يونيرسيتي بروني دارالسلام
UNIVERSITI BRUNEI DARUSSALAM

Regression modelling using I-priors

NUS Department of Statistics & Data Science Seminar

Haziq Jamil
Mathematical Sciences, Faculty of Science, UBD
https://haziqj.ml

Wednesday, 16 November 2022

Overview

Examples

- Longitudinal analysis of growth of cattle
- Predicting fat content in meat samples

Longitudinal analysis of growth of cattle

Aim: Discern whether there is a difference between two treatments given to cows, and whether this effect varies among individual cows.

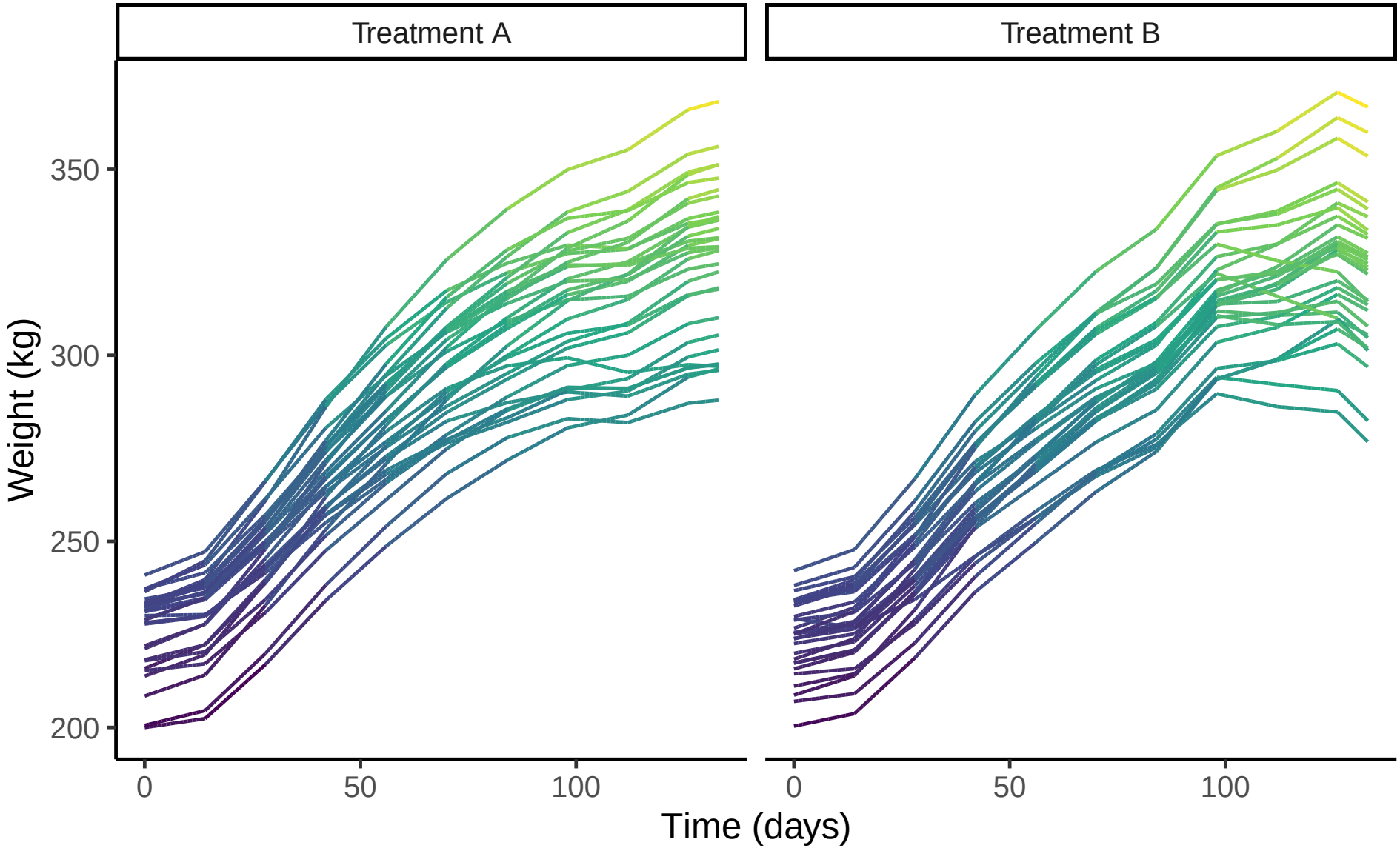
Data consists of a balanced longitudinal set of weights y_{it} for 60 cows. The herd were randomly split between two treatment groups (x_i). Model

$$y_{it} = f_{1t}(i) + f_{2t}(x_i) + f_{12t}(i, x_i) + \epsilon_{it}$$

assuming smooth effect of time, and nominal effect of cow index and treatment group.

