A robust Bayesian approach to land use modelling

Matthias C. M. Troffaes

in collaboration with Lewis Paton (Durham),

Nigel Boatman (FERA), Mohamud Hussein (FERA), and Andy Hart (FERA)

Department of Mathematical Sciences, Durham University, UK

25 September 2015

work funded by FERA and EPSRC

Outline

Outline

Lakes face 'complex' challenges

Urgent measures are needed to protect lakes in England and Wales from pollution and climate change, according to the Environment Agency.

Fears for England's famous Lakes

The call for action comes as experts gather at Windermere in Cumbria to discuss ways to safeguard England's largest lake.

Windermere faces threats including invasive species and farming pollution.

Durham University

Legal bid to protect England's 'precious rivers' from pollution

27 August 2015 | England

Campaigners say poor land management is washing soils carrying nutrients and pesticides into water habitats

Conservation and angling groups have launched a legal bid after claiming the government is not sufficiently protecting England's "precious rivers".

The joint case by WWF-UK, the Angling Trust and Fish Legal will focus on habitats including Poole Harbour and the Avon, Wye and Eden rivers.

They said the government has failed to stop pollution washing into waterways.

Durham University

Problem Description: Aim

- aim? model and predict agricultural land use
- why? food security, landscape, environmental impact
- this project: focus on crop choice

Main Issue

an abundance of uncertain factors influencing crop choices

- soil type
- previous years' crops
- intensity & time of rainfall
- temperature
- crop price
- fertilizer price
- farmer's attitude towards risk
- farmer's assets
- farm size
- ...

Problem Description: Data

- historical crop data (IACS)
- historical rainfall data (MET office)
- historical fertiliser price data (Dairyco)
- soil type map (LandIS)
- expert information on historic crop profit predicitions (John Nix book)
- expert information on yield level per crop & soil type (John Nix book)
- predictions of future price and climate scenarios (decision maker)

Problem Description: Model

- crop sequences typically follow set patterns

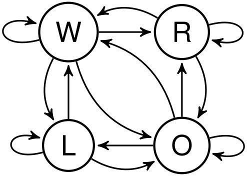
	year 1	year 2	year 3	year 4
field 1	**wheat**	**fallow**	**wheat**	**beans**
field 2	**barley**	**barley**	**sugar beet**	**wheat**
field 3	**grass**	**grass**	**wheat**	**grass**

- patterns not entirey deterministic: we use a Markov chain

Durham University

Problem Description: Model (Simplified)

W = wheat, R = rapeseed, L = legumes, O = other



- key idea: probabilities are a multicategorical logistic function of a linear combination of the continuous factors (climate and price) influencing crop choices [1]

$$p(W \mid L) = \text{function of } \beta_0 + \beta_1 \times \text{price} + \beta_2 \times \text{climate}$$

- aim of statistical inference: identify β_0, β_1, and β_2 from data

Durham University

Problem Description: Likelihood

- J = number of crops
- k = time
- Y_k = current crop $\in \{1, \dots, J\}$
- Y_{k+1} = next crop $\in \{1, \dots, J\}$
- X_k = current economy, climate, … (regressors) $\in \mathbb{R}^M$
- β = regression coefficients $\in \mathbb{R}^{J \times (J-1) \times M}$

$$p_{kj} = \exp(\beta_{Y_k,j} \cdot X_k) \qquad \text{for } j = 1 \dots J-1 \tag{1}$$

$$p_{kJ} = 1 \tag{2}$$

$$Y_{k+1} \sim \mathrm{dcat}(p_{k1}, \dots, p_{kJ}) \tag{3}$$

- soil type not included as regressor:
  separate model for each soil type

Durham University

Problem Description: Prior

Issues

- little prior information about β's
- flat prior $\neq$ prior ignorance
- too flat prior $\rightarrow$ numerical issues (JAGS barks very loudly)
- particular worry about rare crop types & extreme events
  inferences must not purely reproduce arbitrary prior assumptions
- likelihood is from curved exponential family

Problem Description: Prior

A Possible Solution (Part 1): Conjugate Prior

- a conjugate prior for our likelihood is of the form [2]:

$$f_0(\beta|s_0, t_0) \propto \exp\left(\sum_{i=1}^{J} \sum_{x \in \mathcal{X}} s_{0i}(x) \left[\sum_{j=1}^{J} t_{0ij}(x) \beta_{ij} x - \log \sum_{j=1}^{J} \exp(\beta_{ij} x) \right]\right) \quad (4)$$

- $\mathcal{X}$ = finite set of regressor values x where we specify prior knowledge
- $t_{0ij}(x)$ = 'matches' prior transition probability
  for going from crop i to j for regressor value x
- $s_{0i}(x)$ = controls variance (large values = peaked prior)

benefit of conjugacy = interpretability of parameters!
updating trivial, but no closed form for posterior predictive

