Biogeosciences Discuss., 9, C92–C94, 2012
www.biogeosciences-discuss.net/9/C92/2012/


Biogeosciences Discussions

# *Interactive comment on* "Technical Notes: Calibration and validation of geophysical observation models" *by* M. S. Salama et al.

**Anonymous Referee #2**

Received and published: 15 February 2012

General comments:

In this paper, the authors describe a calibration/validation method, which is inspired on the traditional cross-validation framework. The developed method is illustrated based on two datasets from the field of remote sensing. The development of techniques supporting calibration/validation and uncertainty assessment is important and should receive sufficient attention in the literature. However, I have a few questions about the proposed method in this paper:

(1) The proposed framework is claimed to be novel (and I do not necessarily disagree on that), however, it is strongly related to what is called 'bootstrapping' in the classical statistical literature. Since its introduction by Efron in the early eighties, bootstrapping procedures have been studied extensively in the statistical literature as well as in



applied domains. As such, I believe that a new method should be evaluated in comparison with this established framework. Nevertheless, the authors do not mention the existence of this well known procedure. A better situation of the method with regard to existing techniques is essential.

(2) The authors introduce a sampling procedure and use this procedure to approximate the distribution of several statistics and model parameters. In classical linear regression, the sample size is very important with respect to the distribution of a statistic. For instance, in case of a regression model, the distribution of the slope (and its confidence interval) is heavily influenced by the sample size. Typically, knowing the distribution of a statistic for a given sample size is important. The approximated distribution that arises when parameters are repeatedly estimated based on samples of varying size might not be very informative. Stated differently, I do not immediately see what kind of relevant 'statistical' question can be answered based on such a distribution.

(3) The authors illustrate their method using two datasets. In both settings, they use a simple linear regression model. For these models, the distributions of the model parameters can be derived theoretically. Basically, this derivation only holds for normally distributed error terms; however, since both datasets contain at least 75 observations, validity can be assumed based on the central limit theorem. Moreover, the obtained t-distribution further supports this claim (as it can be the result of mixing a large number of Gaussians). Bearing this in mind, why do the authors make use of a more complicated and computationally demanding sampling scheme instead?

Specific comments:

- p317.17: I think this formula should be: 'nr = n-2kmin+1'

- p318.1: The nuci becomes larger for an increasing number of cal pairs ki. Hence, matchups with a large cal set eventually contribute more to the derived distributions. What is the effect of the variable nuci on the distributions, and how would these distributions look like if a fixed nuc would be used?

- p319.17: The best results may possibly show randomness regarding the number of cal/val pairs; however, I expect that the worst results will not show such randomness, and will mainly occur with unbalanced cal/val sets. Therefore, I do not agree that an optimal setup for subdividing matchups into cal/val sets cannot be defined, and that the only objective approach is to evaluate all possible combinations.

Technical corrections:

- p313.22: applications ‘use’ or ‘make use of’

- p320.25: Y is the corresponding remote sensing observations: change to ‘set of’, or contains the ...

---

Interactive comment on Biogeosciences Discuss., 9, 311, 2012.