Geophysical Research Abstracts
Vol. 20, EGU2018-2628, 2018
EGU General Assembly 2018
© Author(s) 2018. CC Attribution 4.0 license.

“Big” Data vs. “Small” data – A Scale Puzzle for Scientists and a Scale Dilemma for Decision Makers

Zamir Libohova (1), Phillip Owens (2), Cathy Seybold (1), Skye Wills (1), Dylan Beaudette (1), and David Lindbo (1)

(1) United States Department of Agriculture, Natural Resources Conservation Service, Soil Scinece Division, National Soil Survey Center, Lincoln, United States (zamir.libohova@lin.usda.gov), (2) United States Department of Agriculture, Agricultural Research Services, Arkansas USA

The most common definitions for “big” data refer to volume of data computationally challenging that requires machines to capture patterns and trends for predictions. “Small” data on the other hand refers to volume of data comprehensible by humans to inform often day to day decisions. The development of sensors and platforms for collecting and delivering high spatial-temporal resolution data on natural phenomenon and processes has brought into focus the “big” vs. ‘small” data dilemma and the role of machine learning in natural sciences. Traditionally, the scientists’ focus has been on finding cause-effect relationships through controlled experiments (small data). However, the complexity of natural processes highlighted by “big” data capabilities has limited its use to often finding correlations only. Apart from exploring and understanding the world around us, one of the main goals of research is to support decisions at every level of human activities. At what point a correlation becomes a cause-effect able to quantitatively guide routine decisions implies a scale factor related to both natural processes and the economy. We argue that “big” and/or “small” data mining and analysis can serve both the scientific community and policy/decision makers. Through examples from soil science discipline we illustrate how the assessment of errors and uncertainties viewed from the scale perspective can feed into policy and management decisions.