# Data reversibility in geology: applications in litho-, magneto- and cyclostratigraphy

Sébastien Wouters[1,2*], Frédéric Boulvain[1], Michel Crucifix[3], Anne-Christine Da Silva[1], Christian Laag[4,5], Matthias Sinnesael[6,7,8], Christian Zeeden[5], Miroslav Zivanovic[9] & Xavier Devleeschouwer[2]

1 Sedimentary Petrology, Liege University, Belgium
2 O.D. Earth and History of Life, Royal Belgian Institute of Natural Sciences, Belgium
3 Earth & Climate, Université Catholique de Louvain, Belgium
4 Institut de Physique du Globe de Paris, Université de Paris, France
5 LIAG – Leibniz Institute for Applied Geophysics, Geozentrum Hannover, Germany
6 Analytical, Environmental and Geo-Chemistry, Vrije Universiteit Brussel, Belgium
7 Department of Geology, Ghent University, Belgium
8 Department of Earth Sciences, Durham University, United Kingdom
9 Electrical Engineering Department, Public University of Navarra, Spain

* Corresponding author: Sébastien Wouters - sebastien.wouters@doct.uliege.be –

**Keywords:** Climatology Sedimentology

Data reversibility is the ability of preserving the data throughout processing. This includes formatting, converting, and analysis. It is particularly challenging in geology due to several factors:

- The unique type of data (nature of the rocks, the fossil content…)
- The necessity of working at high resolution on long datasets
- The lack of data standardisation
- The multiplicity and endemism of software
- The loss of information due to data conversion and/or processing

This makes verification and reproducibility studies particularly difficult. In the framework of a cyclostratigraphy reproducibility study we are establishing, in the open-source programing environment R, tools improving data reversibility in three sub-branches of stratigraphy: litho-, magneto- and cyclostratigraphy.

For lithostratigraphy it is the R package StratigrapheR, designed to generate lithologs and to deal with stratigraphical information (https://CRAN.R-project.org/package=StratigrapheR). It allows to import and process lithological information in a command-line environment. In that environment all the basic data needed to generate lithologs can be accessed, which is a first step towards full data reversibility and curation.

For magnetostratigraphy it is the implementation of a conversion format in R, consisting of a single data table, partly standardised and extensible. It is designed to serve as centralising format, to allow conversion between other formats.

For cyclostratigraphy it is the implementation in R of Empirical Mode Decomposition (EMD) algorithms adapted to astrochronology. As astrochronology is based on the identification of periodic Milankovitch cycles recorded in stratified sedimentary rocks, it uses spectral analysis, mainly through Fourier and wavelet transforms. However the use of these transforms can be awkwardly reversible. To estimate the contribution of each spectral component in the geological record, the signal is often filtered, which allows to assess the behaviour of cycles in a geologically meaningful way. In a similar fashion EMD decomposes a signal into a few components, with specific properties:

- They can be summed back into the signal
- They preserve the original data points
- They can be analysed using the Hilbert transform to compute the instantaneous frequency and amplitude

This allows to perform frequency analysis in a perfectly reversible way.