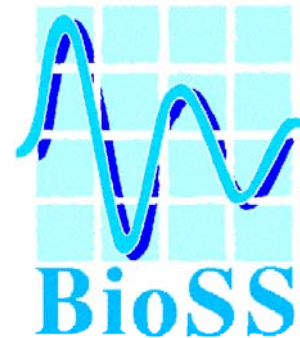


# Bayesian analysis of non-linear differential equation models applied to a gut microbial system

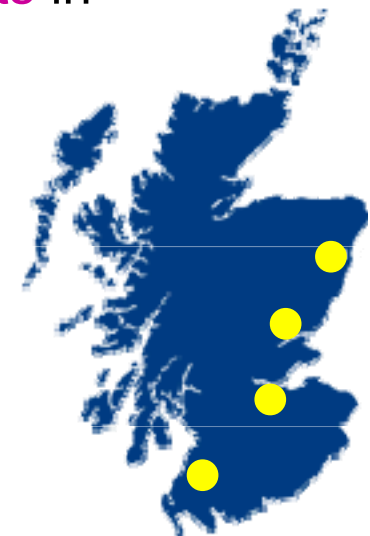
Grietje Holtrop & Daniel Lawson



Biomathematics & Statistics Scotland  
Rowett Institute of Nutrition and Health  
Bucksburn, Aberdeen

- ~ 35 members of staff  
(Statisticians, modellers, bioinformaticians, support staff)
- Funding from Scottish Government to support biological sciences, plus external funding
- Staff located at various biological **research establishments** in Scotland (Aberdeen, Dundee, Edinburgh, Ayr)
- Provision of
  - Consultancy, Research, Training

I am based at the Rowett Institute of Nutrition and Health

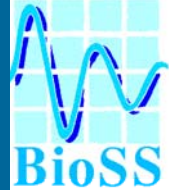


# Digestion of food

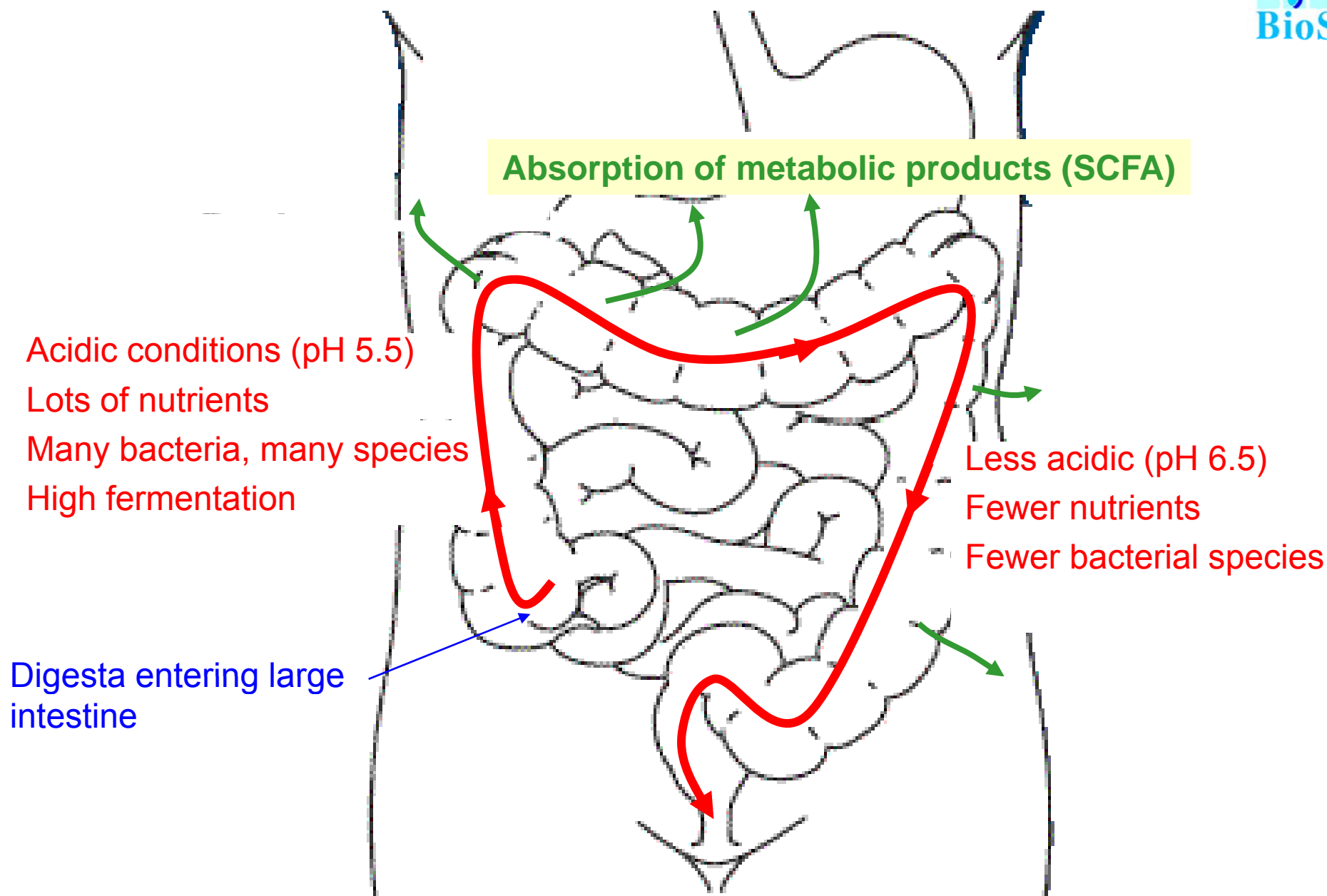


- Food is absorbed in the small intestine
- Undigested food (such as fibres) reach large intestine
  - Home to gut bacteria ( $10^{11}$ /g dry matter, several hundreds species)
  - Undigested food particles are fermented into metabolic products such as Short Chain Fatty Acids (SCFA)
  - These are then absorbed by the host to provide energy (up to 10% of total energy supply)
- Remainder leaves body as faeces

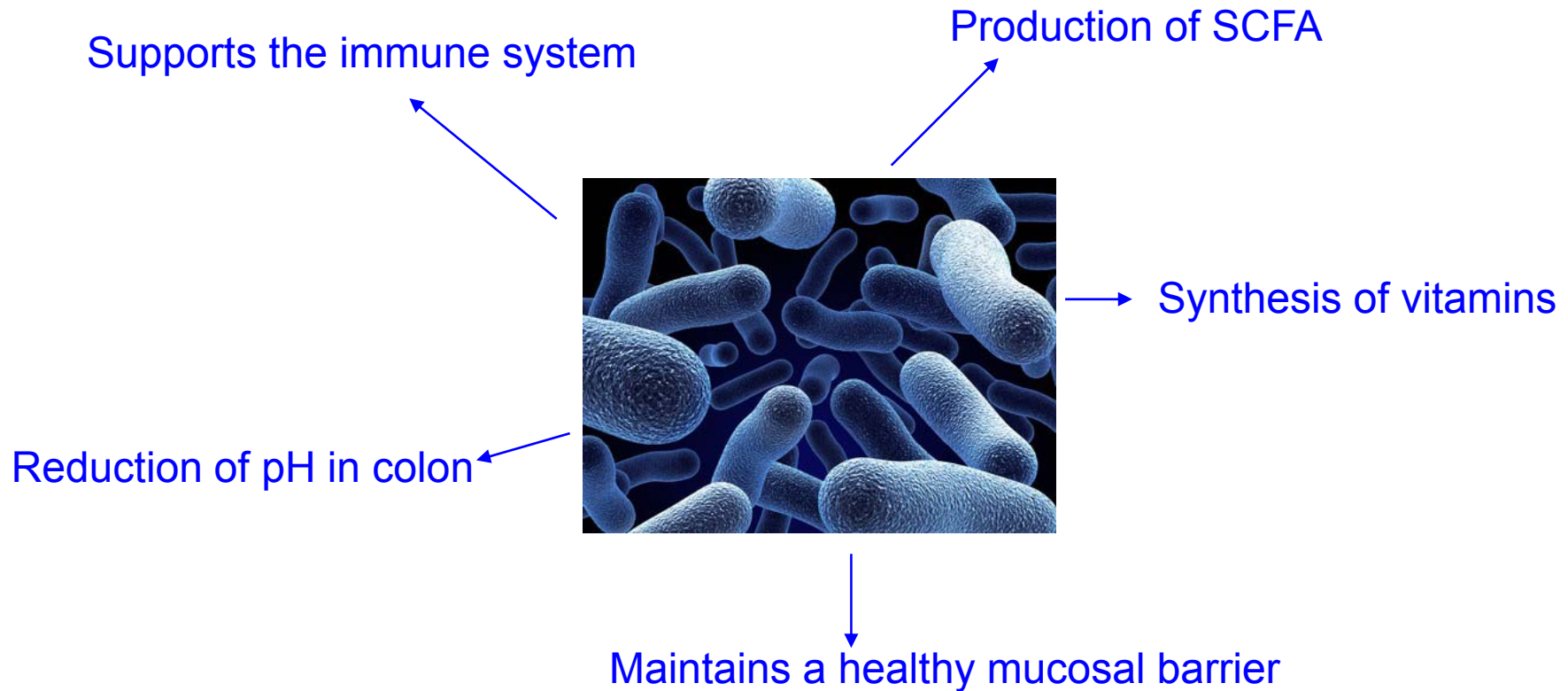
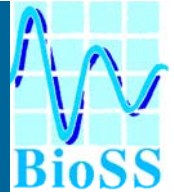
# Digestion of food



- Food is absorbed in the small intestine
- Undigested food (such as fibres) reach large intestine
  - Home to gut bacteria ( $10^{11}$ /g dry matter, several hundreds species)
  - Undigested food particles are fermented into metabolic products such as Short Chain Fatty Acids **SCFA**
  - These are then absorbed by the host to provide energy (up to 10% of total energy supply)
- Remainder leaves body as faeces
- Health aspects:
  - Butyrate is thought to play protective role against colon cancer
  - Lactate associated with gut disorders such as Crohn's disease
  - Irritable bowel syndrome, inflammatory bowel disease



# Healthy gut

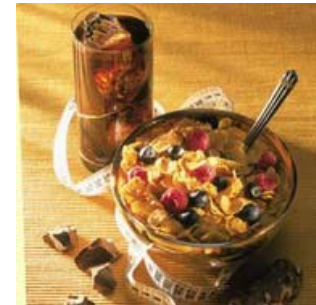


# Influencing gut health via nutrition



Can we create a 'healthy' environment for gut bacteria?

