

Liebig's Law of the Minima: Interpreting N Response in 2 Dimensions

Introduction

Linear-plateau models (graph) for crop yield are simple to interpret and biologically defensible. However, extending these models to two sources of a nutrient across a response surface is challenging both to fit computationally and to visualize conceptually.

Model-fitting with existing software is either sensitive to starting conditions, slow to converge, difficult to implement, or all of the above.

Figures plotted in 3D have perceptual challenges due to the constraints of axonometric projections and non-interactivity on the printed page. Flattening figures to present marginal effects of one variable at levels of another creates an implicit hierarchy. Choices must be made to balance loss-of-information with interpretability.



Data

Corn yield was measured following a factorial of 6 seeding rates of two cover crop species in replacement series (cereal rye 0...168 kg ha⁻¹; hairy vetch 34...168 kg ha⁻¹) across 5 rates of subsurface-banded poultry litter (SSB PL 9...269 kg PAN ha⁻¹). The strip-plot experiment was conducted on two sites (Organic, Conventional; 3 replicate blocks each) and replicated in three years (2012, 2013, 2014). Cover crop mixtures were separated and all residues combusted to compute a net C:N ratio.

Code repository

To see annotated code for this analysis, citations, and an interactive 3D version of these figures, visit:

github.com/brianwdavis/liebig



Software packages

nlme :: nlme()

Advantages

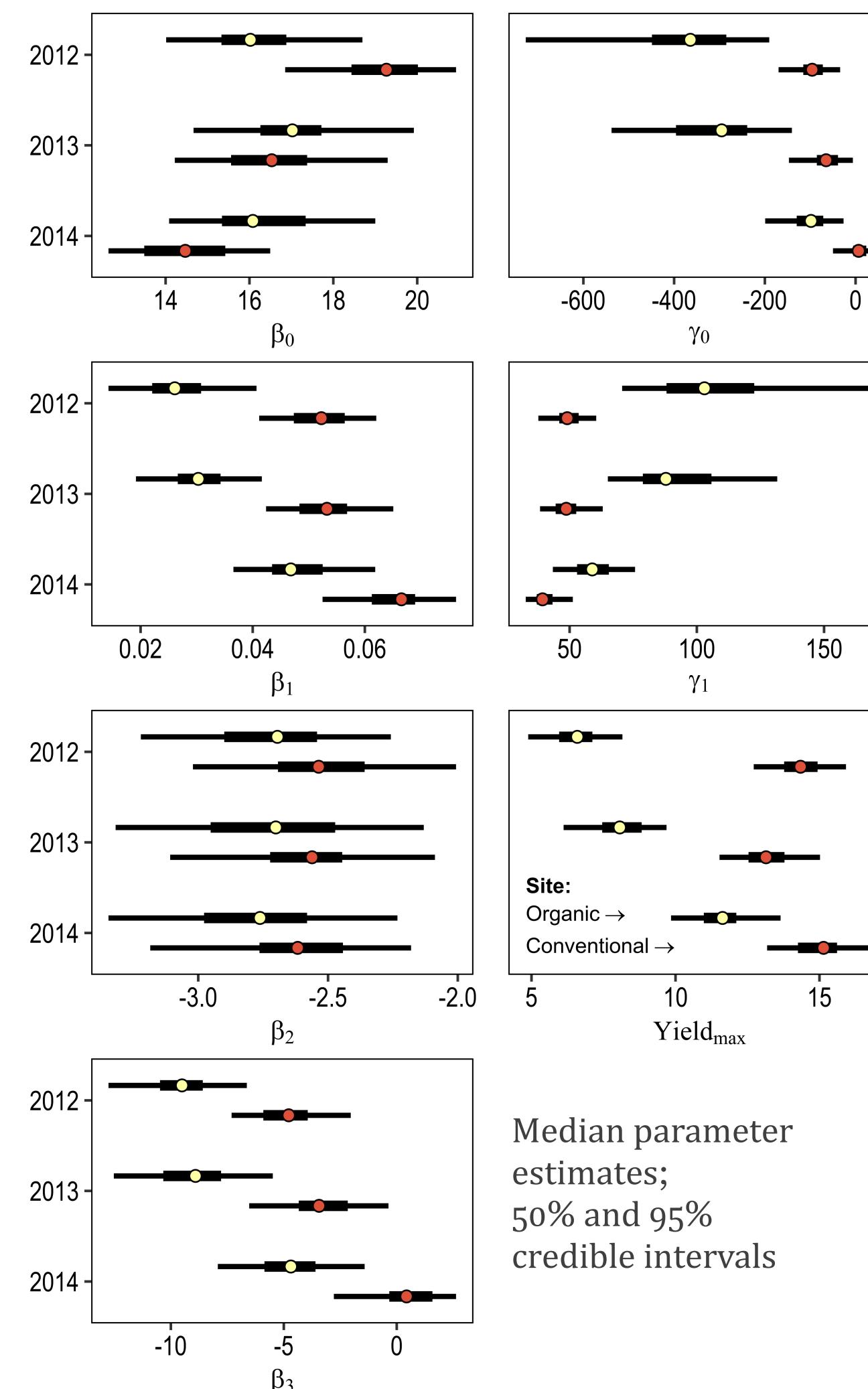
- Familiar syntax to linear model fitting in R
- Allows non-arithmetic functions in formula; e.g. pmin(), SSasymp()
- No additional software needed
- Once starting values are optimized and correct scaling applied, fitting is relatively fast (seconds to minutes)

