Analyzing paleolimnological data with R

Isle of Cumbrae, Scotland, 16-20 August 2012

Diego Navarro[1] and Steve Juggins[2]

[1]Instituto de Investigaciones Marinas y Costeras, CONICET, Universidad Nacional de Mar del Plata, Argentina; dnavarro@conicet.gov.ar
[2]School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, UK

Paleolimnology has grown rapidly over the last two or three decades in terms of the number of physical, chemical, and biological indicators analyzed and the quantity, diversity and quality of data generated. Such growth has presented paleolimnologists with the challenge of dealing with highly quantitative, complex, and multivariate data to document the timing and magnitude of past changes in aquatic systems, and to understand the internal and external forcing of these changes. To cope with such tasks paleolimnologists continuously add new and more sophisticated numerical and statistical methods to help in the collection, assessment, summary, analysis, interpretation, and communication of data. Birks et al. (2012) summarize the history of the development of quantitative paleolimnology and provide an update to analytical and statistical techniques currently used in paleolimnology and paleoecology.

R is both a programming language and a complete statistical and graphical programming environment. Its use has become popular because it is a free and open-source application, but above all because its capability is continuously enhanced by new and diverse packages developed and generously provided by a large community of scientists.

This recent workshop, held in the comfortable facilities of the Millport Marine Biological Station, trained researchers on the theory and practice of analyzing paleolimnological data using R. The course was led by Steve Juggins (Newcastle University) and Gavin Simpson (University College London, recently moved to the University of Regina), two of the researchers that have contributed in the development and application of different statistical tools and packages for paleoecology within the R community.

A total of 31 participants from a range of continents (North and South America, Europe, Asia, and Africa), career stages (PhD students to faculty) and scientific backgrounds (paleolimnology, palynology, diatoms, chironomids, sedimentology) enjoyed four long days of training in statistical tools and working on their own data. Initially, participants were introduced to R software and language, tools for summarizing data, exploratory data analysis and graphics. The following lectures and practical sessions focused on simple, multiple and modern regression methods; cluster analysis and ordination techniques used to summarize patterns in stratigraphic data; hypothesis testing using permutations for temporal data, age-depth modeling, chronological clustering, smoothing and interpolation of stratigraphic data;

and calculation of rates of change. The final lectures dealt with the application of techniques for quantitative environmental reconstructions. The theory and assumptions underpinning each method were introduced in short lectures, after which the students had the opportunity to apply what they had learned, to data sets and real environmental questions, during practical sessions. There was also time in the evenings for sessions on important R tips, advanced R graphics, special topics proposed by the assistants, and for the students to work on their own data.

The course was conveniently organized just prior to the 12th International Paleolimnology Symposium (Glasgow, 20-24 August 2012), which enabled all of the workshop participants to attend the symposium and encouraged further discussions throughout the following week. PAGES covered travel and course costs for five young researchers from developing countries (Turkey, South Africa, Macedonia, and Argentina) all of who were very grateful for the opportunity to attend.

References

Birks HJB, Lotter AF, Juggins S, Smol JP (2012) *Tracking Environmental Change Using Lake Sediments: Data Handling and Numerical Technique, Developments in Paleoenvironmental Research 5.* Springer, 745 pp
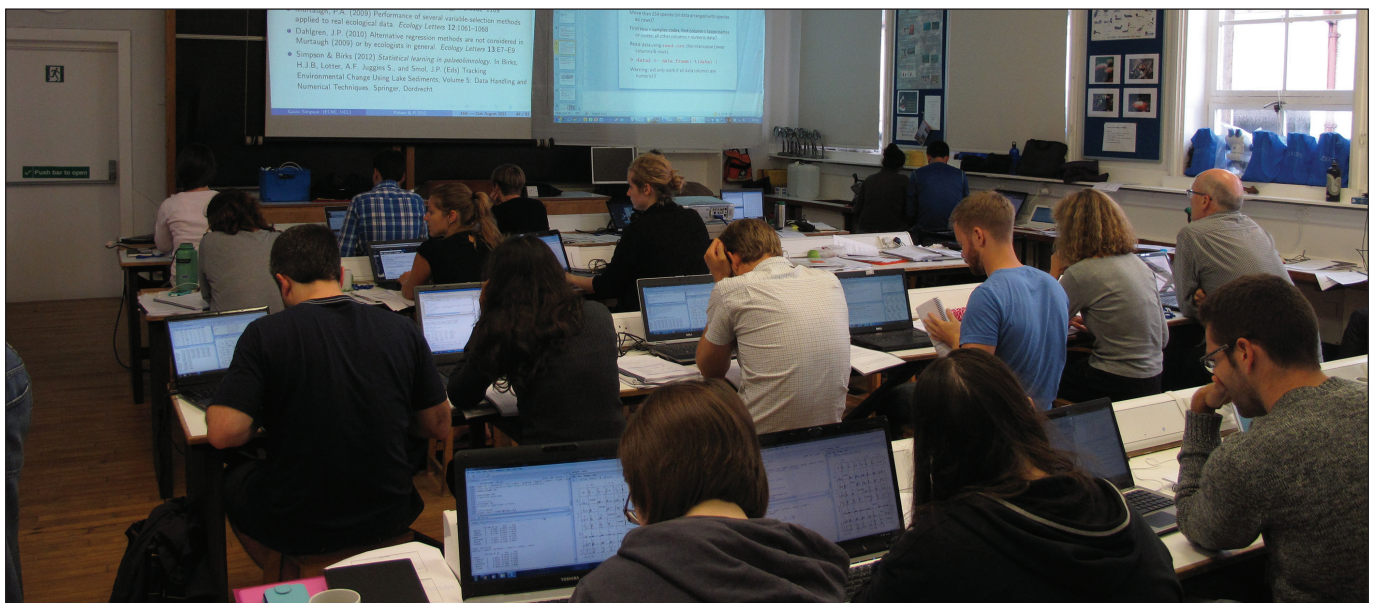
Figure 1: Participants during the R workshop. Photo by S. Juggins.