- Prebiotics such as cereals containing oats to stimulate growth of 'good' bacteria
- Probiotics such as yoghurts containing live biocultures to add 'good' bacteria to gut microbiota



# Influencing gut health via nutrition



Can we create a 'healthy' environment for gut bacteria?

- Prebiotics such as cereals containing oats to stimulate growth of 'good' bacteria
- Probiotics such as yoghurts containing live biocultures to add 'good' bacteria to gut microbiota



- For such approaches to be targeted and successful, **knowledge** is needed:

**Preferred substrates**

**Environmental conditions (pH)**

**Competition between species, interactions**

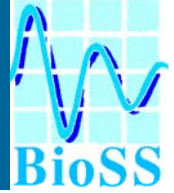


# Experimental studies



- In vivo: feed volunteers a breakfast cereal and collect faecal samples

# Experimental studies



- In vivo: feed volunteers a breakfast cereal and collect faecal samples
- In vitro – simple (batch cultures using 9 ml tubes)
  - tube with **one known bacterial strain** and a substrate of interest, monitor bacterial growth and production of metabolites
  - tube with **several known bacterial strains** and a substrate of interest, monitor bacterial growth and production of metabolites
  - tube with **mixed bacterial population** (from faecal sample) and a substrate of interest, monitor production of metabolites



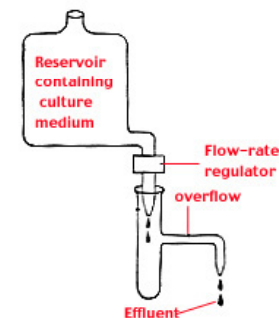
# Experimental studies



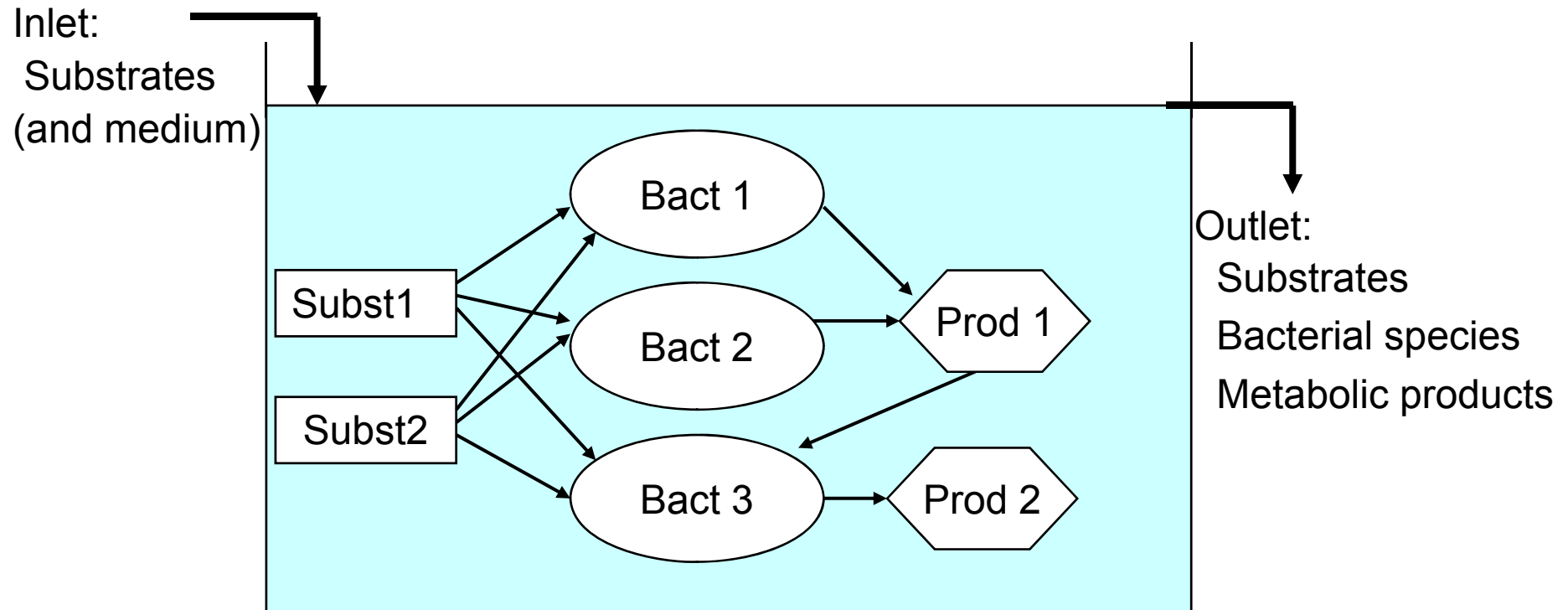
- In vivo: feed volunteers a breakfast cereal and collect faecal samples
- In vitro – simple (batch cultures using 9 ml tubes)
  - tube with **one known bacterial strain** and a substrate of interest, monitor bacterial growth and production of metabolites
  - tube with **several known bacterial strains** and a substrate of interest, monitor bacterial growth and production of metabolites
  - tube with **mixed bacterial population** (from faecal sample) and a substrate of interest, monitor production of metabolites



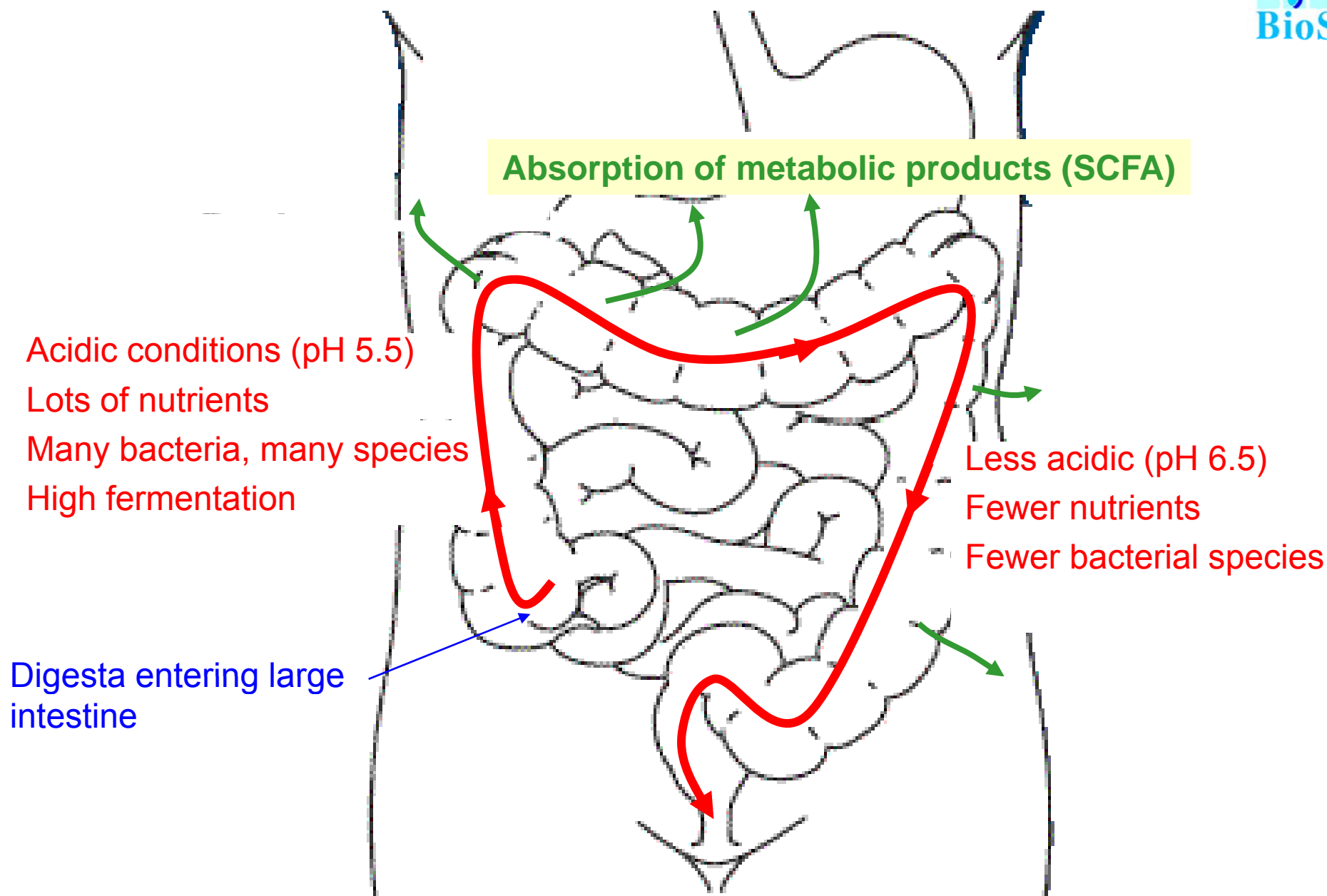
- In vitro – more realistic (continuous fermentor studies)
  - Representation of aspects of colon
  - Continuous supply of substrate



