

## Compositional Data Analysis with Red-R

SERGE-ETIENNE PARENT<sup>1</sup> AND KYLE R. COVINGTON<sup>2</sup>

<sup>1</sup>ERSAM, Department of Soils and Agrifood Engineering, Université Laval, Québec, Canada, G1K 7P4

<sup>2</sup> Translational Biology and Molecular Medicine, Baylor College of Medicine, Houston, Texas, USA, Red-R developer

The compositional analyst must use a series of software to transform raw compositional data and run statistical analyses on them. Tools for compositional data analysis are available in R, an open source widely-used statistical computing environment. However, using R requires prior programming knowledge. Red-R is an open-source, user-friendly visual data flow interface based on R. The interface uses principles of pipeline programming where functions are represented as icons, termed widgets, and data flows from one function to another by drawing lines between them on a canvas. Red-R is able to perform common data analysis tasks (hypothesis tests, analysis of variance, regressions, principal component analysis, data cloud plots, bar plots, biplots, etc.). We have developed a novel Red-R package which implements the *compositions* package in R. Our *compositions* package can be used to perform compositional data operations over raw data (closure, additive, centered and isometric log ratio transformations, perturbations and powering, etc.), and create compositional plots (ternary diagrams, ilr-dendrograms, etc.) without prior programming knowledge, after few basic operations. The objective of this work is to present Red-R and its *compositions* package using an application example for geochemical data. The network of widgets provides an easy-to-follow step-by-step procedure to run a large number of operations available in R, hence facilitating the tasks of the compositional data analyst. Furthermore, the entire analysis network can be saved and reloaded. Reports can be generated from the widget network to document and share results. Non-programmers can have an easy access to the advanced tools available in *compositions* analysis.