

About some crucial issues in temporal data analysis

Mireille GETTLER-SUMMA – Bernard GOLDFARB

Keywords:

Functional data, multidimensional time series, pre-processing, metrics, relational graph, structural models

Abstract:

Temporal data analysis is a specific context for complex data.

The most relevant particularities are:

- discrete-continuous approaches transitions
- time is an ordered variable, which allows reversibility

A crucial issue is the pre-processing stage which implies at least to distinguish common time intervals or significant change points, seasons or cycles, stationarity.

In the case of multiple times series, the three-way data tables can be displayed:

- in repeating the individuals at the different times
- in repeating the variables at the different times

We present another paradigm which consists in using the functional framework

In the functional approach, the values as well as the shape are key-points for exploring the data and for choosing relevant metrics for further factorial analysis or clustering analysis. We suggest some indices for comparing times series. In the case of discrete time series, we explore the possible link with the Multiple Factorial Analysis point of view.

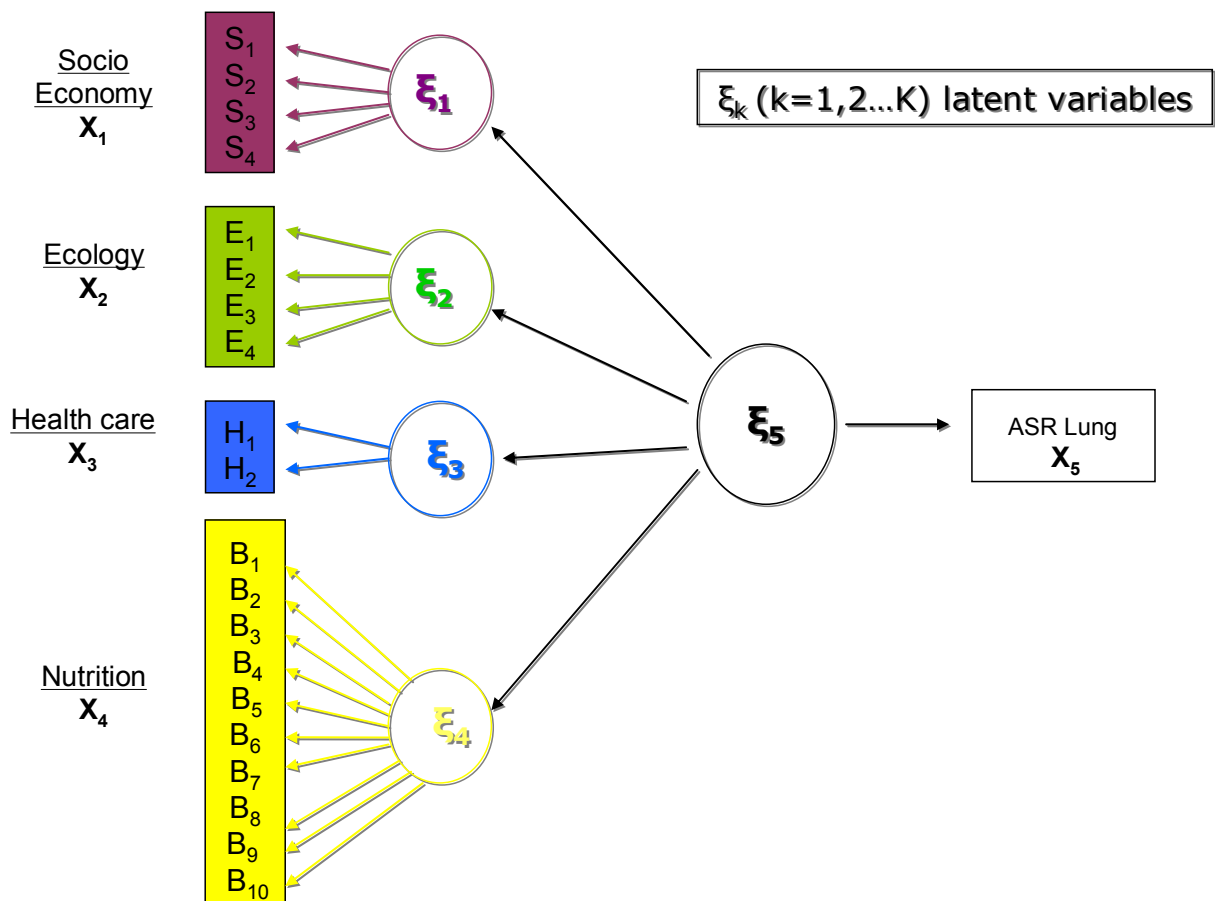
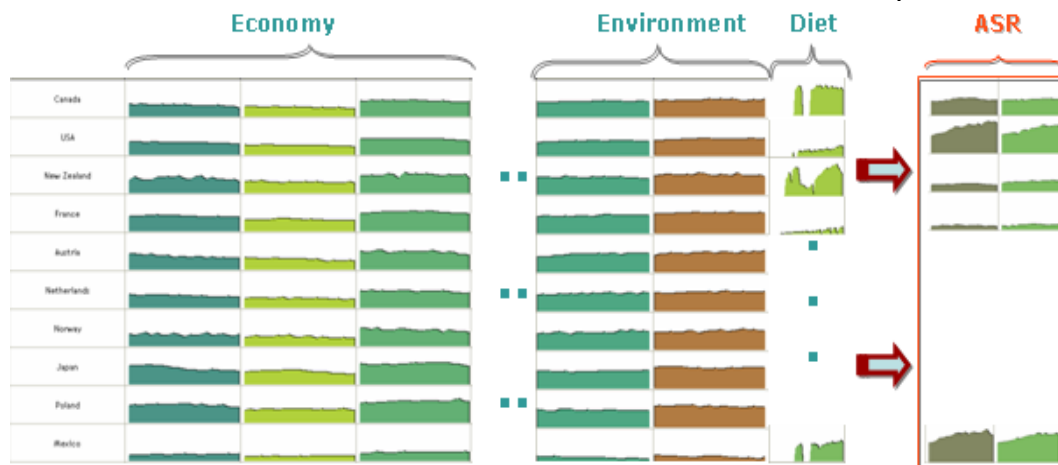
We present in this context the DeltaMetrics® software which has been developed in the CEREMADE laboratory for multiple multidimensional time data. In order to visualise the data table the functional approach has been chosen for the editor.

Exploratory tools in the DeltaMetrics® platform



Among nowadays challenges, a modelling phase problem rises along with a set of multiple time series which are to be learnt from groups of multiple time series, taking into account a relational graph.

As an example comes an epidemiological application which would require a generalization of the structural methods to the functional framework in the case of temporal data



Data stream analysis represents another challenge for learners and for data analysts. In this framework, data is produced over time without any predefined scheme; the issue is a knowledge discovery problem, but classical analysis methods based on fixed patterns hypothesis cannot help in the first stages. However, the key-points of the stability of underlying processes or those of stationarity can be approached using the paradigms corresponding to the pre-processing stage for high-dimensional temporal data.