



Edublog for multivariate statistical analysis of ecological data, using the ade4 graphical interface in R

Edublog para el análisis estadístico multivariado de datos ecológicos, usando la interfaz gráfica ade4 de R

José Luis Sotelo Vásquez¹  

Armando Cervantes Sandoval¹  

Patricia Rivera García¹  

Alejandro J. Perales Ávila¹  

¹ Laboratorio de Aplicaciones Computacionales, FES Zaragoza, UNAM, México.

Submitted: 05-05-2026 | Revised: 23-06-2026 | Accepted: 25-07-2026 | Published: 01-08-2026

Corresponding Author: José Luis Sotelo Vásquez 

ABSTRACT

An edublog is presented, developed to promote the application of multivariate statistical analysis techniques to ecological and biological data. Using the R package ade4 and its graphical interface ade4TkGUI, six of the most commonly used techniques are reviewed: Principal Component Analysis, Simple Correspondence Analysis, Multiple Correspondence Analysis, Canonical Correspondence Analysis, Discriminant Analysis and Cluster Analysis, each illustrated with examples applied to real or simulated datasets. Facilitating the application of multivariate analysis through an accessible graphical interface that allows users to visualize results and interpret graphs without the need for programming. Hosted on the virtual campus of FES Zaragoza, this edublog serves as a teaching aid for the Multivariate Statistics course in the Biology degree program. By comparing programming-based approaches with the use of this graphical interface, it has proven useful in allowing students and academics to access topics considered complex or difficult to understand in a 'user-friendly' manner.

Keywords: Edublog, Multivariate Analysis, ade4.

RESUMEN

Se presenta un edublog desarrollado para promover la aplicación de técnicas de análisis estadístico multivariado a datos ecológicos y biológicos. Utilizando el paquete ade4 de R y su interfaz gráfica ade4TkGUI, se revisan seis de las técnicas de uso más común: Análisis de Componentes Principales, Correspondencias Simples, Correspondencias Múltiples, Correspondencias Canónicas, Análisis Discriminante y Análisis de Clúster, cada una de ellas ilustrada con ejemplos aplicados a conjuntos de datos reales o simulados. Facilitando la aplicación del análisis multivariado, mediante una interfaz gráfica accesible que permite visualizar resultados e interpretar gráficos sin necesidad de programar. Alojado en el campus virtual de la FES Zaragoza, se utiliza como apoyo didáctico para impartir el curso curricular de estadística multivariada de la carrera de Biología, donde mediante una comparación de las opciones de programación y el uso de esta interfaz gráfica ha mostrado su utilidad al permitir a estudiantes y académicos acceder de forma "amigable" a temas considerados complejos o difíciles de entender.

Palabras clave: Edublog, Análisis Multivariado, ade4.

INTRODUCTION

Multivariate analysis groups together a set of statistical techniques that simultaneously examine multiple variables, allowing a more detailed study of the structures, associations, and complex relationships present in the data. Multivariate methods implement procedures that simplify the complexity and dimensionality of the data, offering a more comprehensive and accurate view (1). In ecology, these techniques allow exploring how different environmental and biological factors interact in the distribution and abundance of species, being useful in restoration studies, plant community management, and ecological gradient analysis (2).

Among the most commonly used multivariate techniques are: Principal Component Analysis (ACP), which reduces the number of variables by grouping them into axes that explain a large part of the variation; Simple Correspondence Analysis (ACS), used to detect relationships between species and habitats from contingency tables; Multiple Correspondence Analysis (ACM), oriented toward categorical variables, useful for visualizing frequent combinations and analyzing ecological profiles; Canonical Correspondence Analysis (ACC), which relates environmental and biological variables in order to identify the factors that explain changes in communities; Discriminant Analysis (AD), which assigns observations to predefined groups; and Cluster Analysis, which groups similar observations without the need for previously established categories (3).

The advancement and use of these techniques have been facilitated by computational tools such as the R statistical analysis environment, widely used in ecological studies due to its ability to process large amounts of data, apply complex methods, and create graphical visualizations that facilitate the interpretation of the analysis. Its open-source nature and the support of an active scientific community have facilitated its acceptance, especially due to its capacity to enable reproducible analyses (4,5). In the R environment, among the specialized tools is *ade4*, created specifically for the multivariate analysis of ecological and environmental data. This tool was developed at the Laboratory of Biometry and Evolutionary Biology of Claude Bernard University Lyon and allows graphical visualization of the complex relationships between biological and environmental variables. Furthermore, it has a graphical interface called *ade4TkGUI*, which simplifies its use for users without programming experience, since it allows analyses to be performed and results to be explored through intuitive menus (6).