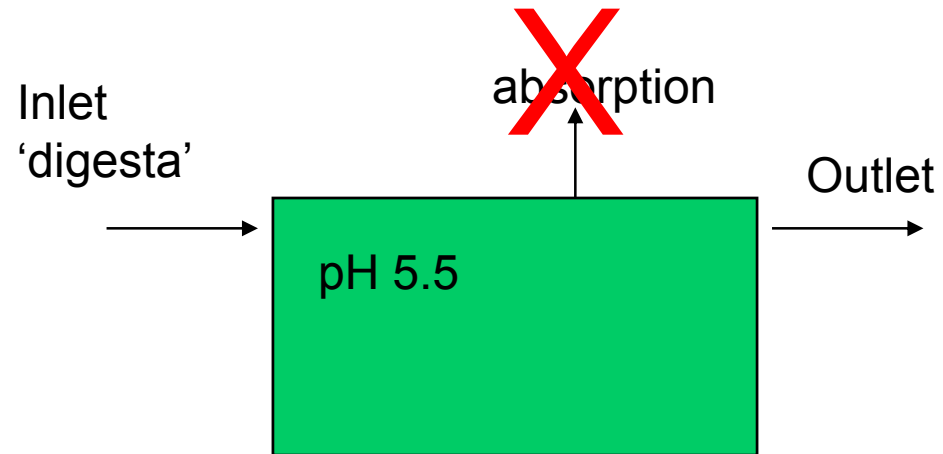
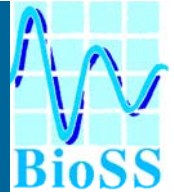
# Fermentor



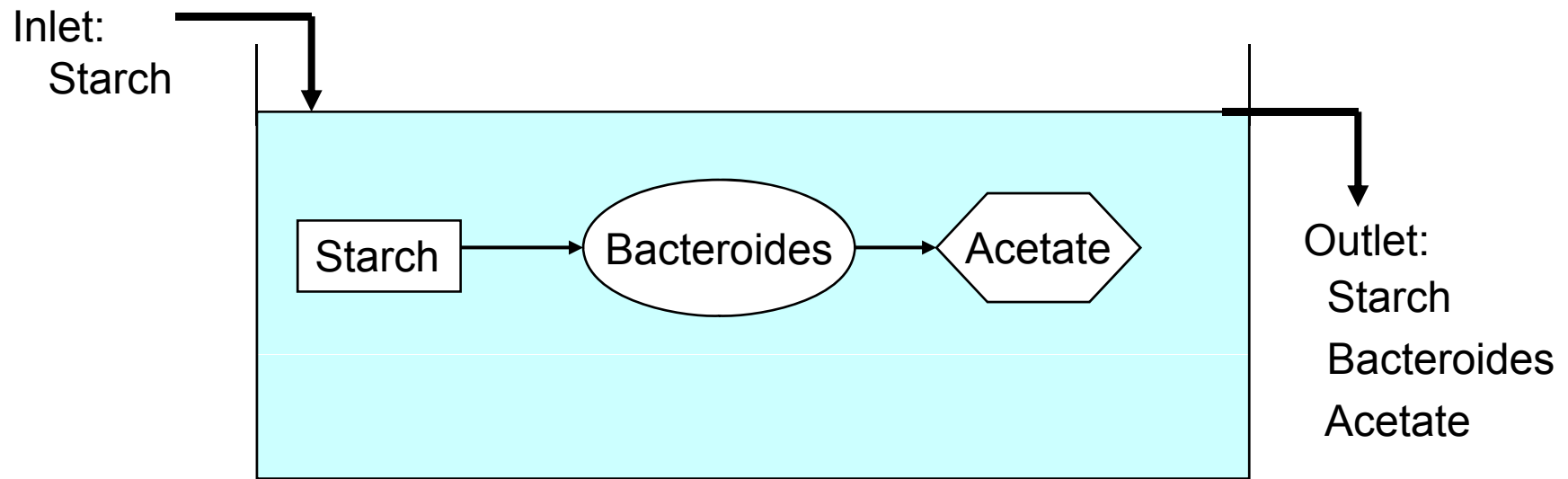
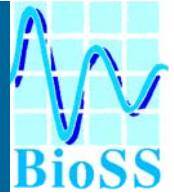
- Vessel with volume of 250 ml
- Faecal sample taken as a 'seed'
- Continuous supply of substrate
- Samples collected from outlet at regular time intervals
- Frequent gentle stirring to ensure mixing
- Controlled conditions (temperature, pH)



# Colon simplified



# Illustration



# Simple model



**Bacteria**

$$\frac{dBac}{dt} = \text{Bacterial growth} - \text{outflow bacteria}$$
$$\frac{dBac}{dt} = \frac{S G}{S + M} Bac - k Bac$$

**Substrate**

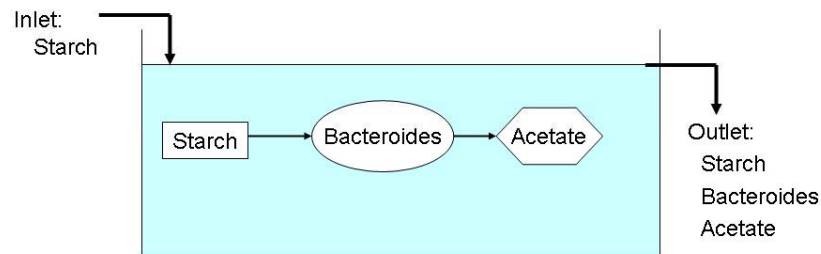
**Products**



# Simulation study

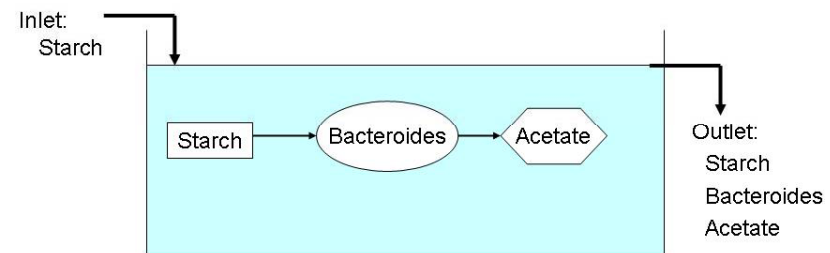


- Two experiments (pH 5.5 and pH 6.5)
- One bacterial strain, one substrate
- Assume maximum growth rate is different for the two pH levels
- All other parameters are assumed the same for both experiments



