Geophysical Research Abstracts
Vol. 19, EGU2017-15937, 2017
EGU General Assembly 2017

Impact of spatial variability and sampling design on model performance

Charlotte Schrape (1), Anne-Kathrin Schneider (2), Boris Schröder (2,3), and Loes van Schaik (4)

(4) Technical University of Berlin, Institute of Ecology, Ecohydrology and landscape evaluation, Berlin, Germany (l.vanschaik@tu-berlin.de), (1) 1) Universität Potsdam, Institute for Earth and Environmental Science, Karl-Liebknecht-Str. 24-25 14476 Potsdam-Golm, Germany, (2) 2) Technische Universität Braunschweig, Landscape Ecology and Environmental Systems Analysis, Langer Kamp 19c, 38106 Braunschweig, Germany, (3) 3) Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), 14195 Berlin, Germany

Many environmental physical and chemical parameters as well as species distributions display a spatial variability at different scales. In case measurements are very costly in labour time or money a choice has to be made between a high sampling resolution at small scales and a low spatial cover of the study area or a lower sampling resolution at the small scales resulting in local data uncertainties with a better spatial cover of the whole area. This dilemma is often faced in the design of field sampling campaigns for large scale studies. When the gathered field data are subsequently used for modelling purposes the choice of sampling design and resulting data quality influence the model performance criteria.
We studied this influence with a virtual model study based on a large dataset of field information on spatial variation of earthworms at different scales. Therefore we built a virtual map of anecic earthworm distributions over the Weiherbach catchment (Baden-Württemberg in Germany). First of all the field scale abundance of earthworms was estimated using a catchment scale model based on 65 field measurements. Subsequently the high small scale variability was added using semi-variograms, based on five fields with a total of 430 measurements divided in a spatially nested sampling design over these fields, to estimate the nugget, range and standard deviation of measurements within the fields. With the produced maps, we performed virtual samplings of one up to 50 random points per field. We then used these data to rebuild the catchment scale models of anecic earthworm abundance with the same model parameters as in the work by Palm et al. (2013).
The results of the models show clearly that a large part of the non-explained deviance of the models is due to the very high small scale variability in earthworm abundance: the models based on single virtual sampling points on average obtain an explained deviance of 0.20 and a correlation coefficient of 0.64. With increasing sampling points per field, we averaged the measured abundance of the sampling within each field to obtain a more representative value of the field average. Doubling the samplings per field strongly improved the model performance criteria (explained deviance 0.38 and correlation coefficient 0.73). With 50 sampling points per field the performance criteria were 0.91 and 0.97 respectively for explained deviance and correlation coefficient. The relationship between number of samplings and performance criteria can be described with a saturation curve. Beyond five samples per field the model improvement becomes rather small. With this contribution we wish to discuss the impact of data variability at sampling scale on model performance and the implications for sampling design and assessment of model results as well as ecological inferences.