gosset: An R package for analysis and synthesis of ranking data in agricultural experimentation

## Kauê de Sousa1,2[\*], David Brown3,4, Jonathan Steinke2,5, Jacob van Etten2

1 Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2318 Hamar, Norway  
2 Digital Inclusion, Bioversity International, Parc Scientifique Agropolis II, 34397, Montpellier Cedex 5, France  
3 Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Droevendaalsesteeg 3, 6708 PB, Wageningen, The Netherlands  
4 Digital Inclusion, Bioversity International, 30501, Turrialba, Costa Rica  
5 Humboldt University Berlin, Berlin, Germany

[\*] Correspondence should be addressed to: [kaue.desousa@inn.no](mailto:kaue.desousa@inn.no)

# Abstract

Appropriate data management and analysis are necessary to produce practical information from agricultural experimentation data. There is also an ongoing trend advocating for programmatic tools that supports reproducible workflows in scientific research. We developed the R package gosset, providing functionality to support analysis workflows with rank-based models, such as Plackett-Luce and Bradley-Terry. The gosset package facilitates data preparation, modelling and results presentation stages. We demonstrate the functionality of the package with a case of on-farm evaluations of common bean (*Phaseolus vulgaris* L.) genotypes in Nicaragua.

**Keywords**: Bradley-Terry, Plackett-Luce, data science, on-farm trials, participatory-research, tricot

# Required Metadata

# Current code version

*Code metadata (mandatory)*

| **Nr.** | **Code metadata description** | **Please fill in this column** |
| --- | --- | --- |
| C1 | Current code version | 0.4.003 |
| C2 | Permanent link to code/repository used for this code version | <https://github.com/AgrDataSci/gosset> |
| C3 | Code Ocean compute capsule |  |
| C4 | Legal Code License | MIT |
| C5 | Code versioning system used | Git |
| C6 | Software code languages, tools, and services used | R |
| C7 | Compilation requirements, operating environments & dependencies |  |
| C8 | If available Link to developer documentation/manual | <https://agrdatasci.github.io/gosset/> |
| C9 | Support email for questions | [kaue.desousa@inn.no](mailto:kaue.desousa@inn.no) |

# Motivation and significance

Participatory experimentation approaches have been increasingly applied in agricultural research (De Roo, Andersson, and Krupnik 2019). While collecting data in ranking format is uncommon in general agricultural research settings, it is often collected in participatory experiments (Coe 2002). Recently developed approaches for on-farm experimentation, such as the tricot methodology are based on the collection of data in ranking format (van Etten, Beza, et al. 2019). On the other hand, newly proposed approaches for synthesis of crop variety evaluation data largely depend on the analysis of ranking data (Brown et al. 2020). The analysis of ranking data requires the use of appropriate statistical models such as Plackett-Luce (Luce 1959; Plackett 1975) or Bradley-Terry (Bradley and Terry 1952). Functionality for fitting Bradley-Terry and Plackett-Luce models are available in R with the packages BradleyTerry2 and PlackettLuce respectively (H. Turner and Firth 2012; H. L. Turner et al. 2020). However, extended functionality was required for the entire data science workflow, which usually includes: (1) Data preparation and cleaning, (2) modelling and validation, and (3) results presentation. For (1) gosset provides functions for converting and preparing data into ranking or pairwise format required by the packages PlackettLuce and BradleyTerry2 respectively. For (2), gosset provides functions for model selection and validation using cross-validation. In the case of (3), enhanced functionality for plotting model results is provided by the gosset package.

# Software description

The R package gosset provides functionality supporting the analysis workflows in agricultural experimentation, especially with rank-based approaches. The package is available in The Comprehensive R Archive Network (CRAN) and can be installed by executing install.packages("gosset"). The package is named in honor to William Sealy Gosset, known by the pen name ‘Student’, a pioneer of modern statistics in small sample experimental design and analysis (Ziliak 2019).

## Software Architecture

The R package gosset is structured following the guidelines described in the manual for creating R add-on packages (Team 2022). This structure basically consist of files DESCRTIPTION, LICENSE, NAMESPACE and NEWS, and directories data, dev, docs, inst, man, R, and vignettes. The package functions were developed following the S3 methods style (Team 2022) and are contained in the R sub-directory.

