

Interactive comment on “Optimal estimator for assessing landslide model efficiency” by J. C. Huang and S. J. Kao

E. Foufoula (Editor)

efi@umn.edu

Received and published: 10 August 2006

This paper proposes to revisit a common index for evaluating landslide model efficiency with a new proposed index and presents a simulation study in which these indices are used as metrics in a model parameter optimization framework, for purpose of comparison. Specifically, the commonly used index called the success ratio SR, is the conditional probability of predicting a cell as unstable given that it is truly unstable $SR = p(\hat{u}|u)$. The new index, called modified SR or MSR, includes also the conditional probability of predicting a cell stable given that it is indeed stable. The proposed index is $MSR = \frac{1}{2} [p(\hat{u}|u) + p(\hat{s}|s)]$ where the $\hat{\cdot}$ indicates prediction (via a model). The paper spends several pages to convey the above message and in my opinion this is not only confusing the issue and effectiveness of the message, but also casts uncertainty in the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

subsequent presentation. A few comments follow.

1. In the simulation study, the authors create landslide maps by choosing the landslide to cover randomly 5, 10, and 15% of the total area – however, we know that there is some coherence in the landslide coverage and ignoring this coherence or spatial dependence is bound to affect the results. Then, a large number of susceptibility maps is generated for each artificial landslide map. Two parameters control the generation of these maps: success in stable and unstable cell prediction, that is $p(\hat{u}|u)$ and $p(\hat{s}|s)$. My understanding is that what is explicitly prescribed in the artificial susceptibility maps is nothing but the two indices, SR and MSR, we seek to evaluate (see the formulae above). Thus, the performance of each map is already prespecified and it is not surprising that the two performance indices give different results. In fact, no simulation is even needed to determine these performances. For example, note in Fig. 1 that the (g), (h) and (i) plots are nothing than $\beta = p(\hat{u}|u)$ vs. $\alpha = p(\hat{s}|s)$ and it is expected that $MSR = \frac{1}{2}(\alpha + \beta)$ will be around the 1:1 line with a band around it that depends on the unconditional number of landslides. At the same time, the top figures that gives only SR are expected to be horizontal bands since SR is not dependent on $\alpha = p(\hat{s}|s)$ which is the parameter on the horizontal axis.
2. I do not see the point in presenting the Kappa index and showing how bad it does.
3. The other contribution of this paper is proposing this modified index for adoptive model calibration (i.e., calibrate model parameters based on past observed cases and model performance optimization as judged by these metrics). I have some reservations with the process (pg. 10) of randomly selecting parameter combinations as these parameters are definitely dependent on each other and this can significantly influence the results.
4. Some other comments:

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

- The conclusion presented at the last sentence of the abstract seems obvious to me given the formulae of SR and MSR as given in the review above. That is, if the number of true stable cells is small then the extra term that the modified index adds has minimal effect and SR will compare well with MSR, etc. Yet the authors present elaborate simulations to arrive at such conclusions?
- The precise number of 3969 landslide maps, etc. escapes me.
- Pages 3 and 4 – confusing discussion of very simple concepts.
- The test on effect of cell aggregation can be interesting but not in the limited way the authors perform it. Unless there is some significant coherence in the landslides (which is true in reality) aggregation effects will be minimal. Given that randomness has been used in the simulations, I believe that the results on aggregation can be misleading (lower bounds, if anything, of reality).

Overall, I think the paper has some merit (in introducing the prediction of stable cells in the evaluation process), but due to the limitations of the presentation and the simulation results, I would not recommend publication of this paper as-is.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 1125, 2006.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper