Introduction to GLLVM and multivariate GLMM

Provided by: Highland Statistics Ltd

Organised by: Marta Rufino (marta.rufino@ipma.pt), IPMA, Portugal

The central theme of this course is the analysis of multiple correlated response (or dependent) variables using GLMs and GLMMs. Rather than applying multiple univariate GLMs or GLMMs, we will focus on multivariate GLMMs, particularly generalised linear latent variable models (GLLVMs), for the simultaneous analysis of all variables.

During the course, we cover a large number of exercises with examples such as trait variables from turtle hatchlings from multiple clutches, biomass data from fish species sampled at multiple sites, count data from 250 freshwater benthic species sampled at 200 sites, abundances of multiple parasite species on fish, counts of 60 different debris types in water samples, abundances of multiple spider species in traps, multiple morphometric variables sampled from honeybees, and absence/presence of diet variables from faecal samples of brown bears.

In all these examples, we can analyse each variable with a univariate GLM(M). Although these analyses are relatively simple, there are also some problems:

- **Extra Work**: Individual analyses are computationally less efficient and require separate validation, interpretation, and reporting.
- **Lack of Multivariate Relationships**: Analysing the variables individually neglects the interconnected relationships and interactions between them.
- **No Shared Variation**: Univariate models might overlook consistent residual patterns across species, while multivariate models can capture shared variations due to common environmental factors.
- **Multiple Testing**: Conducting separate analyses increases the risk of Type I errors, especially when the response variables are highly correlated.
- **Loss of Community-Level Insights**: Analysing species separately misses out on a comprehensive, community-level viewpoint and can lead to inconsistent conclusions.

Hybrid course onsite/online

Venue: Portuguese Institute for the Sea and the Atmosphere (IPMA). Av. Dr. Alfredo Magalhães Ramalho. Lisbon, Portugal
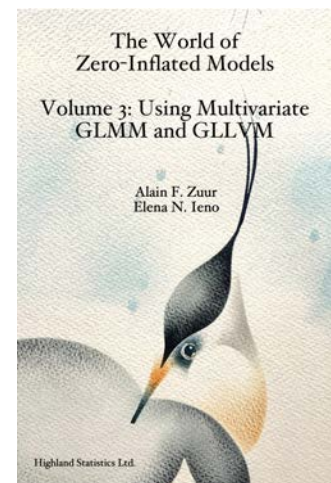
Dates:
- Monday 28 April - Friday 2 May 2025.

Price: 425 GBP.

Included: 1 hour face-to-face video chat about your data.

Instructors:
- Dr. Alain Zuur.
- Dr. Elena Ieno.

COURSE CONTENT

Preparation material with on-demand video:
- Exercise on Poisson and negative binomial GLM.
- DHARMa for model validation.
- Matrix notation.

Monday:
- General introduction.
- A short theoretical presentation revising linear mixed-effects models.
- One exercise on linear mixed-effects models.
- Theory presentation on multivariate GLM and multivariate GLMM.
- One exercise on multivariate GLM(M).

Tuesday:
- Theory presentation on generalised linear latent variable models (GLLVM) for the analysis of multivariate data.
- Three exercises on GLLVM for the analysis of count data (Poisson/negative binomial).

 Wednesday:
- Catching up
- Theory presentation on constrained GLLVM (reduced rank regression and concurrent ordination).
- Three exercises on constrained GLLVM.

Thursday & Friday morning:
- Three exercises using GLLVM with Tweedie, Gamma, and Bernoulli distributions. The Tweedie distribution can be used for continuous data with zeros (e.g., biomass), and the Gamma distribution for continuous data without zeros. The Bernoulli distribution can be used for the analysis of absence/presence data.
- Time allowing: Adding spatial correlation to GLLVMs.

PRE-REQUIRED KNOWLEDGE
You need to have a good understanding of data exploration, multiple linear regression, and Poisson and negative binomial GLMs. Working knowledge of R is required. Familiarity with linear mixed-effects models is not required but is an advantage.

GENERAL
- Please ensure that you have system administration rights to install R, and R packages on your computer. Instructions what to install is on the course website.
- We reserve the right to change the exercises. Pdf files of all theory material will be provided. All exercises consist of data sets and annotated R scripts. Access to the course website is for 12 months. The material covered from Monday to Friday does not contain on-demand video (yet).

For terms and conditions, see:
https://www.highstat.com/index.php/component/hikashop/checkout/termsandconditions/step-3/pos-6/tmpl-component

GENERAL INFORMATION

COURSE FEE: £425
Credit card payments are charged in GBP currency.
VAT Charge:
- UK participants are charged a 20% VAT.
- Non-EU participants (including Norway and Switzerland) are not subject to VAT.
- We do not charge VAT to EU participants who provide their institutional VAT number.
- EU participants who do not provide a VAT number will be charged VAT at their national rate.
- The course fee does not contain coffee, tea, lunch or accommodation.
- Access to the course website is 12 months.

COURSE TIMES:
- Monday - Thursday: 09.00am to 16.30pm including a 1 hour lunch break and a 20 minutes break both morning and afternoon.
- Friday: 09.00- 13.00

COURSE LOCATION
Sala Luis Saldanha (brown room)
Portuguese Institute for the Sea and the Atmosphere (IPMA)
Division of Modelling and Management of Fisheries Resources
Av. Dr. Alfredo Magalhães Ramalho, 6, 1495-165 Lisboa
Just in front of the train stop 'Algés'.

FREE 1-HOUR FACE-TO-FACE MEETING
The course fee includes a 1-hour face-to-face meeting with one or both instructors. The meeting needs to take place within 12 months after the course. You can discuss your own data, but we strongly advice that the statistical topics are within the content of the course. The 1-hour needs to be consumed in one session, and will take place at a mutual convenient time.

CANCELLATION POLICY:
Once participants are given access to course exercises with R solution codes, pdf files of certain book chapters, and pdf files of presentations, all course fees are non-refundable.

REGISTRATION
http://highstat.com/index.php/courses
highstat@highstat.com

Dr. Alain F. Zuur
Highland Statistics Ltd.
9 St Clair Wynd,
AB41 6DZ Newburgh,
UK