Hydrol. Earth Syst. Sci. Discuss., 3, 1125–1144, 2006
www.hydrol-earth-syst-sci-discuss.net/3/1125/2006/
© Author(s) 2006. This work is licensed
under a Creative Commons License.

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Optimal estimator for assessing landslide model efficiency

J. C. Huang and S. J. Kao

Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan

Received: 20 April 2006 – Accepted: 22 May 2006 – Published: 23 June 2006

Correspondence to: S. J. Kao (sjkao@gate.sinica.edu.tw)

Abstract

The often-used success rate (SR) in measuring cell-based landslide model efficiency is based on the ratio of successfully predicted unstable cells over total actual landslide sites without considering the performance in predicting stable cells. We proposed a modified SR (MSR), in which we include the performance of stable cell prediction. The goal and virtue of MSR is to avoid over-prediction while upholding stable sensitivity throughout all simulated cases. Landslide susceptibility maps (a total of 3969 cases) with full range of performance (from worse to perfect) in stable and unstable cell predictions are created and used to probe how estimators respond to model results in calculating efficiency. The kappa method used for satellite image analysis is drawn for comparison. Results indicate that kappa is too stern for landslide modeling giving very low efficiency values in 90% simulated cases. The old SR tends to give high model efficiency under certain conditions yet with significant over-prediction. To examine the capability of MSR and the differences between SR and MSR as performance indicator, we applied the SHALSTAB model onto a mountainous watershed in Taiwan. Despite the fact the best model result deduced by SR projects 120 hits over 131 actual landslide sites, this high efficiency is only obtained when unstable cells cover an incredibly high percentage (75%) of the entire watershed. By contrast, the best simulation indicated by MSR projects 83 hits over 131 actual landslide sites while unstable cells only cover 16% of the studied watershed.

1 Introduction

Landslides triggered by intensive rainfall caused serious damage producing thousands of deaths and billions of dollars in property losses every year. To mitigate these damages, deterministic and non-deterministic (stochastic) models have been developed to generate maps of susceptibility, danger or risk for landsliding (e.g. Ayalew and Yamagishi, 2005; Bora et al., 1998; Pack et al., 2001). To ensure the model efficiency all

modeled landslide maps must be validated by actual landslide map, which is the end-product resulted from real landform processes. To validate a landslide model, to have a proper index or estimator for measuring performance (or efficiency) is crucial.

Meanwhile, the efficiency indicator functions as likelihood measure in model calibration. To optimize parameters, consistent sensitivity is required when estimator is applied onto model results with various landslide patterns. Previous studies (Montgomery and Dietrich, 1994; Dietrich et al., 1995; Borga et al., 1998; Duan and Grant, 2000; Lee, 2005) had been using "success rate" (hereafter, SR) to evaluate the model efficiency. The SR is defined as a ratio of how many actual landslide locations are successfully predicted. However, the SR estimation does not include the success (or failure) in stable cell prediction inherited in the model. In this study, we modify the SR (hereafter, MSR) by combining the success of stable cell prediction into efficiency estimation. We generate 3969 landslide susceptibility maps by using C^{++} language and create 9 man-made actual landslide maps. These two groups of maps are used to reveal how SR and MSR respond to model results that have wide degree of performance under different landslide pattern. The kappa statistic, which is used for measuring interobservers, and logic concepts in image similarity analysis are discussed. For further annotation, we present a real case to illustrate the advantage of this newly proposed MSR.

2 Methods and logics in image similarity analysis

As comparing the actual landslide maps with the predicted maps, four types of outcome in terms of logic aspect may appear: Type 1) actual landslide cells are predicted as unstable; Type 2) actual landslide cells are predicted as stable; Type 3) actual stable cells are predicted as unstable; and Type 4) actual stable cells are predicted as stable. The Type 1 and Type 4 are taken as success in model prediction. Both Type 2 and Type 3 are regarded as failure in prediction. Generally speaking, the performance assessment in landslide susceptibility map is similar to the similarity measurement in

image analysis. Thus, we drew the kappa statistic for comparison. Below, we overview methods with different types of logics in assessing image similarity and proposed a new one.

2.1 The previous success rate

The previous success rate calculation only takes Type 1 outcome into account and ignores the other three types. The equation of SR:

$$\mathrm{SR} = \frac{\text{successfuly predicted landslide numbers}}{\text{actual landslide numbers}} \quad . \tag{1}$$

In this equation landslide number instead of landslide cells is used in efficiency calculation. The reason is that one can only expect the prediction partially overlaps the actual landslide areas that resulted from all complex processes such as triggering and transporting. Nevertheless, in this equation only unstable cell prediction is considered.

2.2 The kappa statistic

Kappa value is a precise index to quantitatively measure the magnitude of agreement among interobservers (e.g. Viera and Garrett, 2005). The equation of kappa:

$$\mathrm{kappa} = \frac{\mathrm{K}_a - \mathrm{K}_e}{1 - \mathrm{K}_e} \tag{2}$$

where K_a is actual agreement and K_e is expected agreement. Concept of this calculation is based on the difference between how much agreement is actually presented (actual agreement) compared to how much agreement would be expected by chance alone (expected agreement). When kappa is applied to landslide prediction actual agreement is the sum of the probability of Type 1 and Type 4 outcomes. Theoretically, the maximum value of actual agreement is one; thus, the lower the actual agreement

the less outcomes of Type 1 and Type 4. Expected agreement is the sum of the expected values of predicted actual unstable and stable cells. The expected agreement explains the expected value of the two interobservers in each class in a random space. Two images are in moderate agreement when the kappa value is larger than 0.4. Perfect congruence presents when kappa value equals 1.0.