pH 5.5

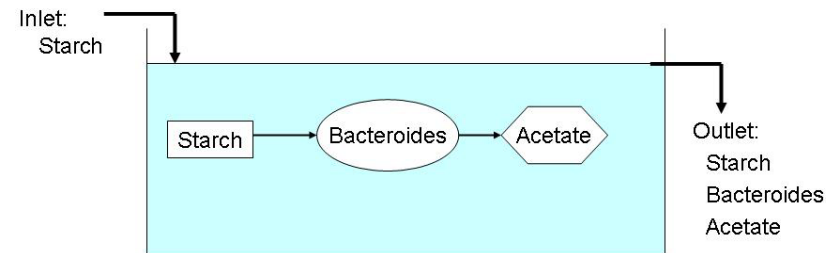
Bacteroides do not grow well



pH 6.5

Bacteroides grow well

# Simulation study

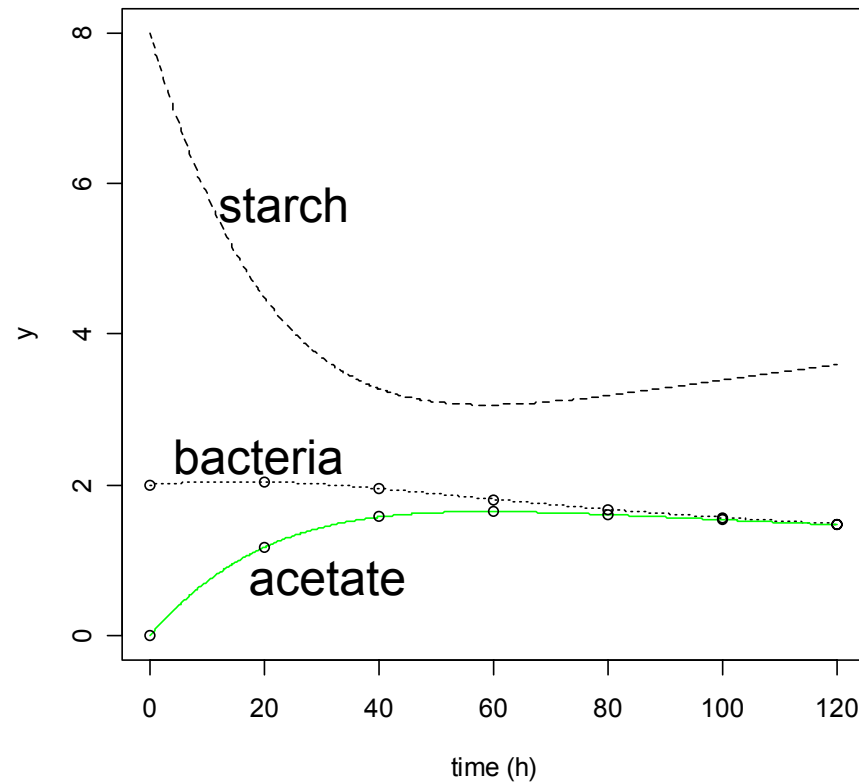


- To be inferred:
    - Max growth rate for pH 5.5
    - Max growth rate for pH 6.5
    - Michaelis-Menten factor
    - Inefficiency (use of substrate/g growth)
    - Acetate production (production/g growth)
- } The same for both experiments
- Use simulated data to compare Bayesian approaches for inference

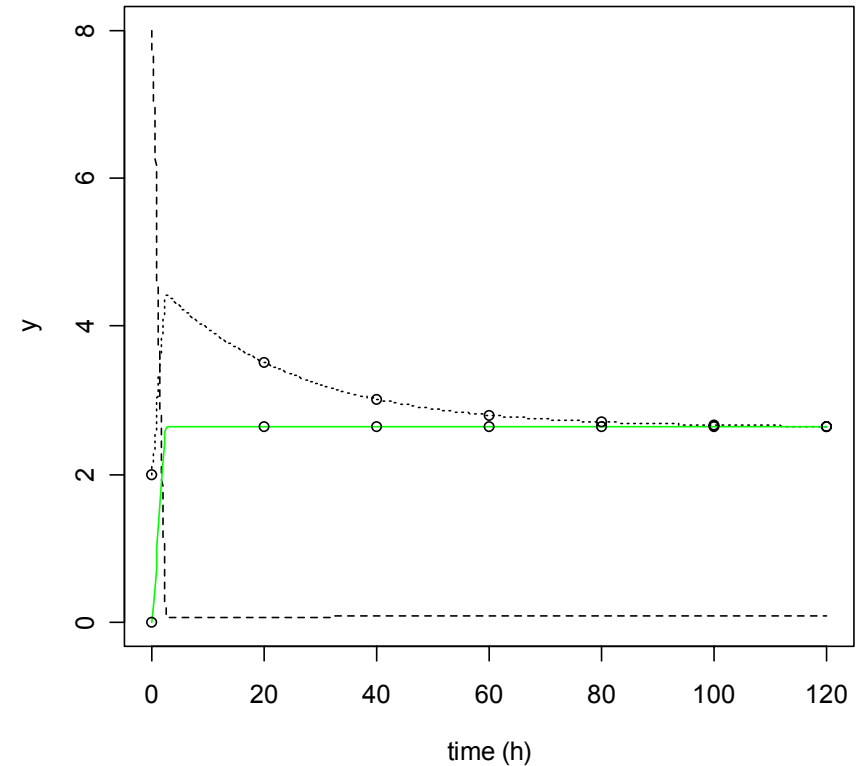
# Simulated time series for fermentor



pH 5.5



pH 6.5



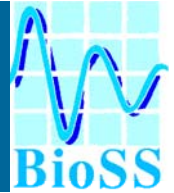
## Parameter settings:

Starch input 5.6 g/d; turnover of 1/d;

$G_{max}=0.05/h$  for pH 5.5;  $G_{max}=0.5/h$  for pH 6.5;

$M=1$ ; Starch use 3g/g bacterial growth; Acetate production 1g/g bacterial growth

# Statistical model



System of differential equations

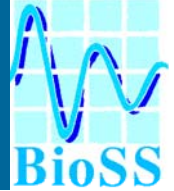
$$y_{abc} | \theta_a, t_{abc}, \sigma_{abc}^2 \sim N(f_a(\theta_a, t_{abc}), \sigma_{abc}^2)$$
$$\theta_a | \mu \sim N(X_a \mu, \Sigma)$$
$$\mu \sim \text{Prior}$$

Indices relate to study  $a$ , measurement  $b$ , time point  $c$ .  $X_a$  is a so-called design matrix for experiment  $a$  to allow for experiment-specific parameters. Variance components  $\sigma_{abc}^2$  and  $\Sigma$  assumed known.

Priors chosen reasonably informative, based on biological knowledge.  
Metropolis-Hastings algorithm used to generate samples from posterior.

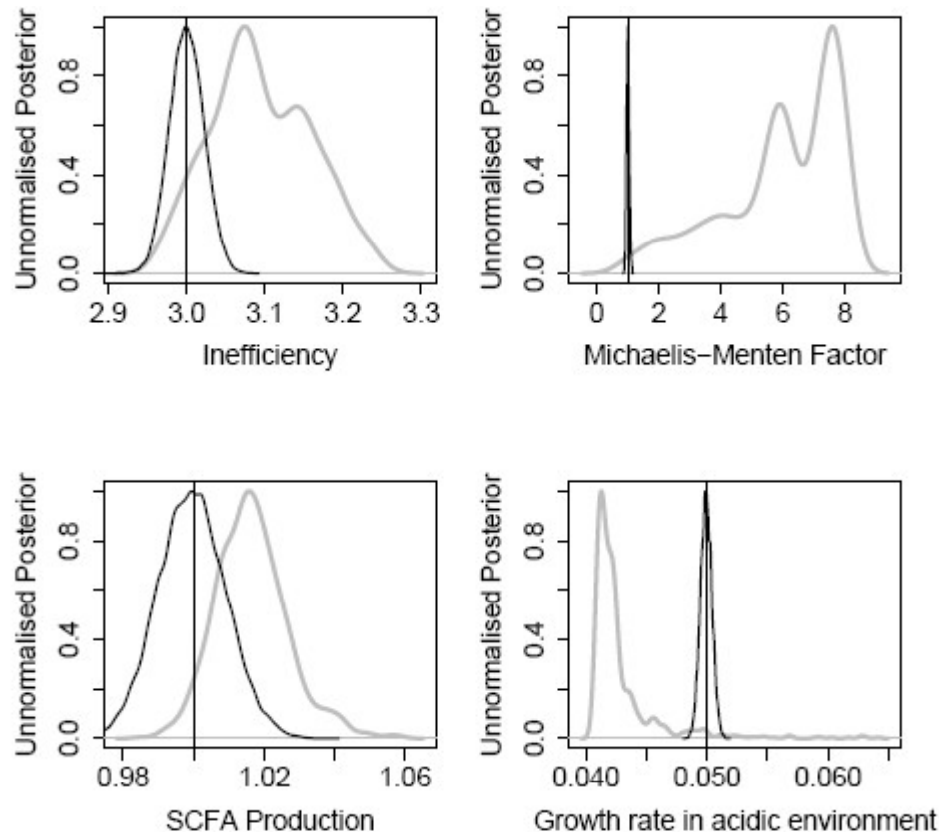
Time-consuming as for each parameter proposal the system of differential equations needs to be evaluated. [R-code on Beowulf cluster](#).

# Two scenarios for estimation



- Scenario A:
  - simultaneous inference using data from both pH 5.5 and pH 6.5 studies
- Scenario B:
  - analyse data from pH 6.5 first and summarise posterior density using mean and variance (ignoring correlations)
  - Use this summary as prior for data analysis at pH 5.5 (for those parameters that are assumed identical across both studies)

# Posterior densities for both scenarios



Scenario A: simultaneous  
Scenario B: sequential

**Fig. 3.** Densities for the posterior probability of parameters using simulated data (Study 1, single bacterial strain) for (black line) Scenario A using simultaneous inference, and (grey line) Scenario B using summary information and separate inference. The true parameter values are indicated as a vertical line. Scenario A has a converged MCMC chain ( $> 200$  effective samples per parameter) whereas no parameters have a large effective sample size in Scenario B.

# Findings



Sequential inference doesn't work well

- Poor mixing

$$\text{Observed growth rate} = G_{\max} \text{ Substrate} / (\text{Substrate} + M)$$

- At pH 6.5 the bacteria grow well so that the substrate concentration is low. The observed growth rate simplifies to  $(G_{\max}/M) \text{ Substrate}$ , and  $G_{\max}$  and  $M$  cannot be separated from this data set alone.
- At pH 5.5 the bacteria don't grow well and substrate concentration is high. Hence the observed growth rate is (approximately) equal to  $G_{\max}$ . When at the second stage the pH 5.5 data are analysed we are working from the 'wrong' prior for  $M$  and for some of the other parameters, and hence poor results.