Challenges

- Very sensitive to scaling of Y, X1, and X2; guess-and-check iteration
- Predictions, coefficients must be back-scaled
- Sensitive to manually specified starting values to achieve convergence
- Must be parameterized such that $\beta_i \sim N(\mu, \sigma)$

Best use: Exploratory analysis

Model fit and interpretation



Monte Carlo simulation

Regardless of the package used, methods exist for both to extract model components, namely the means and variance-covariance matrices. Coefficient estimates can be then generated from a joint-normal distribution. Alternately, brm fit objects store a population of posterior samples that can be used.

These can be used to derive the Gaussian-ratio-distributed frontier parameters γ_0 & γ_1 and account for covariance in the normally-distributed plateau parameter, Yield_{max} = $\beta_0 + \beta_3$.

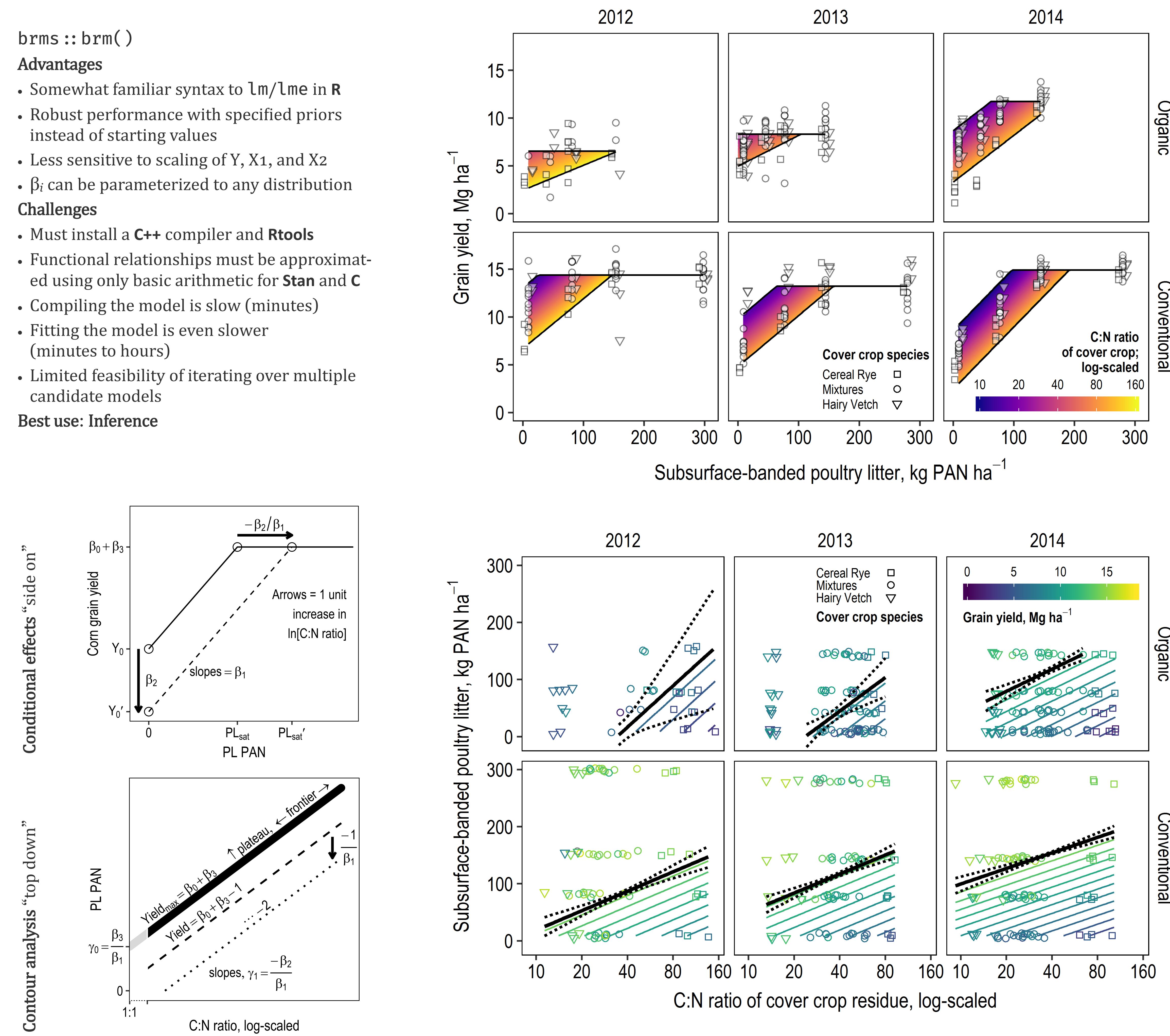
$$\text{Yield} = \beta_0 + \min \left\{ \frac{\beta_1 \times \text{PL}}{\beta_3} + \frac{\beta_2 \times \ln(\text{C:N ratio})}{\beta_3} \right\} + \varepsilon_N$$

$$\text{PL}_{\text{sat}} = \frac{\hat{\beta}_3}{\hat{\beta}_1} + \frac{-\hat{\beta}_2}{\hat{\beta}_1} \times \ln(\text{C:N ratio})$$

$$\text{PL}_{\text{sat}} = \hat{\gamma}_0 + \hat{\gamma}_1 \times \ln(\text{C:N ratio})$$

These figures illustrate two perspectives of the model, generated from the same fit object in R:

- "side-on", the conditional effect of PL at each possible level of C:N ratio, showing the linear-plateau functional relationship up to Yield_{max} and an envelope of the response space as a shaded region