## Software Functionalities

### Data management and preparation

* rank\_binomial transforms a ranking object into a binary comparisons, as required by package BradleyTerry2 (H. Turner and Firth 2012).
* rank\_numeric converts numeric values into rankings. The parameter ascending = indicates if the rankings should be made considering the numeric values in ascending order. The default is asceding = FALSE. This function is useful when the data were collected as numerical observations, for instance, in a experiment measuring crop yield.
* rank\_tricot transforms data in tricot format into rankings.

### Modelling

* AIC computes the Akaike Information Criterion Akaike (1974)] for a Bradley-Terry model fitted with BradleyTerry2 (H. Turner and Firth 2012) or a Plackett-Luce model fitted with (H. L. Turner et al. 2020).
* btpermute deviance-based forward variable selection (Lysen 2009) procedure for Bradley-Terry models.
* compare compare the agreenment between methods.
* crossvalidation performs k-fold cross-validation, where k could be specified by the user.The default is 10-fold. Folds can be provided as a vector for a custom cross-validation, such as blocked cross-validation.
* forward executes forward variable selection with cross-validation.
* kendallTau computes the Kendall-tau rank correlation coefficient between two rankings (Kendall 1938).
* kendallW computes Kendall’s W (coefficient of concordance) among observed rankings and those predicted by the Plackett-Luce model.
* pseudoR2 computes goodness-of-fit measure McFadden’s pseudo-R2 (McFadden et al. 1973).

### Visualization and results presentation

* plot provides an alternative to the default plot function of the PlackettLuce package.
* regret computes the regret [Loomes and Sugden (1982); ].
* reliability computes the reliability of crop varieties, the probability to outperform the reference variety (Eskridge and Mumm 1992).
* worth\_bar creates a bar plot of the estimated worth values for each evaluated variety.
* worth\_map creates a heatmap plot of the estimated worth values for all varieties considering each of the evaluated traits.

# Illustrative Examples

To demonstrate the functionality of the gosset package, we use the nicabean dataset, which was generated with decentralized on-farm trials of common bean (*Phaseolus vulgaris* L.) varieties in Nicaragua over five seasons (between 2015 and 2016). Following the tricot approach (van Etten, Beza, et al. 2019), farmers were asked to test in their farms three varieties of common bean. The varieties were randomly assigned as incomplete blocks of size three (out of 10 varieties). The farmers assessed which of the three varieties has the best and worst performance in nine traits (vigor, architecture, resistance to pests, resistance to diseases, tolerance to drought, yield, marketability, taste) and overall appreciation). The farmers also provided their overall appreciation about the varieties, i.e., which variety has the best and the worst performance based on the overall performance considering all the traits.

We used the Plackett-Luce model implemented in the R package *PlackettLuce* (H. L. Turner et al. 2020). To consider the effect of climate on the variety performance, we used climate data. For obtaining the climate data, we used the *nasapower* package (Sparks 2018). Climatic indices were computed with the *climatrends* package (de Sousa, van Etten, and Solberg 2020).

First, we load the required packages and data. The nicabean dataset contains the tricot rankings (trial), the climate covariates (covar) and the traits observations (traits).

library("gosset")  
library("PlackettLuce")  
library("climatrends")  
library("nasapower")  
library("ggplot2")  
  
data("nicabean", package = "gosset")  
  
dat <- nicabean$trial  
  
covar <- nicabean$covar  
  
traits <- unique(dat$trait)

Make a PlackettLuce rank using the function rank\_numeric.

R <- vector(mode = "list", length = length(traits))  
  
for (i in seq\_along(traits)) {  
   
 dat\_i <- subset(dat, dat$trait == traits[i])  
   
 R[[i]] <- rank\_numeric(data = dat\_i,  
 items = "item",  
 input = "rank",   
 id = "id",   
 ascending = TRUE)  
}

The following lines show the first 10 rankings for the trait vigor, in the format used by the PlackettLuce package.

head(R[[1]], 5)

## 1   
## "INTA Ferroso > Amadeus 77 > IBC 302-29"   
## 2   
## "BRT 103-182 > IBC 302-29 > SJC 730-79"   
## 3   
## "INTA Ferroso > INTA Precoz > INTA Ma ..."   
## 4   
## "BRT 103-182 > IBC 302-29 > ALS 0532-6"   
## 5   
## "BRT 103-182 > IBC 302-29 > ALS 0532-6"

