Using machine learning to estimate snow distribution

David Buckingham Christian Skalka, Ph.D. Josh C. Bongard, Ph.D.

Abstract. Estimating the amount of snow on the ground over a mesoscale area, e.g. a water drainage basin, has uses from predicting spring runoff to the study of climate. Single-point snow water equivalent (SWE) sensors are installed at sites throughout the world. Single-point SWE measurements from such sensors are commonly used as estimates for areal SWE, the amount of snow distributed over a larger area. However, because snow distribution over an area is usually non-homogenous, SWE at one point may not be representative of SWE over an area, causing high error with this simple method. Many factors, such as meteorology, ground topography, and tree coverage may interact in complex, non-linear ways to determine how snow is distributed over an area, and hence the relationship between point SWE and areal SWE. This work in progress uses machine learning to better understand this relationship and improve areal SWE estimates. We use symbolic regression methods, such as Genetic Programming (GP) and Fast Function Extraction (FFX) to develop models of the non-linear relationships between point SWE, various factors that influence snow distribution, and areal SWE. We expect that such models can provide more accurate areal SWE estimates that do other methods.