Due to its visualization-focused design, the graphical interface has contributed not only to the incorporation of multivariate analysis in research, but also to its use as an educational tool to teach statistical techniques in digital environments. This interface reduces the difficulties inherent in programming for students and beginner or occasional users, making it an option for integration into learning platforms supported by digital technologies. Its use allows data exploration and the development of analytical skills through interactive environments centered on the active construction of knowledge (7).

The incorporation of digital technologies in the educational field has generated new ways of acquiring, developing, and applying knowledge. Information and Communication Technologies (ICT) are digital resources that facilitate the storage, processing, and transmission of information. However, their effective implementation depends not only on their availability but on the instructional approach with which they are integrated (8,9). In response to the need to go beyond an approach focused solely on access to information, the concept of Learning and Knowledge Technologies (TAC) emerges. Unlike ICT, TAC seek to transform that access into meaningful and practical learning. These technologies promote the active use of digital resources to strengthen skills, stimulate participation, and support the construction of knowledge. Among them are educational platforms, interactive content, and collaborative tools, whose integration favors student participation and the development of competencies such as problem-solving and collaborative work (11).

In this framework, edublogs represent a specific application of TAC in the educational environment. They are virtual spaces that promote communication, reflection, and collaboration. Their structure allows the dissemination of content, the expression of ideas, and the recording of learning processes in an open and constant manner. In the university context, they have proven effective in encouraging active participation, collaborative work, and the strengthening of digital skills (12,13). Likewise, they stand out for their accessibility, ease of use, and applicability in teaching, constituting a concrete example of how TAC favor the active participation of the student in their own learning (14). The edublog offers a specialized virtual space for the multivariate analysis of ecological and biological data, using the *ade4* statistical package together with its graphical interface. In a complementary manner, the *vegan* library, focused on the analysis of vegetation, was used.

Aimed at students and professionals in Ecology, this digital resource provides an interactive environment that integrates organized content, practical examples, and additional materials. Its

purpose is to facilitate access to advanced statistical tools and the development of skills in the analysis and interpretation of complex problems in ecological and biological contexts (15,16).

This resource is used as supporting material in the multivariate statistics course taught in the Biology degree program at FES Zaragoza.

METHODS

Initially, a documentary and applied work was developed, aimed at developing an edublog as a tool to facilitate the application of multivariate analysis to ecological and biological data, using the *ade4* package of R, particularly in its graphical interface *ade4TkGUI*. The method was structured in three main stages ranging from bibliographic review to digital implementation.

The first stage consisted of a literature review and collection of materials. To this end, a search and analysis of relevant information on the *ade4* library and its main functions was carried out. Specialized books, scientific articles, user manuals, and online resources such as forums, tutorials, and academic platforms were consulted. Likewise, examples developed in R were reviewed, both from the *ade4* graphical interface and from R Commander.

In the second stage, the selection and adaptation of examples was carried out. Suitable examples were identified to illustrate the multivariate techniques of ACP, ACS, ACM, ACC, AD, and cluster. These examples were adapted using ecological and biological datasets included in the *ade4* package, and were documented step by step through the use of *ade4TkGUI*. They are presented as a guide.

The guide was adapted to an interactive digital environment by creating an educational blog that allows navigation through thematic sections, visualization of graphs, code snippets, datasets, and results. This edublog is intended as a complementary resource for those beginning in multivariate analysis and seeking a tool with practical examples oriented toward ecological interpretation.

For the development of the project, various materials were used. Regarding software, the R environment (version 4.3), the *ade4* package, the *ade4TkGUI* interface, R Commander, and the Microsoft Office suite were used. Bibliographic sources included *ade4* manuals and official documentation, scientific articles on multivariate analysis in ecology, specialized books, open-access academic theses, and digital resources such as blogs, forums, and tutorials. Among the complementary tools, web browsers and academic search engines such as Google Scholar, Semantic Scholar, and SciELO were used. The WordPress platform was used for publication.

Regarding the data, we worked with datasets included in the *ade4* library, in addition to simulated databases and examples specifically adapted to illustrate the analyses. Datasets were also generated with artificial intelligence tools to represent practical cases.

RESULTS

The edublog was designed and developed on the digital platform of the Center for the Teaching of Advanced Technologies (CETA) of FES Zaragoza <https://blogceta.zaragoza.unam.mx/ade4> (Figure 1). This space is aimed at students and teachers interested in the use of multivariate analysis of ecological and biological data. For the edublog, 22 pages were designed, showing a minimum of information and the development of examples.