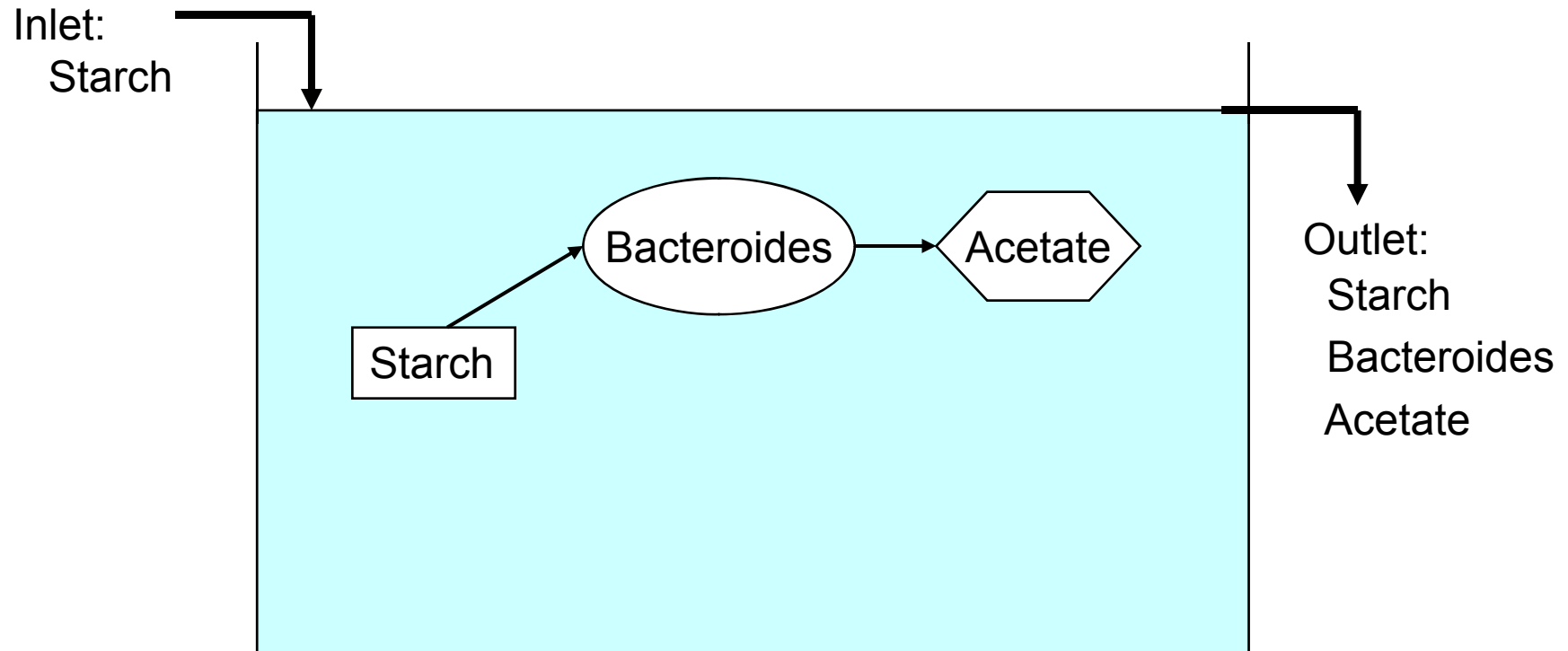
# Findings



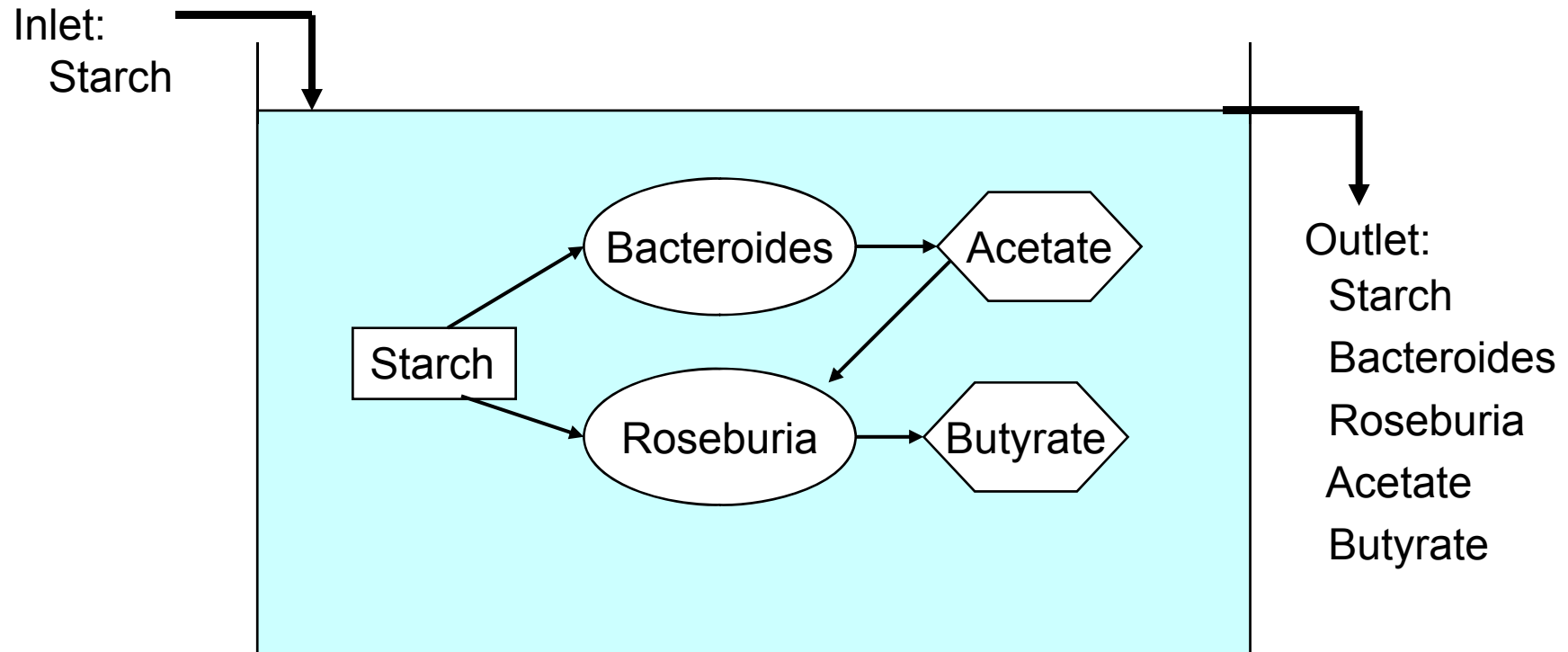
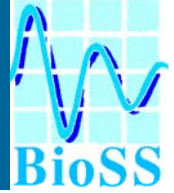
- Simple illustration that Bayesian approach of updating posterior distributions when data become available from new studies [conducted under different experimental conditions](#), does not work here.
- Further investigations with two bacterial species.



# Two bacterial species



# Two bacterial species



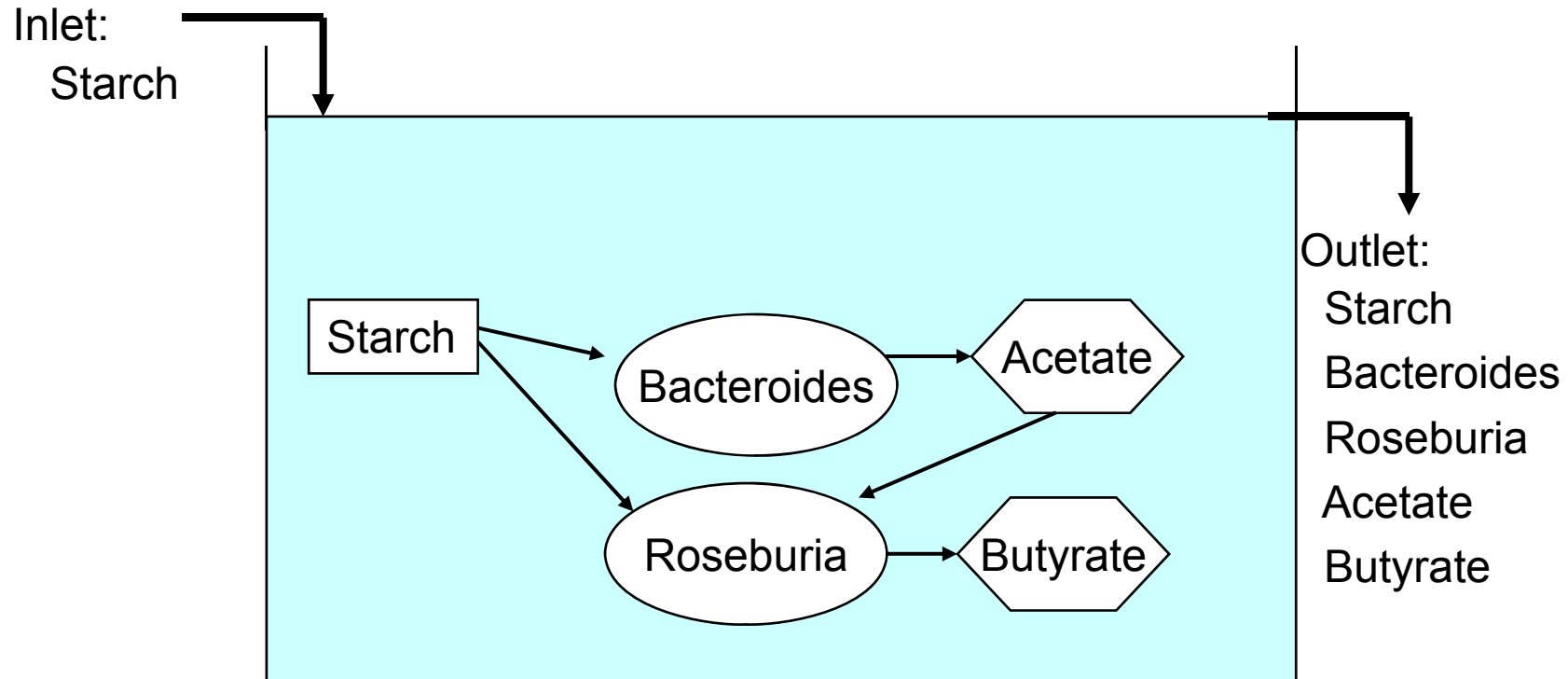
Simulation study to check performance of MCMC algorithm → OK  
(simultaneous inference of pH 5.5 and pH 6.5 'data')

