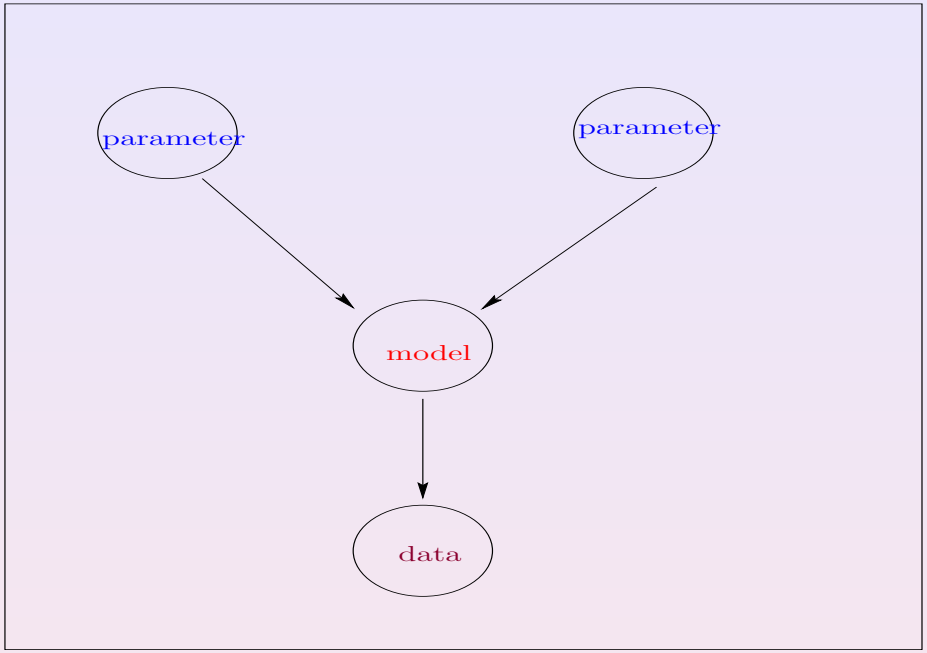
High Resolution Bayesian Space-Time Modelling for Environmental Data

Sujit Sahu
S^{3}RI: Southampton Statistical Sciences Research Institute
http://www.soton.ac.uk/∼sks/

National Oceanography Centre, Southampton
April 2013

A Simple Example of Bayesian Modelling



- For example, data can be independently normally distributed.
- The model can assume that all data points have a common mean and a variance (parameter).
- The specification is completed by assuming prior distributions for unknown parameters.

Bayesian Hierarchical Modeling, BHM

Bayesian models are generally specified (i.e. distributions are written) in hierarchical stages, e.g.

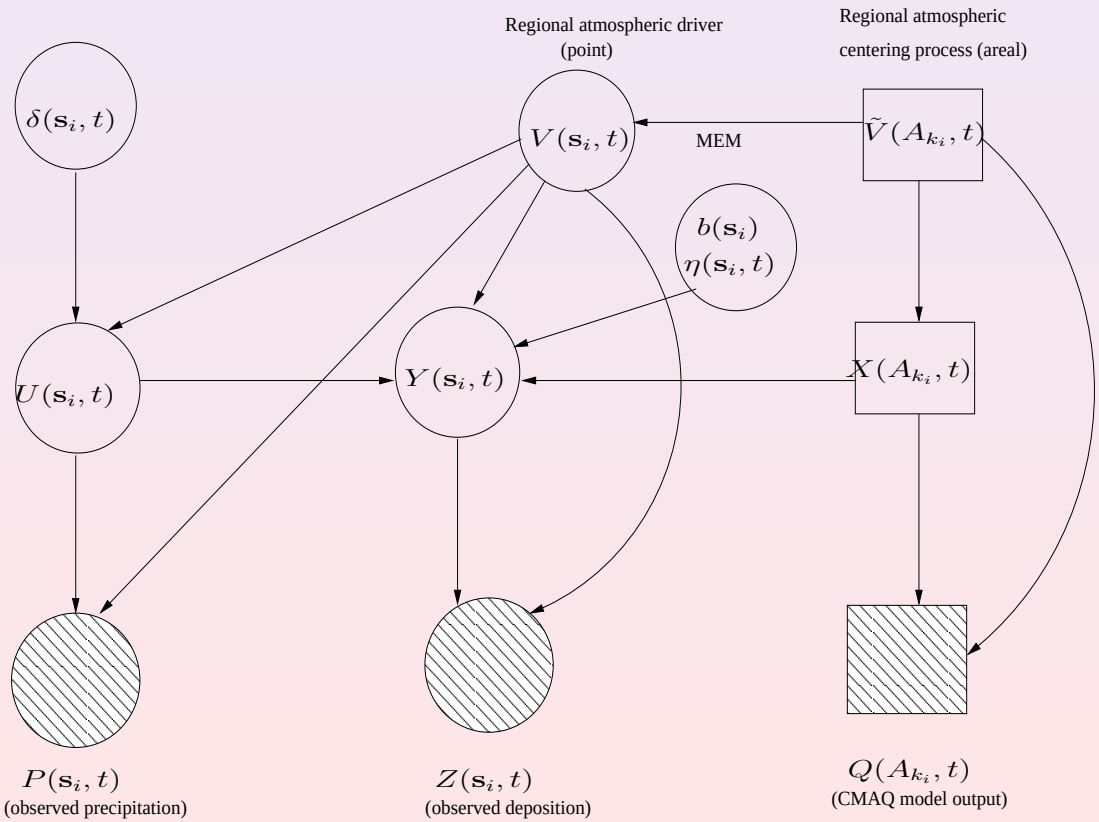
1. first stage: $[\text{Data} \mid \text{model}, \text{parameters}]$
2. second stage: $[\text{model} \mid \text{parameters}]$
3. third stage: $[\text{parameters}]$

