

## Bayesian Estimation of Hierarchical Models for Genotypic Differentiation of *Arabidopsis thaliana*

**Authors :** Gautier Viaud, Paul-Henry Cournède

**Abstract :** Plant growth models have been used extensively for the prediction of the phenotypic performance of plants. However, they remain most often calibrated for a given genotype and therefore do not take into account genotype by environment interactions. One way of achieving such an objective is to consider Bayesian hierarchical models. Three levels can be identified in such models: The first level describes how a given growth model describes the phenotype of the plant as a function of individual parameters, the second level describes how these individual parameters are distributed within a plant population, the third level corresponds to the attribution of priors on population parameters. Thanks to the Bayesian framework, choosing appropriate priors for the population parameters permits to derive analytical expressions for the full conditional distributions of these population parameters. As plant growth models are of a nonlinear nature, individual parameters cannot be sampled explicitly, and a Metropolis step must be performed. This allows for the use of a hybrid Gibbs-Metropolis sampler. A generic approach was devised for the implementation of both general state space models and estimation algorithms within a programming platform. It was designed using the Julia language, which combines an elegant syntax, metaprogramming capabilities and exhibits high efficiency. Results were obtained for *Arabidopsis thaliana* on both simulated and real data. An organ-scale Greenlab model for the latter is thus presented, where the surface areas of each individual leaf can be simulated. It is assumed that the error made on the measurement of leaf areas is proportional to the leaf area itself; multiplicative normal noises for the observations are therefore used. Real data were obtained via image analysis of zenithal images of *Arabidopsis thaliana* over a period of 21 days using a two-step segmentation and tracking algorithm which notably takes advantage of the *Arabidopsis thaliana* phyllotaxy. Since the model formulation is rather flexible, there is no need that the data for a single individual be available at all times, nor that the times at which data is available be the same for all the different individuals. This allows to discard data from image analysis when it is not considered reliable enough, thereby providing low-biased data in large quantity for leaf areas. The proposed model precisely reproduces the dynamics of *Arabidopsis thaliana*'s growth while accounting for the variability between genotypes. In addition to the estimation of the population parameters, the level of variability is an interesting indicator of the genotypic stability of model parameters. A promising perspective is to test whether some of the latter should be considered as fixed effects.

**Keywords :** bayesian, genotypic differentiation, hierarchical models, plant growth models

**Conference Title :** ICCSDA 2017 : International Conference on Computational Statistics and Data Analysis

**Conference Location :** London, United Kingdom

**Conference Dates :** July 24-25, 2017