Durham University

Problem Description: Prior

A Possible Solution (Part 2): Near-Vacuous Set of Conjugate Priors

- choose for $\mathcal{X}$ a reasonable prior range for regressors
- fix $s_{0i}(x)$ to a constant value—determines learning speed
- choose set of $t_{0ij}(x)$ covering every possible distribution

restrict to extreme points for computations

	number of extreme points			
# crops	2	3	4	5
# regressors 2	16	256	**65536**	4294967296
# regressors 3	81	6561	43046721	1853020188851841
# regressors 4	256	65536	4294967296	18446744073709551616
# regressors 5	625	390625	152587890625	23283064365386962890

situation similar to Walley's imprecise Dirichlet model [6]
but a lot more extreme points to explore!!

Durham University

Outline

Problem Description

Approach 1: Markov Chain Monte Carlo

Approach 2: MAP Estimation

Decision Analysis

Conclusion

Durham University

Approach 1: Markov Chain Monte Carlo

= build a Markov chain which samples from the posterior

Upsides

- full posterior inference

Downsides

- too slow: single run in JAGS = 2 hours
  - 10000 samples
  - 4 crops
  - 2 regressors
  - 30000 data points

  full analysis = 15 years
- sampling error across runs   can be controlled

Durham University

Outline

Problem Description

Approach 1: Markov Chain Monte Carlo

Approach 2: MAP Estimation

Decision Analysis

Conclusion

Durham University

Approach 2: MAP Estimation

= use posterior mode obtained via numerical optimisation

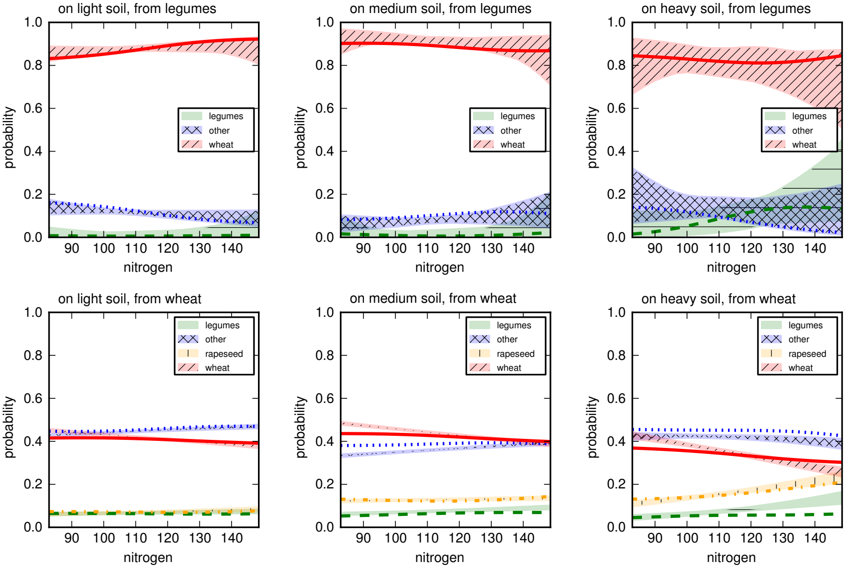
Upsides

- very fast
  single run = fraction of a second
  full analysis typically within 30 minutes
- no sampling error across runs

Downsides

- no full uncertainty quantification
- only reasonable as rough approximation to expectation
  (i.e. linear utility in a decision context)

Approach 2: Validation Against Non-Parametric Estimate



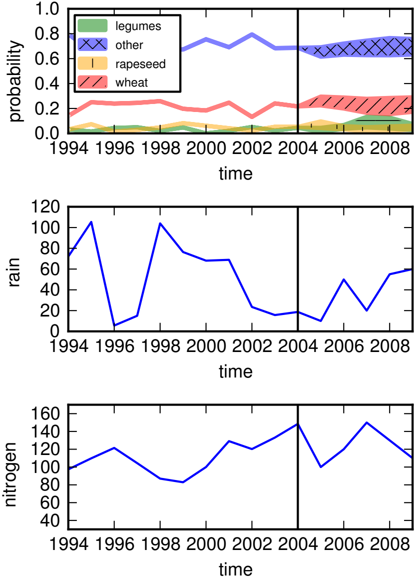
Durham University

Approach 2: Cross-Validation

How good is the model at predicting which crop is grown?

region	deter-minacy	single accuracy	indeterminate output size	set accuracy
Anglia	0.968	0.722	2.008	0.855
Mease	0.988	0.758	2.140	0.929