# Data from fermentor study

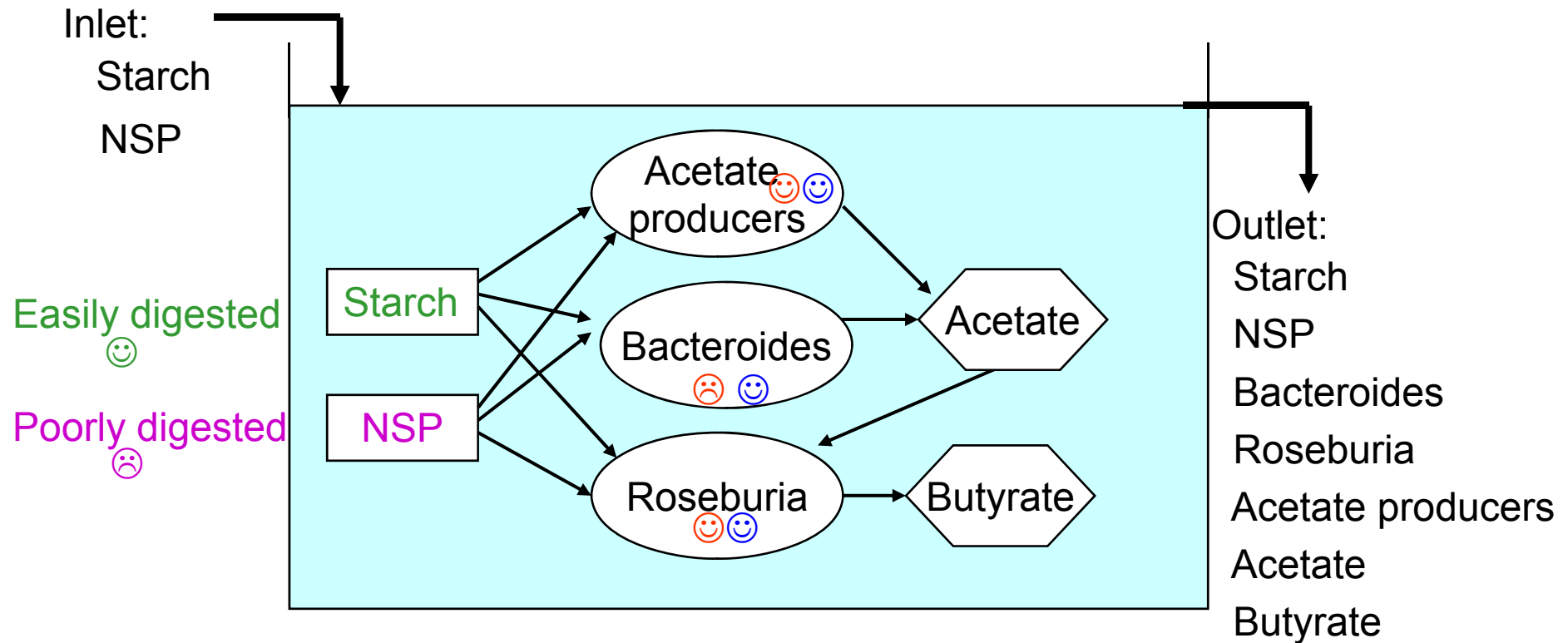
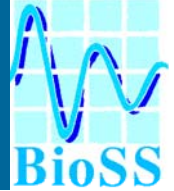


- Mixed bacterial population (from faecal sample)
- Fermentor experiment at pH 5.5 and at pH 6.5
- Two types of substrates were fed to the fermentor
  - Starch: easily digestible, bacteria grow well
  - NSP (fibre): poorly digested, bacteria do not grow well
- Population is assumed to consist of 3 major functional groups

# Data: 3 species-groups, 2 substrates

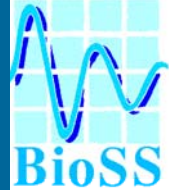


# Data: 3 species-groups, 2 substrates



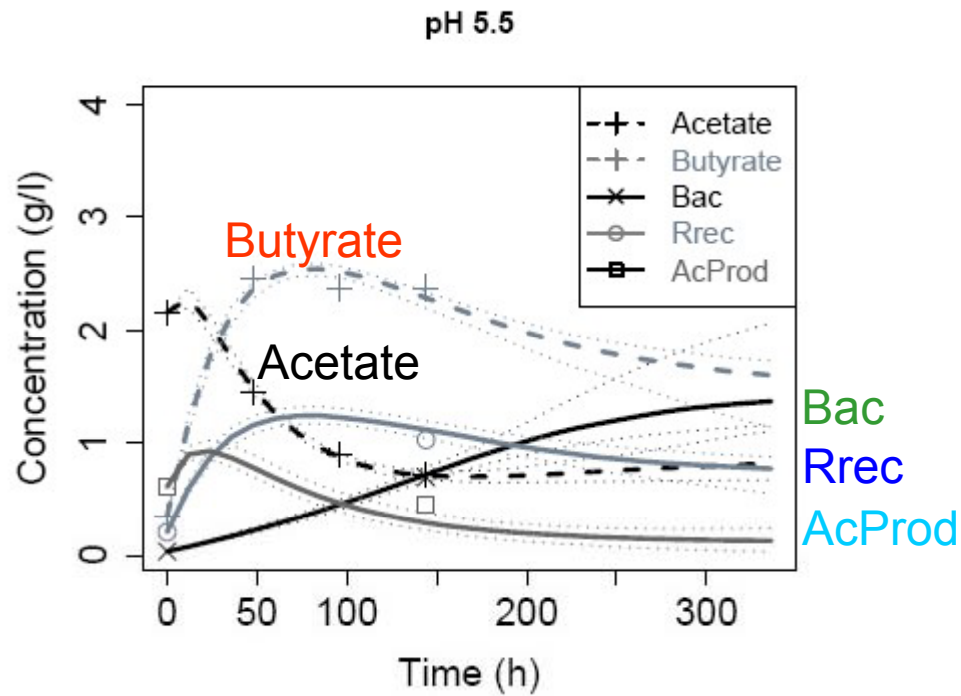
Data from fermentor study at pH 5.5 and from fermentor study at pH 6.5

# Estimation for data from fermentor study



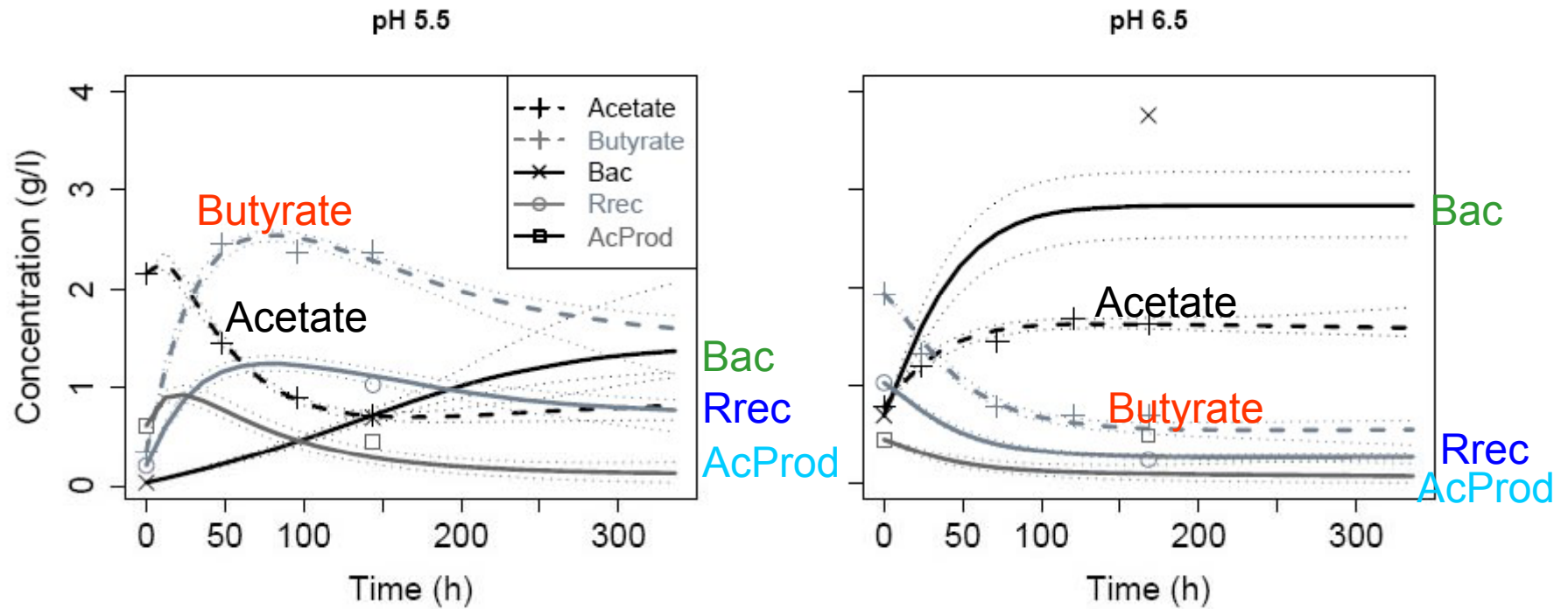
- Data on bacteria (2 time points) and metabolic products (4 time points) for each fermentor
- 29 unknown parameters
- Prior distributions reasonably informative
- Variance components assumed known
- Highly structured model of differential equations, which acts as additional prior information