The edublog is organized into thematic sections that allow the user to explore step by step 6 different multivariate techniques. Each blog entry contains a brief theoretical introduction, screenshots of the *ade4TkGUI* graphical interface, applied examples, code snippets, links to the datasets used, and output graphs, which facilitates autonomous learning and replication of the analyses (Figure 2).

The edublog contains six multivariate statistical techniques and six examples of each, organized into independent entries (Figure 3).

The section dedicated to ACP shows the complete procedure for performing this analysis, including data set selection, standardization, generation of graphs, and visualization of the principal axes. Images of the examples with real data and output graphs are included, accompanied by guidelines for their interpretation.

In the case of AC, it documents how to load a contingency table, select correspondence analysis, and visualize the resulting factorial planes. The publication guides the user in the graphical interpretation of associations between rows and columns, using visual resources and screenshots of the analysis environment.

Figure 1. Edublog Multivariate Analysis Ecology

Análisis Multivariado Ecología

Análisis Multivariado de Datos Ecológicos

PAPIME PE212024

Presentación Introducción a R Análisis de Componentes Principales Análisis de Correspondencias Análisis de correspondencias múltiples Análisis de Discriminante Análisis de Correspondencias Canónicas Cluster Créditos

Presentación

Análisis Estadístico Multivariado en Ecología: Herramientas Clave para el Estudio de los Ecosistemas

En la ecología moderna, los datos rara vez pueden describirse con una sola variable. Los ecosistemas están influenciados por múltiples factores ambientales, interacciones entre especies y procesos ecológicos complejos. Para abordar esta complejidad, el **análisis estadístico multivariado** se ha convertido en una herramienta esencial en la investigación ecológica, permitiendo extraer patrones y relaciones a partir de datos multidimensionales.

En este blog, exploraremos las técnicas multivariadas más utilizadas en ecología, incluyendo:

Entradas recientes

- Ejemplo Cluster
- Cluster
- Ejemplo ACC
- Datos de entrada
- Análisis de Correspondencias Canónicas

Comentarios recientes

No hay comentarios que mostrar.

Archivos

- mayo 2025
- abril 2025

The entry corresponding to ACM presents the use of the method on matrices with multiple categorical variables. It shows how to convert the data to complete disjunctive format from the interface, apply the analysis, and visualize the results.

In the case of ACC, the necessary steps to link 2 tables (environmental and biological) from the graphical interface are described. Applied examples, analysis configurations, and joint visualizations of variables and observations are included. The graphs are accompanied by simple instructions that guide the user in exploring relationships between data sets.

The AD entry shows how to assign factors to known groups, run the discriminant analysis, and project the results onto the factorial space. Screenshots of the selection of variables and functions are shown, along with classification plots by colored groups, focused on facilitating their reading and visual interpretation.

Finally, the section on Cluster Analysis describes how to apply different hierarchical methods, adjust the number of groups, and represent the groupings on planes generated by previous analyses. The use of tools such as the group slider, graphical visualization options, and results export is documented.

This material is made available to students, who are asked to develop their own materials to explain some of the reviewed techniques. These works are presented in the same blog (Figure 4).

Figure 2. a) Theory and data set. b) Example with code and use of the graphical interface. c) Graphs generated by the graphical interface.



Figure 3. Pages of the edublog corresponding to each multivariate technique. a) ACP. b) ACS. c) ACM. d) AD. e) ACC. f) Cluster Analysis.

