

Assignment 6

Linear mixed effects models

Due Friday 2nd June 2023

- The format of your assignment should be similar to the previous ones.
- I expect any discussion to be in the context of the question so comments such as “The slopes are significant” for the questions below will not score as well as a comment that references the context of the question etc.
- The assignment can be submitted at any time on that day.

QUESTION 1 Generalised additive models

For this question use the package `gam` in R to fit Generalised Additive Models (GAMs) to some data. The “Date” variable will need to be converted from a character type to Date in R.

```
CCC05$Date <- as.Date(CCC05$Date, "%d/%m/%Y")
```

Should do this, check that it is now a date using

```
str(CCC05)
```

- Two data sets are provided CCC05.csv and ECAN93.csv
- They are measurements taken over a period of time of Nitrate levels in two rivers, one in Christchurch City (CCC05) and one in the Canterbury region (ECAN93)
- Each data set includes
 - `Lawa_or_NRWQN_ID` – an identifier of the site on the river
 - `Date` – The date of the reading
 - `Indicator` – The Indicator measured (always N)3-N_ppb in these data)
 - `Value` – The value of the reading
- For each data set import the data to R and fit a series of GAMs to the Value using a smoother on Date.
- Use at least 4 different spar parameters and discuss which seems to best describe the data series.
- Your output MUST include graphs of the data and fitted values, and discuss the important features of the data.
- Your comments should be brief but each part of the output must be included. (14 Marks)

QUESTION 2 Multiple Comparisons

The data set “Herbicides.xlsx” is a small part of the data collected in an experiment on various herbicides in Dairy pastures. The Variable “Grass_percent” is the percent cover of grass in each plot in the experiment. Eight different treatments were applied, seven herbicides and Nil which can be assumed to be the control treatment.

- (a) Carry out an analysis of variance on the data with Herbicide as the explanatory variable and Grass_percent” as the response.
- (b) Discuss the residuals.
- (c) Carry out an LSD type analysis comparing all possible pairs of treatments. Note which pairs have a significant difference.
- (d) Carry out pairwise comparisons using Bonferroni, Tukey and Dunnett adjustments and in each case show the pairs with significant differences.
- (e) How do the conclusions in (c) and (d) differ? (6 Marks)