# Data and posterior predictions



- - - Metabolic products
  - Bacterial groups
- Shown: median and 95% credible intervals

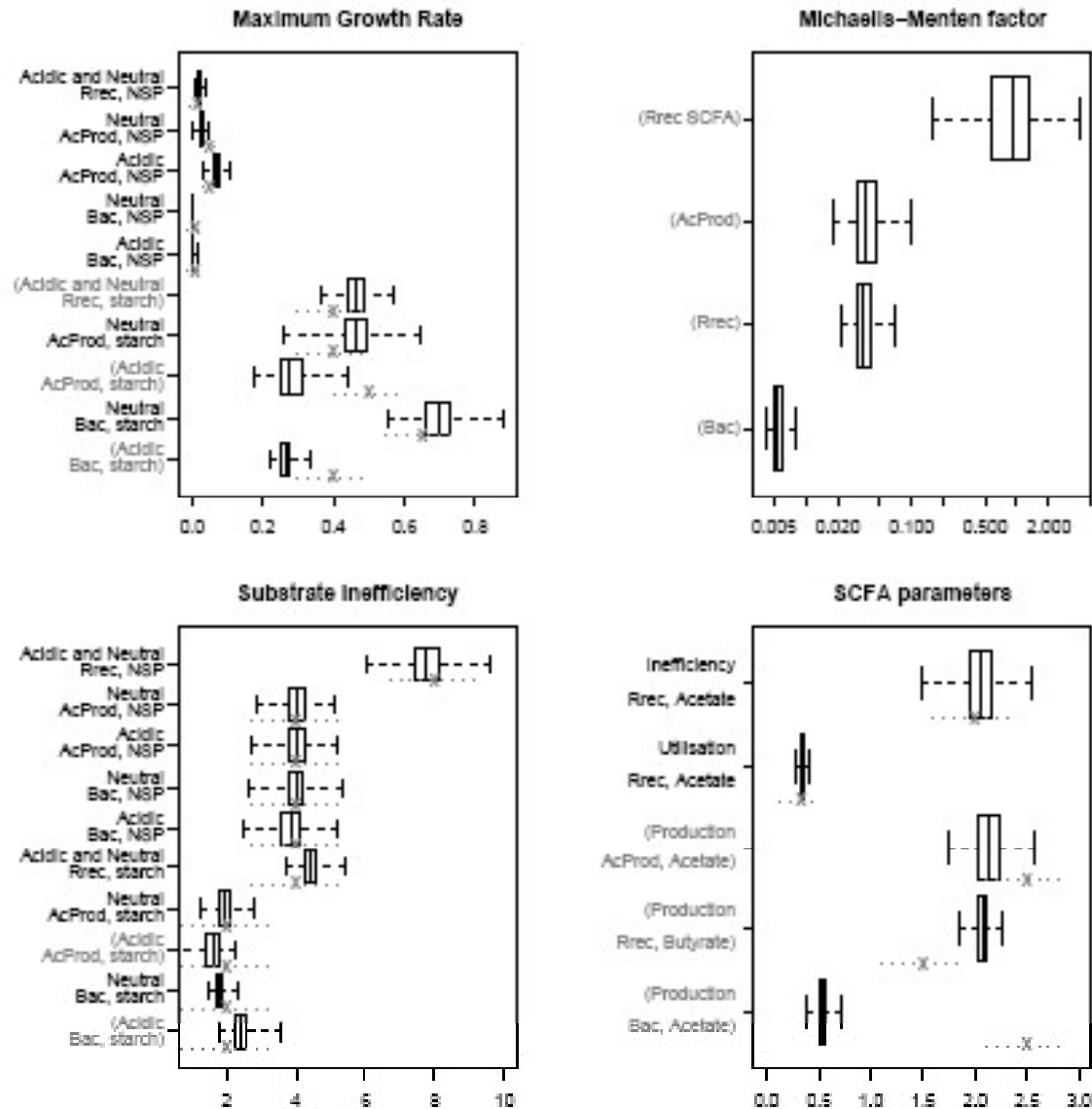
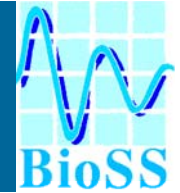
# Data and posterior predictions



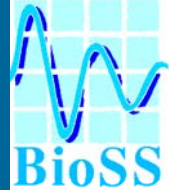
- - - Metabolic products
  - Bacterial groups
- Shown: median and 95% credible intervals



# Posterior distributions

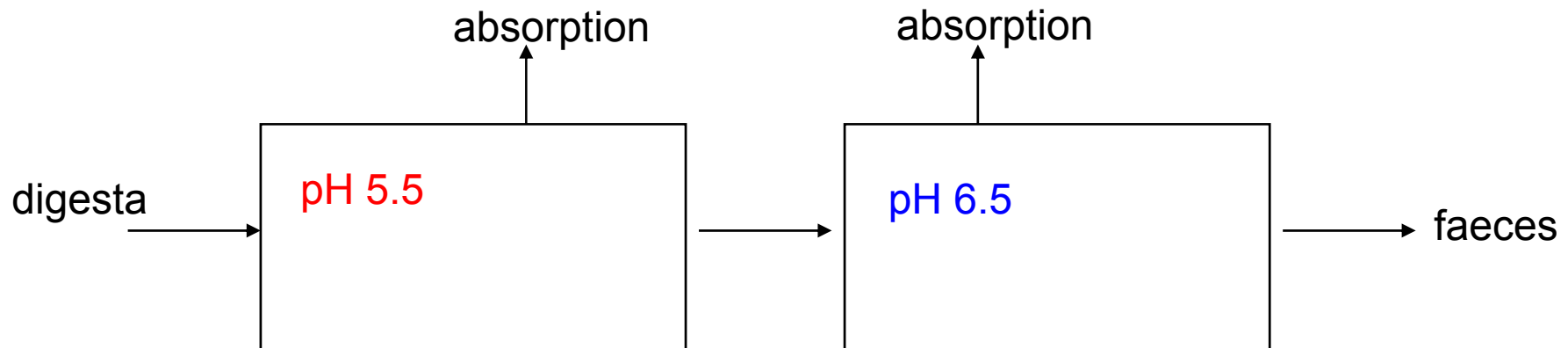


# Predict behaviour in colon



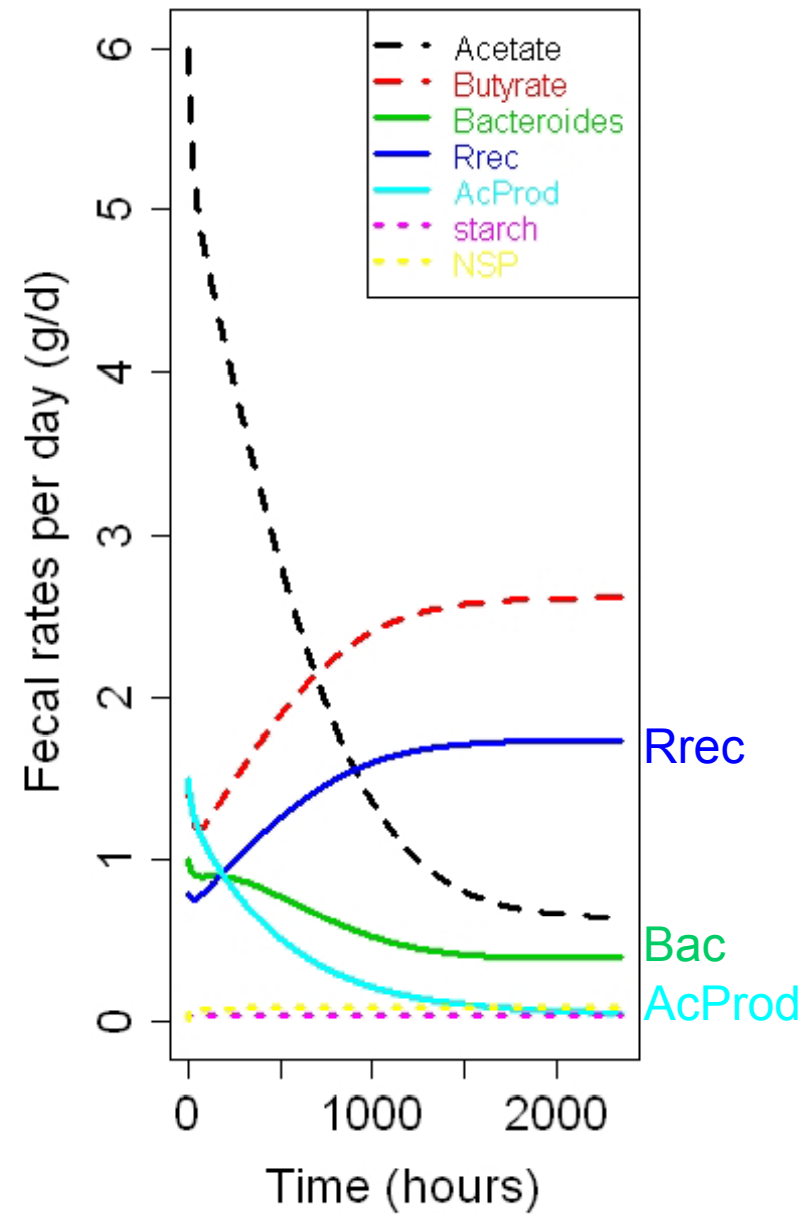
Fermentor studies: no absorption

But this is major aspect of conditions in vivo!