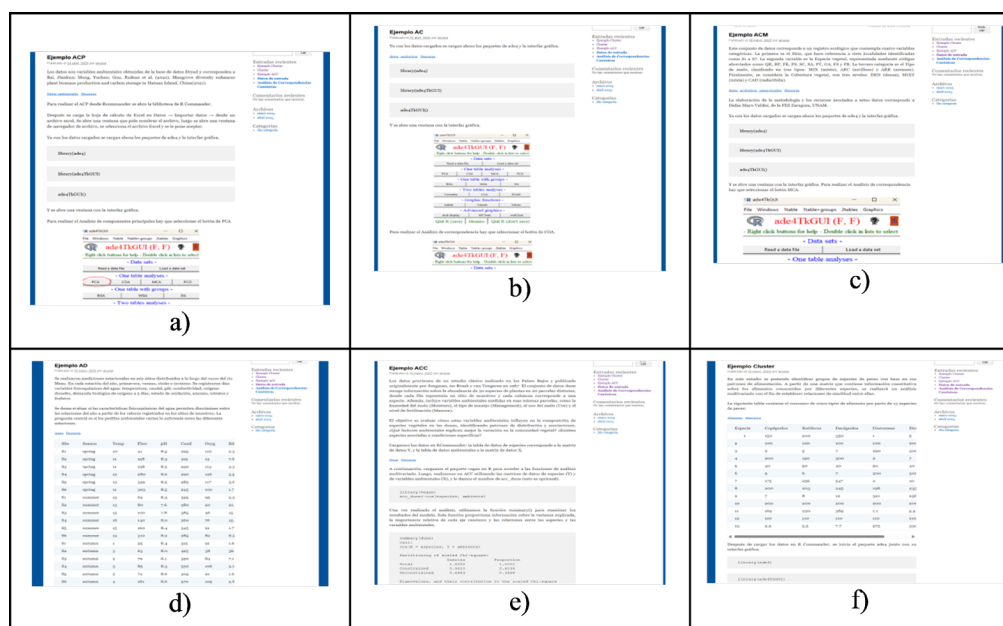
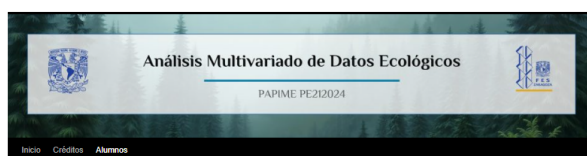


Figure 4. Works developed by students.

Análisis Multivariado Ecología



Alumnos

Se presentan aplicaciones realizadas por alumnos que han cursado la asignatura de estadística multivariada: 1) ACP, 2) Correspondencias, 3) Correlación canónica.

1.- ACP



Buscar

Entradas recientes

- Ejemplo Cluster
- Cluster
- Ejemplo ACC
- Datos de entrada
- Análisis de Correspondencias
- Canónicas

Comentarios recientes

No hay comentarios que mostrar.

Archivos

- mayo 2023
- abril 2023

Categorías

- Sin categoría

DISCUSSION

A didactic resource was generated with the aim of shortening the learning curve by reducing the need to handle R programming code, making multivariate analysis more accessible to new users. By focusing on the use of the graphical interface, the edublog enables students and academics to concentrate on conceptual understanding and analysis of results with minimal programming knowledge. This allows novice and occasional users to apply advanced statistical techniques quickly and with relative ease.

Another contribution of the blog lies in the combination of various multivariate techniques with the visual structure provided by the graphical interface. This tool enables users to perform, within the same environment, analyses such as ACP, ACS, ACM, AD, and Cluster through menus, dialog boxes, and configurable options, facilitating navigation among methods of different complexity.

The review and use of the *vegan* package enables the performance of ACC, expanding the set of available techniques. This combination of packages improves the selection of methods according to the type of data and expands the possibilities for statistical exploration and interpretation. The application of this technique is important in current ecological studies, as it allows relating abundance and diversity data with environmental data.

Although the *ade4* package and its graphical interface were designed for the analysis of ecological and biological data, their application is not limited to these fields. Their versatility and variety of multivariate techniques allow their use to be extended to other disciplines, broadening the potential impact of the edublog and offering the possibility of performing advanced statistical analyses in other contexts.

Although the graphical interface reduces the need for programming to carry out the analyses, it does not completely replace the use of code. In particular, detailed numerical processing of results, such as obtaining coordinates, eigenvalues, or contribution measures, still requires the use of R commands. Therefore, the graphical interface functions as a complementary tool that is especially useful during the exploratory and visual phase of multivariate analyses, but it has limitations in terms of customization and more detailed analysis. Thus, it is aimed primarily at beginner users, although it may fall short for those with a more advanced level.

It has proven its usefulness as a didactic resource, since students have base material to begin studying these topics and performing multivariate analyses on real data or data from the literature, until they reach the development of a topic with the corresponding data handling and application of interpretation criteria.

CONCLUSION

Multivariate analysis is a powerful tool for studying complex data and structures, especially in ecological and biological contexts, where relationships among multiple variables are key to understanding patterns and processes. In this regard, an edublog was developed that not only disseminates knowledge but also promotes the development of analytical skills through the integration of educational technology.

The edublog constitutes an effective application of technology for learning and practicing multivariate analysis by incorporating 6 distinct techniques that allow working with both quantitative and categorical variables. The inclusion of these methodologies expands the possibilities of analysis and facilitates understanding of their theoretical foundations and practical applications.

The strength of the developed material lies in the use of the *ade4* library and its graphical interface, which allows performing multiple analyses within the same platform, reducing the need to write code. Nevertheless, it should be noted that this functionality does not completely replace the use of R programming, since certain advanced analyses and detailed processing of results still require intervention through code.

