

The Time Series Analysis of Continuous Proportions: Univariate Case

Barceló-Vidal, C.¹, Aguilar, L.² and Martín-Fernández, J.A.¹

¹ Dept. Informàtica i Matemàtica Aplicada, Campus de Montilivi, Univ. de Girona, E-17071 Girona, Spain

² Dept. de Matemáticas, Escuela Politécnica, Univ. de Extremadura, E-10071 Cáceres, Spain

ORAL PRESENTATION

Abstract: Compositional time series (CTS), i.e., multivariate time series of vectors of proportions, arise in many areas of application where the focus of attention is on the relative, rather than the absolute, values of their components. Such series are characterized by components which are positive and sum to one at each instance in time. Typical examples are the evolution of market shares or the results from polls on political opinions conducted at different instances in time. Although these kind of data constitute multivariate time series, standard modelling techniques are not applicable due to the positivity of the components and the constant sum constraint. In other words, the problems arise because the sample space is not the D -dimensional real space, \mathbb{R}^D , nor the positive real space, \mathbb{R}_+^D , but the $(D - 1)$ -dimensional simplex space, \mathcal{S}^D .

Some authors (e.g. Grunwald et al., 1993) use the Dirichlet distribution to model CTS. Nevertheless, most authors choose to apply the transformations introduced by Aitchison (1986) to the original compositional data, and then analyze the transformed data using the standard techniques of time series analysis. Thus, for instance, Quintana and West (1988) use the centered (or symmetric) log-ratio transformation (*clr*), while Brundson and Smith (1989) prefer the additive logistic transformation (*alr*) for modelling CTS. However, as far as we are aware, nothing has been published in the literature which compares the efficiency or appropriateness of the different transformations for specific case studies or the approaches to statistical modelling.

On the other hand, in the general context of compositional data analysis, it has recently been shown (Barceló-Vidal et al., 2001) that the transformation methodology introduced by Aitchison (1986) is based on a Euclidean structure for the simplex space \mathcal{S}^D in which the perturbation and the powering operations play important and decisive roles. Within this framework, Egozcue et al. (2003) introduced the isometric log-ratio transformation (*ilr*), one that does not suffer from the disadvantages of other compositional transformations. For instance, the *alr*-transformation is not isometric, and the *clr*-transformation is isometric but is zero-sum constrained.

The aim of this work is to update the analysis of compositional time series from a new perspective based on Euclidean geometry and the two operations defined

on the simplex \mathcal{S}^D . This approach will then be used as a means of comparing the methodologies based on the *alr* and *clr* transformations and to establish when they are equivalent. Initially, we will focus on the special case of univariate compositional time series, studying from a compositional perspective ARIMA models arising out of the use of the logistic (or log odds-ratio) transformation.

Keywords: Compositional time series; Additive log-ratio transformation; Centered log-ratio transformation; Isometric log-ratio transformation.

References

- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. London, New York: Chapman & Hall. 416 pp. Reprinted in 2003 by Blackburn Press.
- Barceló-Vidal, C., Martí-Fernández, J.A. and Pawlowsky-Glahn, V. (2001). Mathematical foundations of compositional data analysis. In: *Proceedings of IAMG'01 – The sixth annual conference of the International Association for Mathematical Geology*, p. 20. CD-ROM.
- Brunsdon, T.M. and Smith, T.M.F. (1988). The time series analysis of compositional data. *Journal of Official Statistics*, **14**(3), 237-253.
- Grunwald, G.K., Raftery, A.E., and Guttorp, P. (1993). Time series of continuous proportions. *Journal of the Royal Statistical Society, Series B*, **55**(1), 103-116.
- Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G. and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, **35**(3), 279-300.
- Quintana, J.M. and West, M. (1988). Time series analysis of compositional data. In: *Bayesian Statistics 3*. 747-756, Oxford University Press.