The Bayesian model gives us:

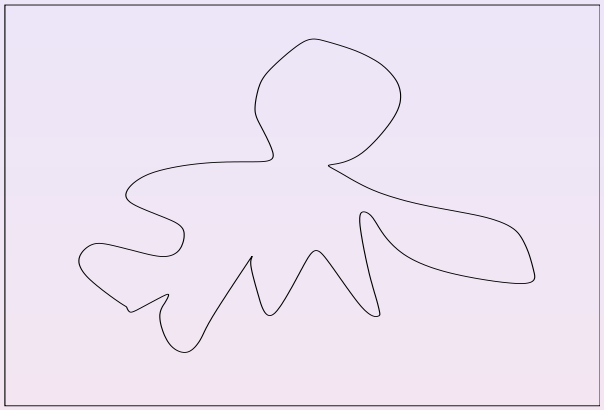
- Posterior distribution, $[\text{parameters} \mid \text{Data}]$ – for quantifying uncertainty.
- Predictive distribution, $[\text{Future Data} \mid \text{Data}]$ – for forecasting and prediction.

My Research

1. Propose new models for problem solving.
2. Develop estimation methods, model based inference and prediction methods.
3. A complex BHM (Sahu et al., 2010). Note the structure, ignore the details.

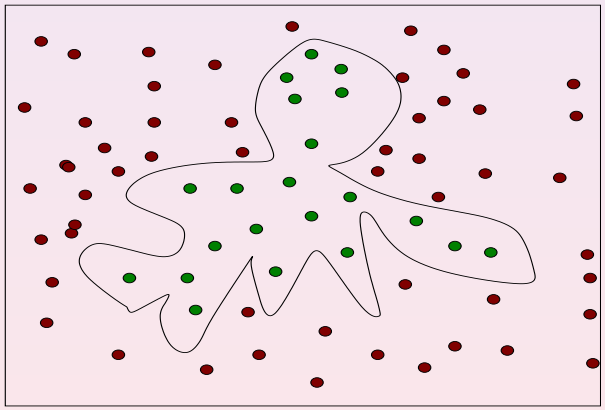


Simulation Methods for Fitting and Prediction



- Problem: Evaluate the area inside the curve.

- Solution: Use simulation methods.
- For example, the Markov chain Monte Carlo (MCMC) techniques.



Advantages of BHM & MCMC

1. They free us from having to use simple, and possibly un-realistic, models.
2. They allow accurate assessment of uncertainty arising from various sources.
3. They provide a unified and coherent approach for solving a wide range of real world problems.

