EGU2020-5468, updated on 25 Feb 2021
https://doi.org/10.5194/egusphere-egu2020-5468
EGU General Assembly 2020


# How 'hot' are the hotspots: Statistical approach to localize the high activity areas on soil images

**Nataliya Bilyera**[1,2], Irina Kuzyakova[3], Bahar S. Razavi[2], Sandra Spielvogel[1], and Yakov Kuzyakov[4]
[1]Dept. of Soil Science, Institute of Soil Science and Plant Nutrition, Christian-Albrechts-University of Kiel, 24118 Kiel, Germany (nbilyera@yahoo.com)
[2]Dept. Soil and Plant Microbiome, Institute of Phytopatology, Christian-Albrechts- University of Kiel, 24118 Kiel, Germany
[3]Dept. Ecoinformatics, Biometrics and Forest Growth, University of Göttingen, 37077, Göttingen, Germany
[4]Dept. of Agricultural Soil Science, Dept. of Soil Science of Temperate Ecosystems, University of Göttingen, 37077, Göttingen, Germany

The recently raised topic of microbial hotspots in soil needs not only visualizations of their spatial distribution and biochemical analyses, but also statistical approaches to segregate these hotspots and separate them from the background. We hypothesized that each type of hotspots (e.g. hotspots of root exudation, enzyme activities, root water uptake, pesticides accumulation in plant) is a result of processes driven by biotic or abiotic factors, and consequently corresponds to a statistical distribution follows of a composite functions (e.g. normal/Gaussian), which is significantly different from the background. Consequently, the elucidation of microbial hotspots should be based on statistical separation of the distributions or segregate of maximal values within one distribution. As examples, we collected 3 groups of published images: 1) $^{14}C$ images on carbon input by roots into the rhizosphere, $^{14}C$ localization in roots and glyphosate accumulation in the plant, 2) zymogram on leucine aminopeptidase, 3) neutron image on root water uptake. Each of the images was analyzed for statistical distribution of activity and its area. In the next step, respective distribution parameters (means and standard deviations) were calculated, the modeled distribution was fit, and the background was removed. For the parameters with one distribution, we identified hotspots as the areas outside of the “Mean+2SD” values (corresponding to the upper ~ 2.5% of activity being over 95.5 % of background values). Finally, images of solely hotspots locations were visualised. Comparison with previously used decisions of the hotspot intensity (i.e. Top-25% intensity) thresholding showed advantages of the “Mean+2SD” approach. The advantages (suitable for “time-specific” hotspots in temporal sequence of images, identification of hotspots with different level of activity, unification of thresholding approach for several imaging methods with different principles of activities distribution) and limitations (loss of hotspot areas at low quality images, several thresholding rounds for two or more distributions at on image) of the suggested approach and the potentials of its further development were discussed. We conclude that objective elucidation and separation of the hotspots is case specific and should be based on statistical tools of distribution analysis, which will also help to understand the processes responsible for the highest activities.