

# Extension to mixed models of the Supervised Component-based Generalised Linear Regression

Jocelyn Chauvet<sup>1</sup>      Catherine Trottier<sup>1,2</sup>      Xavier Bry<sup>1</sup>  
Frédéric Mortier<sup>3</sup>

<sup>1</sup> Institut Montpelliérain Alexander Grothendieck,  
CNRS, Univ. Montpellier, France.

`jocelyn.chauvet@umontpellier.fr, xavier.bry@umontpellier.fr`

<sup>2</sup> Univ. Paul-Valéry Montpellier 3,  
F34000, Montpellier, France.

`catherine.trottier@univ-montp3.fr`

<sup>3</sup> Cirad, UR BSEF, Campus International de Baillarguet  
TA C-105/D - 34398 Montpellier,

`frederic.mortier@cirad.fr`

## Abstract

We address the component-based regularisation of a multivariate Generalized Linear Mixed Model (GLMM). A set of random responses  $Y$  is modelled by a GLMM, using a set  $X$  of explanatory variables, a set  $T$  of additional covariates, and random effects used to introduce the dependence between statistical units. Variables in  $X$  are assumed many and redundant, so that regression demands regularisation. By contrast, variables in  $T$  are assumed few and selected so as to require no regularisation. Regularisation is performed building an appropriate number of orthogonal components that both contribute to model  $Y$  and capture relevant structural information in  $X$ . To estimate the model, we propose to maximise a criterion specific to the Supervised Component-based Generalised Linear Regression (SCGLR) within an adaptation of Schall's algorithm. This extension of SCGLR is tested on both simulated and real data, and compared to Ridge- and Lasso-based regularisations.

*Keywords:* Component-model, Multivariate GLMM, Random effect, Structural Relevance, Regularisation, SCGLR.