The *ade4* library represents an effective tool for the statistical analysis of ecological data, thanks to the diversity of methods it makes available to the user. Although it was developed with a specific focus on biology and ecology, its versatility allows its use to be extended to other disciplines, which broadens the impact and potential reach of the edublog. Overall, it was possible to demonstrate how the strategic use of digital tools can facilitate the learning of complex statistical techniques and their application in various fields of knowledge.

As a study resource, it has proven useful, since for some students it is the first time they encounter these topics or use this type of software tools; by the end of the course, they have developed materials on a specific technique with a good academic level that are useful to those who consult them.

REFERENCES

1. Macías WOS, Anchundia JM, Arellano WIM, Córdova FJG. Herramientas estadísticas avanzadas para el análisis de datos en investigaciones cuantitativas: Una revisión sistemática. *Ciencia y Educación*. 2024;5(12):71-89.
2. Martínez M, Pérez J, López R. Análisis multivariado para diseñar la restauración de la llanura aluvial del río Mezquital. *Rev Mex Biodivers*. 2008;79(1):1-12.
3. Flores-Tolentino M, Camacho-Sánchez M, Muñoz-Rodríguez A. Análisis multivariado para el estudio de la distribución geográfica de especies en ambientes ribereños del centro de México. *Darwiniana, nueva serie*. 2023;11(1):391-409.
4. Peng RD. Reproducible research in computational science. *Science*. 2011;334(6060):1226-1227.
5. Lai J, Lortie CJ, Muenchen RA, Yang J, Ma K. Evaluating the popularity of R in ecology. *Ecosphere*. 2019;10(1):025-67.
6. Thioulouse J, Dray S, Dufour A-B, Siberchicot A, Jombart T, Pavoine S. *Multivariate Analysis of Ecological Data with ade4*. Springer; 2018.
7. Thioulouse J, Dray S. Interactive Multivariate Data Analysis in R with the ade4 and ade4TkGUI Packages. *J Stat Softw*. 2007;22(5):1-14.
8. Sánchez-Jiménez M, Vergara D, Barrientos-Fernández A, Nieto-Sobrinó M. Análisis del conocimiento de las Tic, Tac, Tep y Tric en el profesorado universitario. *Rev Eduweb*. 2025;19(1):21-39.
9. Solórzano-Barberán GM. Tecnologías de información y comunicación (TIC) en la educación. *Polo Conocimiento*. 2021;6(3):2246-2260.
10. Rojas Carrasco O, Martínez-Fuentes M, Campbell L. Tecnologías del Aprendizaje y el Conocimiento (TAC) para mejorar los procesos de enseñanza virtual. *Educare Digital*. 2023;23:85-115.
11. Alcívar Bravo EE, García Bucheli CC, Zambrano Ormaza DI, Cedeño Zambrano LP, Segovia García MS. Tecnologías de aprendizaje y conocimiento (TAC) en el proceso de enseñanza-aprendizaje para el desarrollo de las competencias digitales en los estudiantes de Tercero de Bachillerato. *Polo Conocimiento*. 2023;8(6):977-994.
12. Dapía Conde MD, Escudero Cid R. Aprendizaje colaborativo mediante el uso de edublog en la enseñanza universitaria. *Enseñanza Teach*. 2014;32(2):53-72.
13. Molina P, Antolín Jimeno L, Pérez-Samaniego V, Devís-Devís J, Villamón M, Valenciano Valcárcel J. Uso de blogs y evaluación continua del aprendizaje del alumnado universitario. *EDUTEC Rev Electrón Tecnol Educ*. 2013;(43).
14. González A, Rodríguez J. Uso de blogs en la formación inicial de maestros de Educación Infantil: Un estudio sobre la intención de uso futuro. *EDUTEC Rev Electrón Tecnol Educ*. 2019;(68).
15. Braga PHP, Hébert K, Hudgins EJ, Scott ER, Edwards BP, Sanchez Reyes LL, et al. Not just for programmers: How GitHub can accelerate collaborative and reproducible research in ecology and evolution. *Methods Ecol Evol*. 2023;14(6):1364-80.
16. Cabrera CE. El edublog en el contexto educativo universitario. *Cienc Soc*. 2019;44(3):7-23.

FUNDING

PAPIME project PE212024, UNAM

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: José Luis Sotelo Vásquez, Armando Cervantes Sandoval, Patricia Rivera García, Alejandro J. Perales Ávila.

Writing–review and editing: José Luis Sotelo Vásquez, Armando Cervantes Sandoval, Patricia Rivera García, Alejandro J. Perales Ávila.