The correlation among rankings can be assessed using the function kendallTau.

baseline <- which(grepl("OverallAppreciation", traits))  
  
kendall <- lapply(R[-baseline], function(X){  
 kendallTau(x = X, y = R[[baseline]])  
})  
  
kendall <- do.call("rbind", kendall)  
  
kendall$trait <- traits[-baseline]  
  
print(kendall)

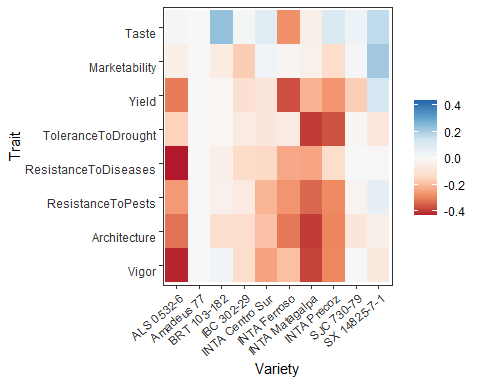
## kendallTau N\_effective trait  
## <dbl> <dbl> <chr>  
## 1: 0.439 58.312 Vigor  
## 2: 0.393 58.312 Architecture  
## 3: 0.463 58.312 ResistanceToPests  
## 4: 0.449 58.312 ResistanceToDiseases  
## 5: 0.411 58.312 ToleranceToDrought  
## 6: 0.749 58.312 Yield  
## 7: 0.639 58.312 Marketability  
## 8: 0.653 58.312 Taste

For each trait, we fit a Plackett-Luce model using the function PlackettLuce from the package of the same name. We use lapply with the list of rankings and the PlackettLuce function as input parameters.

mod <- lapply(R, PlackettLuce)

The worth\_map function can be used to visually assess variety performance based on different traits. The values represented in a worth\_map are worth estimates in the log scale.

worth\_map(mod[-baseline],  
 labels = traits[-baseline],   
 ref = "Amadeus 77") +  
 labs(x = "Variety",  
 y = "Trait")



# Impact

Reproducible and efficient workflows are fundamental in scientific research (Lowndes et al. 2017). The gosset package provides functionality that was not previously available from other R packages and which enabled scientific studies based on the analysis of ranking data. This functionality enables making the entire workflow to be reproducible and more efficient. The utility of the gosset package has been demonstrated by enabling studies based on the analysis of ranking data. For instance, van Etten, de Sousa, et al. (2019), Moyo et al. (2021) and de Sousa et al. (2021) applied the Plackett-Luce model in combination with recursive partitioning (H. L. Turner et al. 2020; Zeileis, Hothorn, and Hornik 2008). In these studies, the gosset package supported data preparation, model validation and results presentation tasks.

# Conclusions

We described the functionality of the R package gosset to support the synthesis and analysis of ranking data. The package provide functions not available in existing R packages for analyzing ranking data. We provided an illustrative example covering the main functionality across the stages involved in the analysis workflow.

# Conflict of Interest

No conflict of interest exists: We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

# Acknowledgements

We acknowledge Olga Spellman (Science Writing Service of the Alliance of Bioversity International and CIAT) for English editing of this manuscript

# References

**Zenodo repository for the gosset package:** <https://doi.org/10.5281/zenodo.6339989>

*Please add the reference to the software repository if DOI for software is available.*

Akaike, Hirotugu. 1974. “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control* 19 (6): 716–23.

Bradley, Ralph Allan, and Milton E Terry. 1952. “Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons.” *Biometrika* 39 (3/4): 324–45.

Brown, David, Inge Van den Bergh, Sytze de Bruin, Lewis Machida, and Jacob van Etten. 2020. “Data Synthesis for Crop Variety Evaluation. A Review.” *Agronomy for Sustainable Development* 40 (4): 1–20.

Coe, Richard. 2002. “Analyzing Ranking and Rating Data from Participatory on-Farm Trials.” In.