	Explanation	Model	Log-lik.	No. of param.
1	Growth due to cows only	f_{1t}	-2792.2	3
2	Growth due to treatment only	f_{2t}	-2295.2	3
3	Growth due to both	$f_{1t} + f_{2t}$	-2270.9	4
4	Growth due to both with cow-treatment variation	$f_{1t} + f_{2t} + f_{12t}$	-2250.9	4

Growth curve



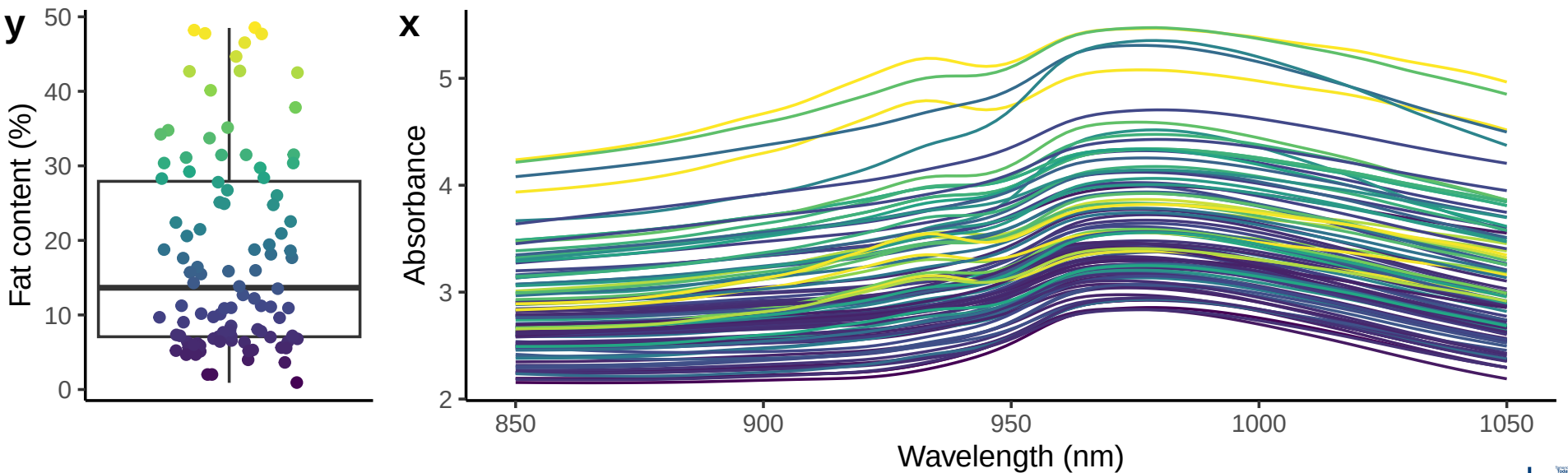
Predicting fat content in meat samples

Aim: Predict fat content of meat samples from its spectrometric curves (Tecator data set).

For each meat sample i, data consist of 100 channel spectrum of absorbances $(x_i(t))$ and its corresponding fat content (y_i). Train/test split is $160 + 55$. Model

$$y_i = f(x_i) + \epsilon_i$$

where x_i is the ith spectral curve.



Results

Model	RMSE Train	RMSE Test
I-prior		
Linear	2.89	2.89
Quadratic	0.72	0.97
Smooth (fBm-0.70)	0.19	0.63
Others		
Linear functional regression		2.78
Quadratic functional regression		0.80
Gaussian process regression		2.93
Neural networks		0.36
Kernel smoothing		1.49
Multivariate adaptive regression splines (MARS)		0.88
Functional additive regression (CSEFAM)		0.85