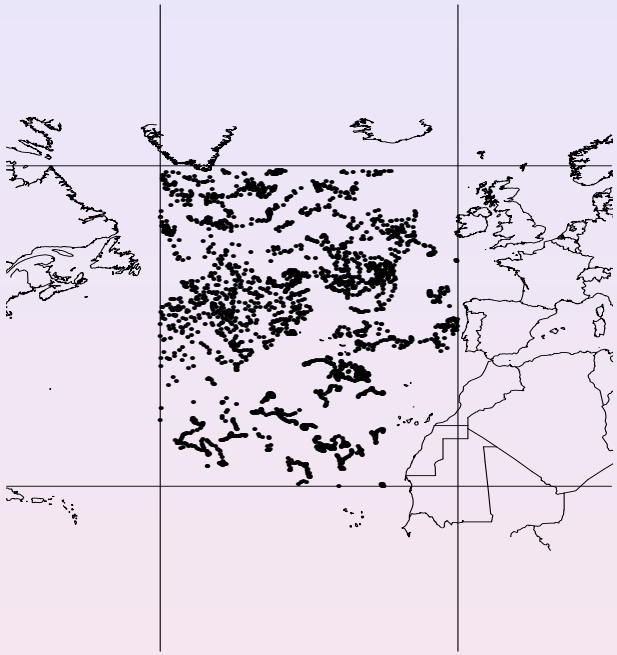
These methods enable one to cross traditional subject boundaries to work in multi-disciplinary research.

Advantages of BHM & MCMC

1. They free us from having to use simple, and possibly un-realistic, models.
2. They allow accurate assessment of uncertainty arising from various sources.
3. They provide a unified and coherent approach for solving a wide range of real world problems.

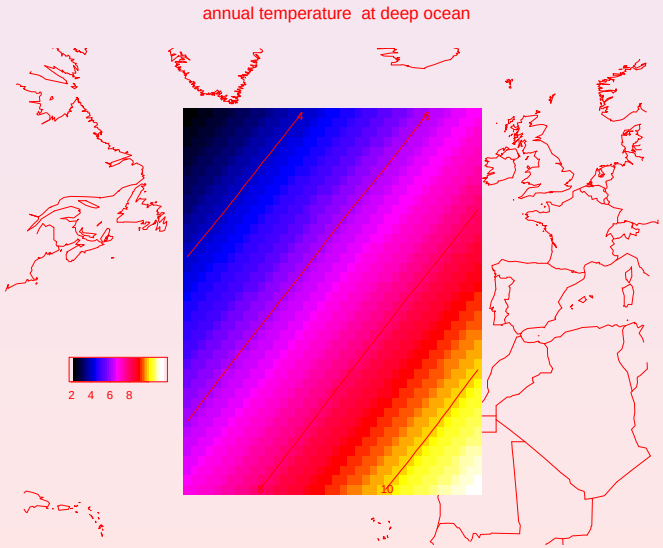
These methods enable one to cross traditional subject boundaries to work in multi-disciplinary research.

Example: Oceanography

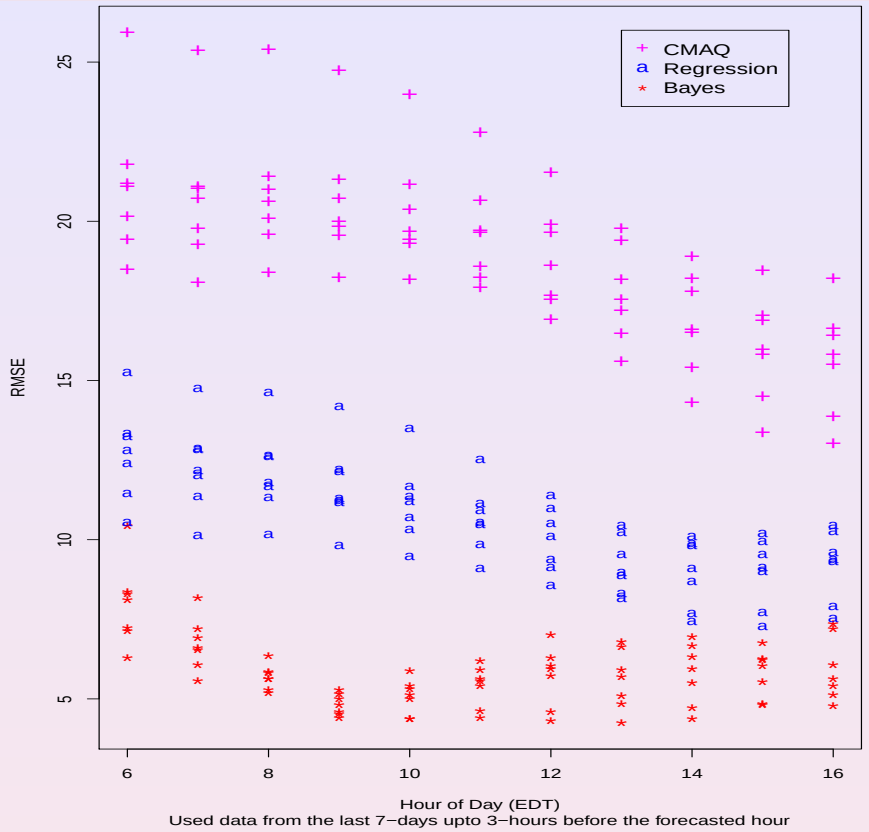


- Need a spatio-temporal model for:
- annual prediction of the temperature and salinity surfaces,
- with their uncertainties.