De Roo, Nina, Jens A Andersson, and Timothy J Krupnik. 2019. “On-Farm Trials for Development Impact? The Organisation of Research and the Scaling of Agricultural Technologies.” *Experimental Agriculture* 55 (2): 163–84.

de Sousa, Kauê, Jacob van Etten, Jesse Poland, Carlo Fadda, Jean-Luc Jannink, Yosef Gebrehawaryat Kidane, Basazen Fantahun Lakew, et al. 2021. “Data-Driven Decentralized Breeding Increases Prediction Accuracy in a Challenging Crop Production Environment.” *Communications Biology* 4 (1): 1–9.

de Sousa, Kauê, Jacob van Etten, and Svein Ø. Solberg. 2020. *Climatrends: Climate Variability Indices for Ecological Modelling*. <https://CRAN.R-project.org/package=climatrends>.

Eskridge, K. M., and R. F. Mumm. 1992. “Choosing Plant Cultivars Based on the Probability of Outperforming a Check.” *Theoretical and Applied Genetics* 84 (3): 494–500. <https://doi.org/10.1007/BF00229512>.

Kendall, M. G. 1938. “A new measure of rank correlation.” *Biometrika* 30 (1-2): 81–93. <https://doi.org/10.1093/biomet/30.1-2.81>.

Loomes, Graham, and Robert Sugden. 1982. “Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty.” *The Economic Journal* 92 (368): 805–24. <https://doi.org/10.2307/2232669>.

Luce, R. Duncan. 1959. *Individual Choice Behavior.* Individual Choice Behavior. Oxford, England: John Wiley.

Lysen, Shaun. 2009. “Permuted Inclusion Criterion: A Variable Selection Technique.” *Publicly Accessible Penn Dissertations*, 28.

McFadden, Daniel et al. 1973. “Conditional Logit Analysis of Qualitative Choice Behavior.”

Moyo, Mukani, Reuben Ssali, Sam Namanda, Mariam Nakitto, Eric K. Dery, Daniel Akansake, Joseph Adjebeng-Danquah, et al. 2021. “Consumer Preference Testing of Boiled Sweetpotato Using Crowdsourced Citizen Science in Ghana and Uganda.” *Frontiers in Sustainable Food Systems* 5. <https://doi.org/10.3389/fsufs.2021.620363>.

Plackett, R. L. 1975. “The Analysis of Permutations.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 24 (2): 193–202. <https://doi.org/10.2307/2346567>.

Sparks, Adam H. 2018. “Nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for r.” *The Journal of Open Source Software* 3 (30): 1035. <https://doi.org/10.21105/joss.01035>.

Team, R Core. 2022. “Writing r Extensions.” R Foundation for Statistical Computing Vienna, Austria. <https://cran.r-project.org/doc/manuals/r-release/R-exts.html>.

Turner, Heather L, Jacob van Etten, David Firth, and Ioannis Kosmidis. 2020. “Modelling rankings in R: the PlackettLuce package.” *Computational Statistics*. <https://doi.org/10.1007/s00180-020-00959-3>.

Turner, Heather, and David Firth. 2012. “Bradley-Terry Models in R: The BradleyTerry2 Package.” *Journal of Statistical Software* 48 (9): 1–21. <https://www.jstatsoft.org/v48/i09/>.

van Etten, Jacob, Eskender Beza, Lluís Calderer, Kees Van Duijvendijk, Carlo Fadda, Basazen Fantahun, Yosef Gebrehawaryat Kidane, et al. 2019. “First Experiences with a Novel Farmer Citizen Science Approach: Crowdsourcing Participatory Variety Selection Through on-Farm Triadic Comparisons of Technologies (Tricot).” *Experimental Agriculture* 55 (S1): 275–96. <https://doi.org/10.1017/S0014479716000739>.

van Etten, Jacob, Kauê de Sousa, Amílcar Aguilar, Mirna Barrios, Allan Coto, Matteo Dell’Acqua, Carlo Fadda, et al. 2019. “Crop variety management for climate adaptation supported by citizen science.” *Proceedings of the National Academy of Sciences* 116 (10): 4194–99. <https://doi.org/10.1073/pnas.1813720116>.

Zeileis, Achim, Torsten Hothorn, and Kurt Hornik. 2008. “Model-Based Recursive Partitioning.” *Journal of Computational and Graphical Statistics* 17 (2): 492–514. <https://doi.org/10.1198/106186008X319331>.

Ziliak, Stephen T. 2019. “How Large Are Your g-Values? Try Gosset’s Guinnessometrics When a Little ‘p’ Is Not Enough.” *The American Statistician* 73 (sup1): 281–90. <https://doi.org/10.1080/00031305.2018.1514325>.