Hybrid course (onsite for IMR staff and online for external participants)

Analysing aquaculture data in R

Provided by: Highland Statistics Ltd

Organised by: The IMR Academy, Institute of Marine Research, Bergen, Norway

This is an onsite course for IMR staff and students, with the option for non-IMR participants to join online via Zoom.

This (live) online course consists of 5 modules representing a total of approximately 40 hours of work. Each module consists of live teaching, followed by exercises using real aquaculture data sets.

The exercises will address key biological and environmental processes affecting aquatic organisms and aquaculture systems, including disease dynamics, host–pathogen interactions, environmental stressors, and the impacts of climate change and human activities on aquatic health and productivity.

The course will also address experimental and study design considerations, such as determining appropriate numbers of cages, fish specimens, and parasite load levels (power analysis).

Brief outline

We begin with an introduction to R and provide a protocol for data exploration to avoid common statistical problems. We will discuss how to detect outliers, deal with collinearity and transformations.

Date & Venue

Dates & times:

- 4-8 May 2026
- 08.30-15.30 (Norwegian time)

Price for non-IMR participants: £400

Venue: IMR, Bergen, Norway

Instructors: Dr. Alain Zuur
Dr. Elena Ieno
Authors of 10 books and providers of over 300 courses

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

An important statistical tool is multiple linear regression. Various basic linear regression topics will be explained from a biological point of view. We will discuss potential problems and show how generalised linear models (GLM) can be used to analyse continuous and presence-absence data.

In the second part of the course, we present an introduction to linear mixed effects models and generalised linear mixed-effects models (GLMM) to analyse hierarchical or clustered data, e.g. multiple observations from the same fish, site, area, cage, tank, lake, transect, etc. These statistical techniques are designed to take care of dependency in your data. GLMMs are applied on continuous, binary (e.g. absence/presence of a disease) and proportional data using the Gaussian, Bernoulli, binomial and beta distribution.

Did you already attend our Data Exploration and Linear Regression course?

You're welcome to join us for days 3, 4, and 5, where you'll deepen your skills by learning GLMM modelling using real aquaculture datasets.

COURSE CONTENT

Day 1:
- General introduction.
- Introduction to R.
- Theory presentation on data exploration (outliers, collinearity, relationships, interactions, etc.).
- Two exercises.
- Theory presentation on linear regression.

Day 2:
- Catching up from day 1.
- Bivariate linear regression.
- One exercise.
- Linear regression with multiple categorical covariates.
- Two exercises.
- Multiple linear regression, interactions, visualising covariate effects.
- One exercise.

Day 3:
- Catching up from day 2.
- Theory presentation on logistic regression.
- One exercise (Bernoulli GLM).
- Theory presentation on linear mixed-effects models.
- Two exercises using linear mixed-effects models.
- Time allowing: power analysis.

Day 4:
- Catching up from day 3.
- Two exercises showing how to apply a Bernoulli GLMM for the analysis of presence–absence data.
- Time allowing: power analysis.

Day 5:
- Exercise showing how to apply a binomial GLMM for the analysis of proportional data.
- Exercise showing how to apply a beta GLMM for the analysis of proportional data.
- Time allowing: power analysis.

FREE 1-HOUR FACE-TO-FACE MEETING

The course includes a 1-hour face-to-face meeting with one or both instructors. The meeting needs to take place within 12 months after the last live zoom meeting. You can discuss your own data, but we strongly advise 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.

GENERAL INFORMATION

COURSE TIMES:

- **Monday - Thursday:** 08.30am to 15.30pm including 45 minutes lunch break and a 20 minutes break both morning and afternoon.
- Friday: 08.30am to 14.30pm including 45 minutes lunch break and a 20 minutes break in the morning.

COURSE MATERIAL:

- Pdf files of all presentations are provided.

PRE-REQUIRED KNOWLEDGE:

Basic statistics (e.g. mean, variance, normality). No R knowledge is required. You will learn R 'on the fly'. This is a non-technical course.

GENERAL

- Please ensure that you have system administration rights to install R and required R packages on your computer.
- Instructions on what to install will be provided before the start of the course.
- Course participants will be given access to the course website one week before the start of the course. The website contains course materials, data sets, annotated R scripts, and selected on-demand video materials.
- This is a hybrid course (onsite and online). Online participation will take place via Zoom. Participants will receive a Zoom link and password prior to the course.
- Selected video materials are available on demand and can be watched online multiple times at any time of the day for a period of 12 months.
- A discussion board allows for daily interaction between instructors and participants.
- We will predominantly use the R package glmmTMB.

INFORMATION ON COURSE CONTENT

Dr Alain F Zuur
highstat@highstat.com
www.highstat.com

External participants can sign up at: highstat.com