- Macronutrient cycles: NERC grant: PI Duncan Purdie.
- Marine Geosciences: Russell Wynn.

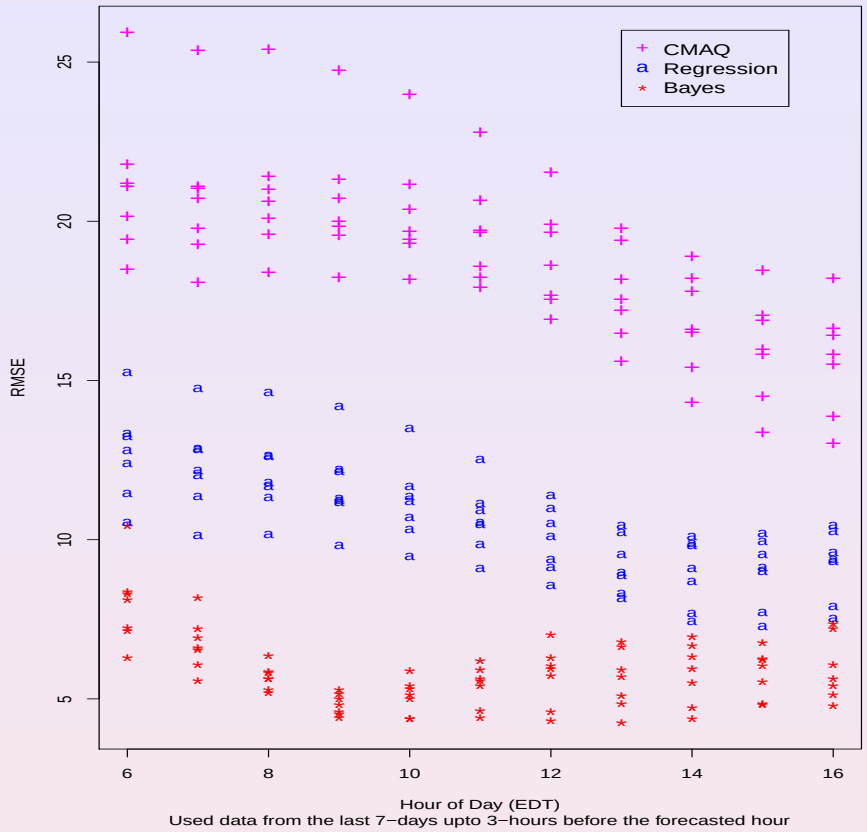


Example: Air Pollution Forecasting



- Root Mean Square Errors (RMSE) of real time forecasts for 8-hr averages.
- The US Environmental Protection Agency has adopted our forecasting method.

Example: Air Pollution Forecasting



- Root Mean Square Errors (RMSE) of real time forecasts for 8-hr averages.
- The US Environmental Protection Agency has adopted our forecasting method.

Why Statistics?

- Often there is a ton of uncertainty in the conclusions and decisions. Stochastic modelling allows us to reduce these.
- Such model based methods are essential in deciding what is signal, which is often weak, and what can be attributed to noise.
- Statistical methods are often the results of a series of dialogues between the applied scientists, like you, and statisticians.

Uncertain knowledge	+	Knowledge of the extent of the uncertainty in it	=	Useable knowledge

How would I contribute to the DTP?

- Research Excellence:
  - Innovative & transformative, e.g. adoption by USEPA.
- Training Excellence:
  - Research led, but accessible, statistical modelling short-courses run by myself.
  - These underpin the analysis methods based on sound statistical theory and practice.
  - A number of NOCS and SOES students/postdocs benefitted from these courses.
- Multidisciplinary.
  - Mathematics and Statistics PhD students will solve the practical modelling problems.
- Partnership and Operational Management.
  - Excellent research practice, environment and management of S3RI.
  - Help in recruiting Mathematics students.