasonal variation in NPP
Fluxes	
Bubbles	Ignored completely
Diffusion	Ignored completely
Plant-mediated	Ignored Completely
transport	Constant temperature dependent growth factor
	Constant vertical root density factor

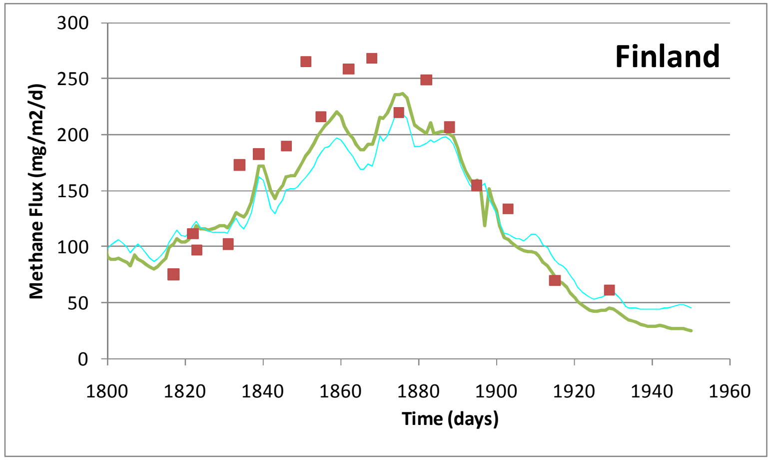
Required	Required at some sites	Indifferent	'Noisy'

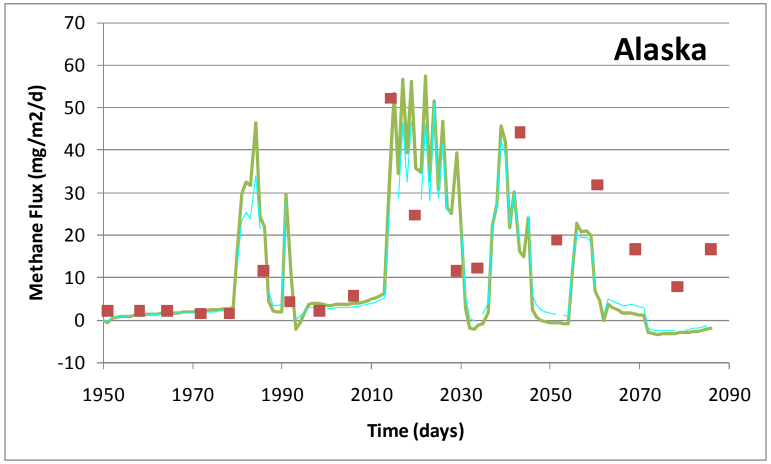
‘Minimum’ Methane Model

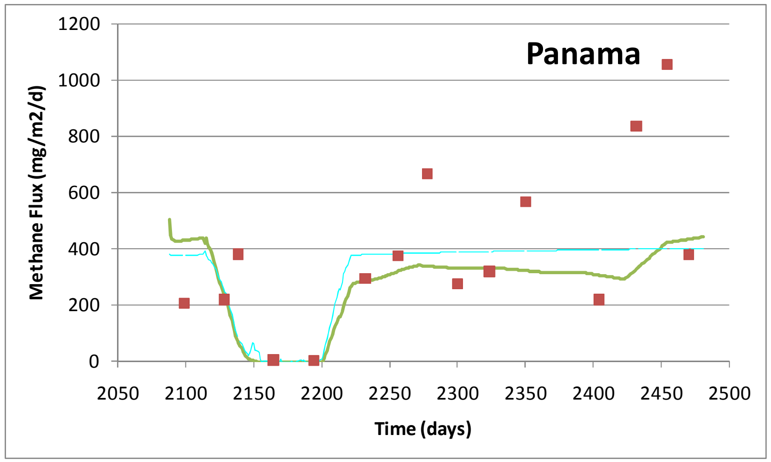
- Variable water table
- Temperature dependent production in saturated zone
- Oxidation rate in aerobic zone dependent on methane concentration
- Plant mediated transport dependent on methane concentration
- Diffusion required, especially if site has ‘rootless’ vegetation
- Important limitation of this analysis
  - Not considered the implications for ‘global’ parameterisation (Walter et al, JGR, 2001)











Comparing the full model (▬▬)
With reduced model (▬▬) where
'noisey' and redundant variables
replaced (except diffusion)

Other applications…

- Applied the same/similar approach to other models, e.g.
  - Marine Ecology Models (carbon cycling)
  - Soil-plant radiocaesium models
  - FARM-ADAPT – very large farm management model
  - Wheat simulation model
  - Soil carbon nitrogen ecosystem model
- Noisy and redundant variables are ubiquitous

More broadly...

- Model formulation is usually uncertain
- Model reduction provides a way to test (brutally) model formulation
- Aim is to test the ideas that make the models
  - Not to test the model as whole

- …all models are wrong, some bits of models are useful…

Useful references?

- NMJ Crout, D Tarsitano, AT Wood (2009). Is my model too complex? Evaluating model formulation using model reduction. Environmental Modelling & Software, 24:1-7.
- JM Gibbons, GM Cox, ATA Wood, J Craigon, SJ Ramsden, D Tarsitano, NMJ Crout (2008). Applying Bayesian Model Averaging to mechanistic models: an example and comparison of methods. Environmental Modelling & Software, 23:973-985.
- Cox GM, Gibbons JM, Wood ATA, Craigon J, Ramsden SJ, Crout NMJ (2006). Towards the systematic simplification of mechanistic models. Ecological Modelling 198:240-246.
- Bernhardt, K. 2008. Finding alternatives and reduced formulations for process based models. Evolutionary Computation 16:1-16. *Alternative approach*.
- Asgharbeygi, N., Langley, P., Bay, S., Arrigo, 2006. Inductive revision of quantitative process models. Ecol. Mod., 194:70-79. *Alternative approach to model reduction, albeit without a statistical framework*
- Brooks RJ, Tobias AM (1996) Choosing the best model: level of detail, complexity, & model... Mathl. Comput. Mod. 24:1-14. *Interesting article on model utility*
- Anderson, T.R., 2005. Plankton functional type modelling: running before we can walk? J. Plankton Research 27:1073-1081. *Nice discussion of model complexity in context of marine systems. Quite controversial, replies to the replies to the replies.*
- Jakeman, A.J, Letcher, R.A., Norton, J.P., 2006 Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software 21, 602-614. *Very nice discussion on what is good practice in model development*
- Myung, J., Pitt, M. A., 2002. When a good fit can be bad. Trends. Cogn. Sci., 6:421-425. *Good introduction to model selection criteria*
- Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach. 2nd ed. New York: Springer-Verlag. *Ultimate reference, but not without critics.*