

Tagungsnummer

V259

Thema

AG Digital Soil Mapping

Neue Methoden in Forschung und Anwendung

Autoren

M. Ließ

Helmholtz-Zentrum für Umweltforschung - UFZ, Bodenphysik, Halle

Titel

Probenahme mittels clhs - Konzept und Einschränkungen

Abstract

"Conditioned latin hypercube sampling" (clhs) hat sich zu einem Standardansatz für die räumlich repräsentative Probenahme entwickelt. Es beruht auf dem "latin hypercube sampling" (lhs)- Algorithmus. Gemäß diesem wird aus jeder von n Variablen eine Stichprobe der Größe k gezogen, so dass die Dichtefunktion der jeweiligen Variablen gut repräsentiert wird. Dies wird erreicht, indem die Dichtefunktion jeder Variable in k Quantile unterteilt und dann zufällig eine Probe pro Quantil gezogen wird. Anschließend werden die so erhaltenen k Werte jeder Variable mit denjenigen der übrigen Variablen entweder zufällig oder auf andere Weise kombiniert. Sofern jedoch diese Variablen die Bodenbildungsfaktoren in einer realen Landschaft repräsentieren (vergleiche clhs), führt die Auswahl eines bestimmten Wertes einer Variablen simultan zur Auswahl von Werten in den anderen Variablen, also zur Auswahl einer bestimmten Landschaftsposition. Um dennoch eine Stichprobe mit genau einer Probe je Quantil je Variable zu erhalten, ist ein mathematischer Optimierungsansatz notwendig. Um die Anwendung von clhs weiter voran zu treiben und insbesondere Probleme zu beheben, die durch die Auswahl physisch mitunter schwer erreichbarer Standorte entstehen, wurden mehrere Anpassungen vorgenommen. Andere Einschränkungen des Algorithmus blieben hingegen weitgehend unbeachtet: So haben die Werte, die in die verschiedenen Quantile einer Variable fallen, nicht die gleiche Wahrscheinlichkeit ausgewählt zu werden, eine Verletzung des lhs-Ansatzes. Ein Problem, das noch verstärkt wird, wenn mehr Variablen berücksichtigt werden oder mehr Probenahmepunkte ausgewählt werden. Auch wenn die Stichprobe in clhs nach wie vor zufällig gezogen wird, sind zudem die ausgewählten Punkte als Folge des Optimierungsprozesses nicht unabhängig voneinander. Und schließlich können aufgrund der Auswahl einer Wertekombination wichtige andere Landschaftspositionen nicht beprobt werden, so dass sie weder in der Stichprobe enthalten sind noch im mit diesem Datensatz erstellten Modell. Die Fähigkeit des clhs die Heterogenität und damit die Pedodiversität einer Landschaft zu erfassen, ist also sehr begrenzt, denn die hier gezogene Stichprobe zeigt eine zu stark vereinfachte Version der Realität und führt daher höchstwahrscheinlich zu verzerrten Modellen. Stratifizierte Zufallsstichprobenverfahren, die beispielsweise auf Methoden der Mustererkennung beruhen, sind für diesen Zweck besser geeignet.