Earth **Surface**
Dynamics
Discussions

Interactive comment on "**An introduction to learning algorithms and potential applications in geomorphometry and earth surface dynamics**" *by* **Andrew Valentine and Lara Kalnins**

E. Goldstein

evan.goldstein@unc.edu

Received and published: 25 February 2016

In section 4.1 the authors discuss training data. They state, on P. 17 L. 1-2:

"To avoid problems, it is important to choose training data with care, and to develop strategies for evaluating and monitoring the performance of the trained system."

My own experience with machine learning suggests that this is crucial: if the training data is bad (e.g., not enough, too much, too noisy, doesn't cover the solution space adequately, etc.), the learner cannot do its job effectively (hence the expression 'garbage in, garbage out'). My recommendation is the authors present specific guidance on how to select appropriate training data and point the reader toward such guidance in the

Printer-friendly version

Discussion paper

literature (i.e., selection routines). Note that this goes beyond offering advice on how to represent data to the learner (discussed on p. 17, line 2-8).

Respectfully,

Evan B Goldstein

Department of Geological Sciences

University of North Carolina-Chapel Hill

Printer-friendly version

Discussion paper