



## **Combination of decision modelling and data mining to model agricultural primary productivity**

Taru Sandén (1), Aneta Trajanov (2,3), Heide Spiegel (1), Vladimir Kuzmanovski (2), Nicolas P.A. Saby (4), Calypso Picaud (5), Christian Bugge Henriksen (6), Marko Debeljak (2,3)

(1) AGES, Department for Soil Health and Plant Nutrition (BPGE), Vienna, Austria (taru.sanden@ages.at), (2) Department of Knowledge Technologies, Jozef Stefan Institute, Ljubljana, Slovenia (aneta.trajanov@ijs.si), (3) Jozef Stefan International Postgraduate School, Ljubljana, Slovenia (aneta.trajanov@ijs.si), (4) INRA, US 1106, unité Infosol, Orléans France (nicolas.saby@inra.fr), (5) INRA, US 0685, Observatoire du Développement Rural (ODR), Toulouse, France (calypso.picaud@inra.fr), (6) Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, Taastrup, Denmark (cbh@plen.ku.dk)

Agricultural soils provide society with several functions, one of which is primary productivity. This function is defined as the capacity of a soil to produce plant biomass for human use, providing food, feed, fiber and fuel. For farmers, it delivers an economic basis and is a prerequisite for agricultural sustainability. Thus, it is crucial that farmers plan for long-term maintenance of crop yields, without endangering other soil functions such as nutrient cycling or soil being an important habitat for biodiversity.

We develop an agricultural primary productivity decision support model that is designed to help farmers and advisors to assess the primary productivity soil function at a field level as low, medium or high. To obtain an accurate decision support model of primary productivity we addressed the following objectives: i) to construct a qualitative decision support model; ii) to carry out verification, calibration and sensitivity analysis of this model, and iii) to validate the model based on empirical data. Our methodology combined decision modelling for the development of the primary productivity decision model, and data mining, which was mostly used to improve the calibration of the decision model. The developed decision support model has a form of a hierarchical qualitative model constructed of 25 input attributes describing soil properties, environmental conditions, cropping specifications and management practices. For validation, we used an extensive French dataset that contained data from 399 sites. The final overall accuracy of the model was 77% (97%, 71% and 63% for high, medium and low, respectively). Our model aims, as endeavored in the H2020 LANDMARK project, to serve as a generic model for primary productivity that can be used across different environmental zones alongside models for four other soil functions including nutrient cycling, climate regulation, habitat for biodiversity and water regulation and purification. When completed, this approach can contribute to improving integrated sustainability in arable cropping systems.

This study was conducted as part of the LANDMARK (LAND Management: Assessment, Research, Knowledge Base) project. LANDMARK has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635201.