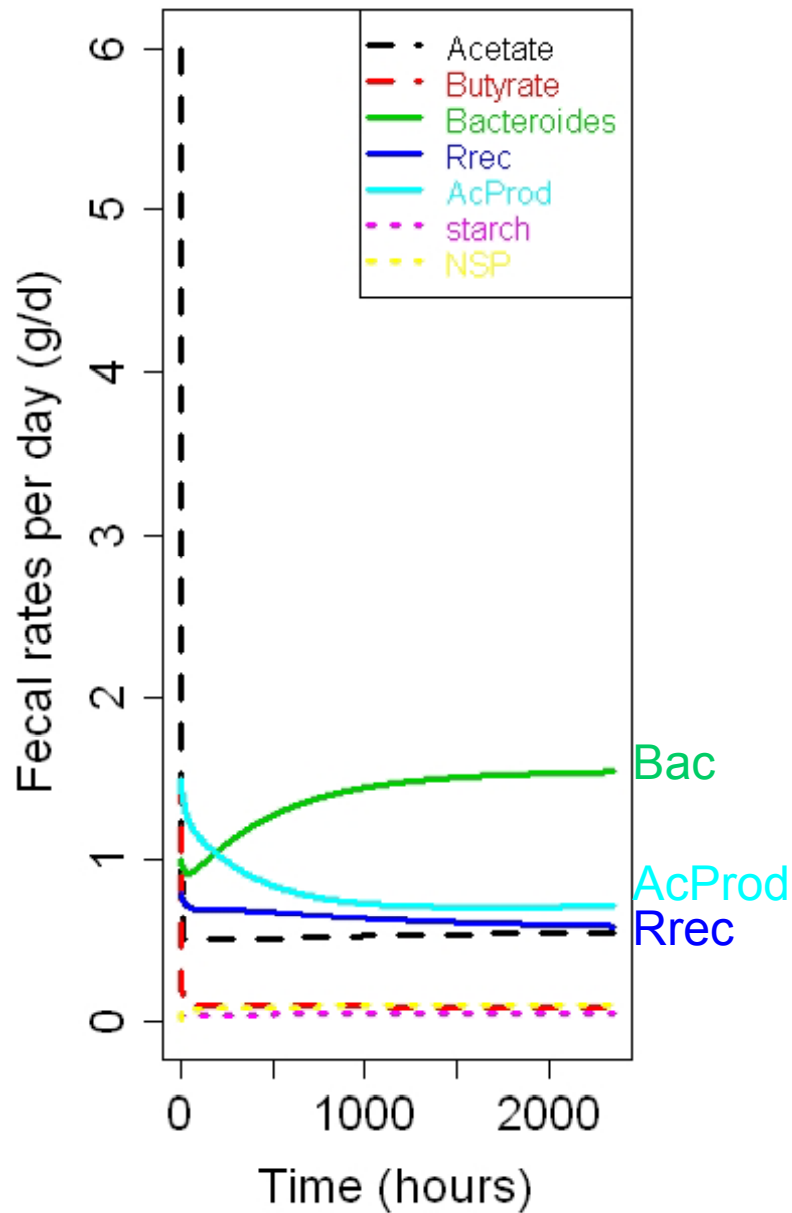


Use model findings to simulate what would happen if absorption of metabolites is allowed for

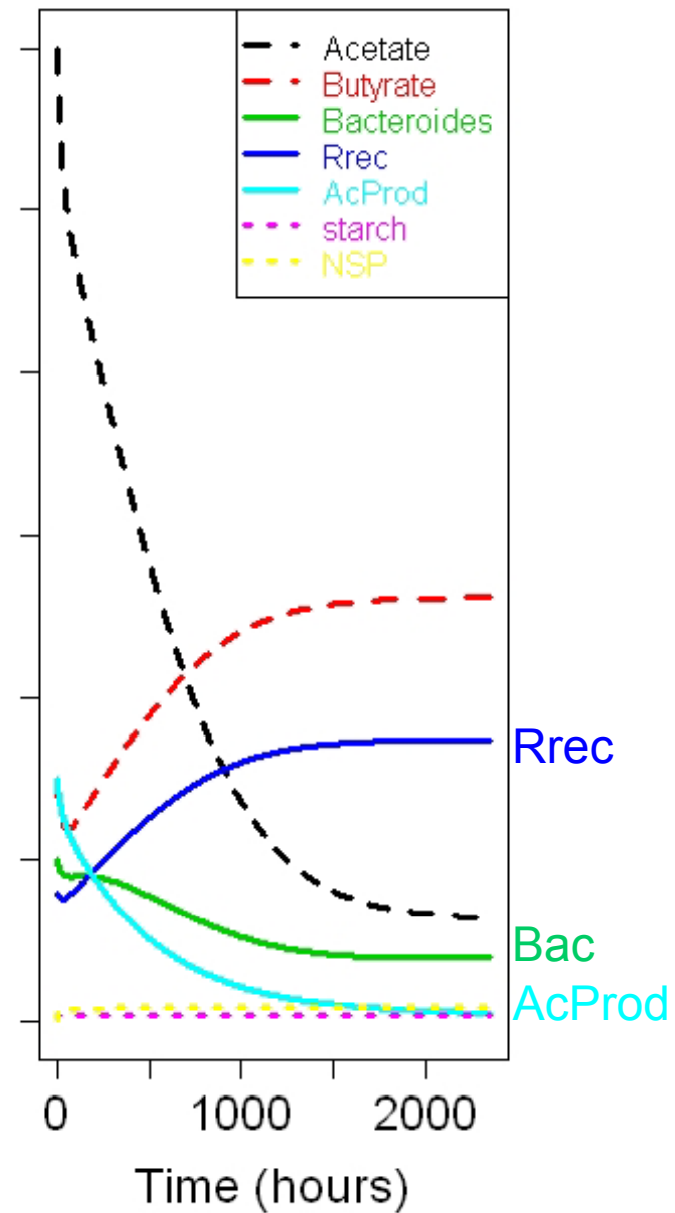
No SCFA absorption: 0g/h



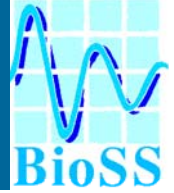
SCFA absorption 0.375g/h



No SCFA absorption: 0g/h



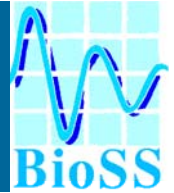
# Predict behaviour in colon



## Results:

- Dominance of bacterial species has changed
- Metabolite concentrations reduced
- This helps to explain why observations from fermentor studies are different from those in vivo (faecal samples)

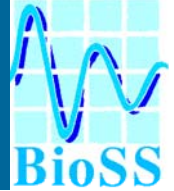
# Conclusions



## Systems with

- limited data
- Well-defined biological model structure (differential equations)
  - Non-linear
  - Non-steady state
  - No analytical solution
- Many unknown model parameters – not measurable
- Often, observations only available from an easily accessible compartment (such as plasma, or faeces), with underlying compartments unobserved
- Usually, estimation of parameters for one data set at a time, which tends to be problematic: (non-)convergence, local? optima, only part of parameter space explored, no or limited information on dependencies between parameters.
- Here we have shown how Bayesian approaches may be useful
- Can take advantage of Bayesian approaches when combining studies (run under different conditions)
- Findings are preliminary!

# Conclusions



## Issues

- Issues with convergence, poor mixing, long chains needed, computational load
- How to select appropriate biological models (ODEs)
- Explore behaviour of ODEs for biological range of parameter values (insights into competition, exclusion, co-existence etc)
- So far only looked at simple bacterial systems

## Future

- More complex bacterial behaviour
- 'Feedback' of products on bacterial growth (SCFA increase acidity)
- Use data from fermentor studies to infer systems parameters, then adapt models to simulate bacterial behaviour in the gut
  - Absorption
  - Movement of digesta and products down the colon
- Faecal samples representative of bacterial processes in colon?

# Acknowledgements



- Daniel Lawson (BioSS, now U.Bristol)
- Harry Flint (Rowett Institute of Nutrition and Health)
- Tony Travis (Rowett Institute of Nutrition and Health)
- Funded by Scottish Government