Title: *Estimating process-based model parameters from species distribution data*

Abstract (350 words max.):

*- Context*: Nowadays, two main types of species distribution models are used to project species range shifts in future climatic conditions: correlative and process-based models. Although there is some continuity between these two types of models, they are fundamentally different in their hypotheses (statistical relationships *vs* cause-to-effect relationships) and their calibration methods (dependent *vs* independent of the species observed distributions).

*- Objectives:* In order to carry out in-depth comparisons of the models*,* our aim here was to calibrate process-based models in the same way as the correlative models, i.e. using the geographic distributions of species. We investigated the feasibility of using an evolutionary algorithm (called covariance matrix adaptation evolution strategy, CMA-ES) to calibrate these models. This method is well established in some fields (robotics, aerospace research, …), but has been seldom, if ever, usedin ecology.

*- Methods:* Using tree species occurrence data across Europe, we adapted the CMA-ES algorithm to find appropriate values of model parameters. We calibrated simultaneously between ten and one hundred parameters of three ecological process-based models (ForCEEPS, PHENOFIT and CASTANEA). We focused our work on three species: *Fagus sylvatica*, *Quercus ilex* and *Abies alba.*

*- Results:* We were able to find a parameter combination that suits best with current species distribution for the three species. CMA-ES was more efficient than a commonly used Approximate Bayesian Computation (ABC) method. For example, with PHENOFIT model, CMA-ES was able to converge to an AUC > 0.9 in less than 24 hours. However, some model parameters and processes were strongly dependent, thus different parameter combinations could lead to high model accuracy.

*- Conclusions:* CMA-ES is an efficient state-of-the-art method to calibrate process-based models with large number of parameters using species occurrence data. This inverse modelling strategy allowed us to efficiently parametrize a distribution-dependent version of each process-based model used in this study. Our next step is to carry out a thorough comparison of different versions of the models, with the final purpose of identifying model weakness and strengths and the causes of their robustness.

Themes: *Forest ecology; Theoretical ecology and ecological modelling; Biogeography*

Keywords: Species distribution model; Process-based model; Calibration; Trees;