Applying machine learning tools to predict trace elements in soils on a global scale

G.D. JONES[1] AND L.H.E. WINKEL[1,2*]

[1] Eawag: Swiss Federal Institute of Aquatic Science and Technology, 8600 Duebendorf, Switzerland
2Institute of Biogeochemistry and Pollution Dynamics, ETH Zurich, 8092 Zurich, Switzerland *lwinkel@ethz.ch

Deficiencies of essential trace elements (i.e., micro-nutrients) affect the health of >2 billion people world-wide. These deficiencies are linked to the uneven distribution of trace elements in soils and differential uptake by plants. Although the factors affecting plant uptake are reasonably well known, global biogeochemical cycling of trace elements as well as the mechanisms controlling their broad scale distributions in soils are poorly understood. To identify these mechanisms, many studies have conducted highly controlled small-scale experiments; however, broad-scale distributions cannot be predicted from such studies. Furthermore, principle component analysis and bivariate correlations are often used to infer cycling mechanisms, but such tools poorly describe synergistic/antagonistic interactions between environmental variables (e.g., clay, pH, etc.) and are thus unsuitable for describing broad scale patterns.

To overcome these limitations, we report on using machine learning tools to investigate the complex relationships between environmental variables and trace element distributions. We will show how the relative importance of mechanisms governing trace element concentrations in soils can be quantified and how concentrations can be predicted for areas where no data on trace element concentrations are available. We use this methodology to infer the dominant mechanisms governing broad-scale trace-element distributions as well as their concentrations in soils. We also demonstrate how anthropogenic changes to the environment may cause widespread changes in micronutrient concentrations and discuss the potential human health impacts.