 - "top-down", the frontier of minimum combinations of PL and C:N ratio that achieve Yield_{max}, along with yield isoquant contour lines along the effective response space
- Each is shown for 6 site-years, with the observed data for context.



Conclusions

Each non-linear model-fitting software package has its respective advantages and challenges, and it's ultimately helpful to be familiar with both for understanding functional relationships within your data. The Monte Carlo method we present to derive parameters of biological interest has applications beyond yield-response curves; however, it's important to observe the empirical distributions of such parameters, as they are generally not normal.

The figure showing the effect of PL conditional on C:N ratio has a more familiar appearance, and is more straightforward to interpret qualitatively. However, the figure showing the contours of yield across the response space illustrates quantitatively the relationship of interest: relative trade-off between the two inputs. Each is appropriate for different audiences. The 3D figure shown on the code repository for this poster helps bridge the gap between the two approaches, but is neither as intuitive as the first nor as objective as the second.

```
library(tidyverse); library(mvtnorm)
rmvnorm(n = 10^6,
        mu = fixef(fit.object),
        sigma = vcov(fit.object)) %>%
  as_data_frame() %>%
  mutate(gamma0 = beta3/beta1,
         gamma1 = -beta2/beta1,
         ymax = beta0+beta3) %>%
  gather(key = term, value = est) %>%
  group_by(term) %>%
  summarise_all(funs(mean, sd))
```