Approach 2: Prediction



Durham University

Outline

Decision Analysis

- interested in stimulating increase in legumes
- utility function:

$$U(a, b) = a - \kappa b$$

a = fraction of legumes across all farms;
function of b and model parameters β
b = subsidy level
κ = weight constant

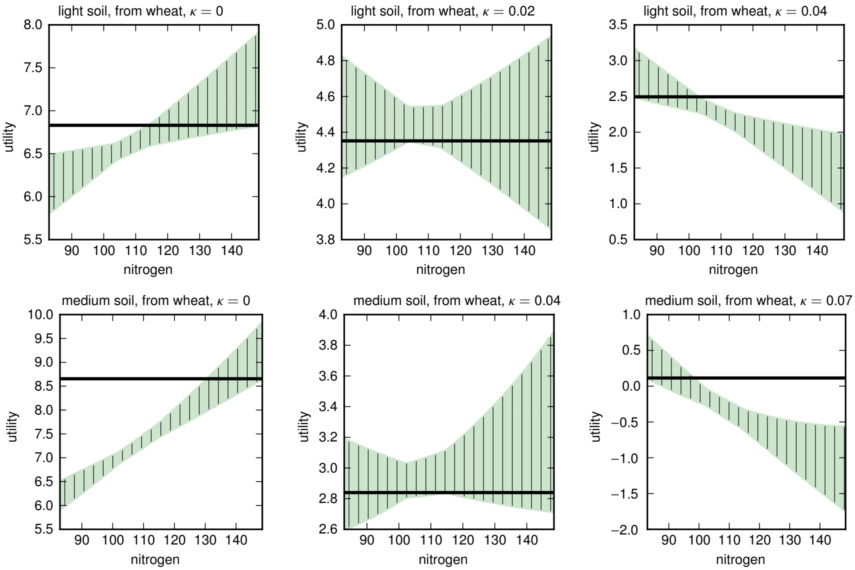
- maximize expected utility by considering all β^* MAP estimates:

$$\left\{ \arg\max_b U(a(b, \beta^*), b) : \beta^* \in B^* \right\}$$

range of optimal policy recommendations
in most cases, actually a unique policy identified

Durham University

Decision Analysis: Results



Outline

Durham University

Conclusion

- conjugacy useful for curved families: interpretation
- robust Bayesian analysis on high-dimensional models is hard: tough trade-offs (forgo MCMC, use MAP)
- you may not know whether there is a prior sensitivity problem until you have done the robust analysis!

Open questions & future work

- reuse a single MCMC run to explore multiple priors at once?
- cleverly explore large quantities of extreme priors (ideas from linear programming?)
- other approximate methods (e.g. ABC)?

Thank you!

References I

[1] Alan Agresti.
Categorical Data Analysis.
John Wiley and Sons, third edition, 2013.

[2] Ming-Hui Chen and Joseph G. Ibrahim.
Conjugate priors for generalized linear models.
Statistica Sinica, 13:461–476, 2003.

[3] Lewis Paton, Matthias C. M. Troffaes, Nigel Boatman, Mohamud Hussein, and Andy Hart.
Multinomial logistic regression on Markov chains for crop rotation modelling.
In Anne Laurent, Oliver Strauss, Bernadette Bouchon-Meunier, and Ronald R. Yager, editors, *Proceedings of the 15th International Conference IPMU 2014 (Information Processing and Management of Uncertainty in Knowledge-Based Systems, 15–19 July 2014, Montpellier, France)*, volume 444 of *Communications in Computer and Information Science*, pages 476–485. Springer, 2014.
doi:10.1007/978-3-319-08852-5_49.

[4] Lewis Paton, Matthias C. M. Troffaes, Nigel Boatman, Mohamud Hussein, and Andy Hart.
A robust Bayesian analysis of the impact of policy decisions on crop rotations.
In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA'15: Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications*, pages 217–226, Pescara, Italy, jul 2015. ARACNE.
URL: http://www.sipta.org/isipta15/data/paper/17.pdf.

[5] Matthias C. M. Troffaes and Lewis Paton.
Logistic regression on Markov chains for crop rotation modelling.
In F. Cozman, T. Denœux, S. Destercke, and T. Seidenfeld, editors, *ISIPTA'13: Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications*, pages 329–336, Compiègne, France, July 2013. SIPTA.
URL: http://www.sipta.org/isipta13/index.php?id=paper&paper=033.html.

[6] Peter Walley.
Inferences from multinomial data: Learning about a bag of marbles.
Journal of the Royal Statistical Society, Series B, 58(1):3–34, 1996.
URL: http://www.jstor.org/stable/2346164.

Durham University