An adaptive sampling method using Hidden Markov Random Fields for the mapping of weeds at the field scale

M. Bonneau[(1) (2)], S. Gaba[(2)], N. Peyrard[(1)], R. Sabbadin[(1)]

(1) INRA-Toulouse, unité de Biométrie et Intelligence Artificielle UR875, BP 52627, 31326 Castanet-Tolosan, France ; {mbonneau ; nathalie.peyrard ; regis.sabbadin}@toulouse.inra.fr

(2) UMR INRA-UB-AGROSUP, 1210 unité de Biologie et Gestion des Adventices, 17 rue Sully, F-21065 Dijon, France ; sabrina.gaba@dijon.inra.fr

The mapping of weeds abundance is usually based on direct observations within fields. Full exploration of the field is often impossible, due to its size. A common approach is to divide the field into a regular grid of *quadrats* and to select a fixed number of them for sampling, once and for all. Sampling is designed considering only practical constraints and not, in general, with the objective of maximizing the quality of the reconstructed weeds map. Moreover, the weeds map is generally reconstructed using the *kriging* technique, under the assumption of Gaussian distribution of weeds abundance, even though this assumption is (often) not satisfied. Here we propose (i) an alternative modeling approach to reconstruct weeds maps and (ii) an adaptive sampling method where the visited *quadrats* are chosen sequentially, so as to maximize a criterion incorporating both exploration costs and the quality of the reconstructed map.

Weeds abundance distributions are characterized using the common image analysis tool of *Hidden Markov Random Fields*. This model is naturally used for mapping discrete variables such as weeds abundance classes. It can represent anisotropic phenomena and takes into account possibly noisy observations. Once observations are available, the complete map of weeds density is reconstructed as the most probable map, given the observations.

We first define the quality of a sampling policy from the quality of the *most plausible* reconstructed map using this sampling policy and the associated cost of exploration. The problem is then to compute an *optimal policy* (which has maximum quality). This optimization problem is intractable, so we propose simulation-based algorithms for computing *approximately optimal policies*.

We evaluate this sampling approach, using data consisting in complete maps of weeds densities, resulting from a six years experiment in Dijon-Époisses (France). The method is tested for five weeds species which exhibit a significant spatial correlation (estimated from the data): *Cirsium Arvense*, *Galium Aparine*, *Alopecurus myosuroides Hudson*, *Fallopia Convolvulus* and oil seed rape.

This work will contribute to help agronomist researchers designing sampling strategies and/or comparing the quality of current sampling protocols when the objective is to map weeds species densities.