2.3 Modified success rate (MSR)

Here we introduce Type 4 outcomes, the success of stable cell prediction, into the previous SR calculation. The efficiencies for landslide site prediction (old SR) and non-unstable cell prediction are equally weighted. The MSR is defined as the following equation:

$$\mathrm{MSR} = 0.5 \times \mathrm{SR} + 0.5 \times \frac{\text{successfully predicted stable cells}}{\text{total number of actual stable cells}} \quad . \tag{3}$$

The value of the MSR can range from 0.0∼1.0. The incorporation of Type 4 outcomes promotes the role of stable area prediction in measuring modeling efficiency, and thus substantially reduces the potential of over-prediction.

3 Generate actual landslide maps and model simulations

To examine differences among these three methods on landslide prediction, we generate three man-made landslide maps in a 20*20 matrix, in which landslide randomly covers around 5.0%, 10.0%, and 15.0% area in the matrix. These artificial landslide maps serve as actual landslide map for model efficiency calibration. The area percentage is limited to 15% since landslide area rarely occupied >15% area in most natural watersheds (Carrara et al., 1995). At each landslide level, we further generate 3 maps with different degree of landslide aggregation. Different aggregation patterns

might lead to changes in landslide numbers (Table 1). Through these cases, we assess effects of aggregation pattern on the efficiency calculation. Features of the nine man-made landslide maps are summarized in Table 1.

As for model results, we generate landslide susceptibility maps by using C^{++} language. Based on the nine man-made landslide maps, we use a generator with duel-parameter, "a" and "b", to create susceptibility maps. The two parameters "a" and "b" represent, respectively, the ratio (from 0.0∼1.0) of success in stable and unstable cell prediction. The spacing for both "a" and "b" are set to be 0.05; thus, 441 susceptibility maps are randomly generated for each artificial landslide map. Since occurrences of unstable and stable cells are fixed in any given man-made landslide map, the generator allows us to create susceptibility maps with full range in model efficiency. For instances, a reverse susceptibility map would be generated as the parameter set is (0.0, 0.0). On the contrary, perfectly congruent case will be derived basing on the parameter set of (1.0, 1.0). The set (0.0, 1.0) will give a map full of unstable cells; therefore this map totally fails in stable cell prediction. A total of 3969 cases are obtained and used to evaluate the feasibility of the three methods.

4 How estimators respond to prediction errors

Model efficiencies derived by three methods are presented in contour patterns (Fig. 1). In each case, we have method-derived model efficiency (Z), which is plotted against the degree of success in predicting stable (X) and unstable (Y) cells. The x-axis is defined as the fraction of successfully predicted stable cells over total actual stable cells, and y-axis is the fraction of successfully predicted unstable cells over total unstable cells (in fact, the X and Y are, respectively, the parameter "a" and "b" in our generator). The interpolation and contour pattern are obtained by using Kriging. The contour pattern enables us to evaluate efficiency distribution respective to prediction errors along both axes.

Since contours are found similar at different aggregation levels, in Fig. 1, we only

present contour at middle level of aggregation for individual method. By contrast, results show significantly different contour pattern among methods (Fig. 1). In all three figures, isopleths show curled feature partially due to Kriging interpolation and partially due to different bases of x-axis (cell number) and y-axis (landslide number). For SR method, Z values increase as the increasing of Y regardless of changes along the x-axis (Figs. 1a–c). Apparently, the old success rate is insensitive to errors in predicting stable cells that are equally important in landslide models. Thus, modelers could obtain very high efficiency even with the worst prediction in stable cells.

Figures 1d, e and f show contours of kappa-derived efficiency. The grayish band of moderate efficiency shifts a little in three plots for kappa, yet, all contours reveal that ~90% Z values are lower than 0.4. Higher Z values appear at the very upper right corner approaching (1.0, 1.0) where only near-perfect success at both axes can reach. The isopleth of zero connected (0.0, 1.0) and (1.0, 0.0) with negative values at the left part. Obviously, this method is too stern in terms of practical usage.

Contours derived from MSR (Figs. 1g, h and i) are much like that derived by kappa yet with smoother gradient. Similarly, the higher efficiencies occurred at the upper right corner. 50% simulated cases have MSR-derived efficiencies higher than 0.5. The moderate efficiency (grayish band) distributed diagonally from (0.0, 1.0) to (1.0, 0.0). Compared to the old SR, the MSR is sensitive to both axes, potentially, useful to avoid over-prediction of unstable cells.

5 Model efficiency as guidance for calibration

Practical steps for calibrating parameters often start with random combinations of parameters (e.g. Duan and Grant, 2000; Zhou et al., 2003). Efficiency indicator is then used to evaluate model outcomes in order to back calibrate parameters. Hence, the efficiency index may guide the direction of model calibration and act as sorter retrieving optimal results

According to contours of model efficiency, we point out two directions in model cali-

bration. The development moving from (0.0, 0.0) toward (1.0, 1.0) is a correct direction for trial and error processes with increasing efficiency on both stable and unstable cell predictions. Unavoidably, random combination leads model toward the direction from (1.0, 0.0) to (0.0, 1.0), of which the success in predicting stable cells is sacrificed to promote the efficiency of unstable cell prediction.

For demonstration, we present two diagonal transects of method-derived values in Fig. 2. Meanwhile, we extract four maps from our simulation as examples to illustrate differences among methods. The coordination (in contour plot) of the four maps can be seen in Fig. 1d. Figures 2a and b show spatial patterns of simulated cases of A1 (0.3, 0.3) and A2 (0.7, 0.7), respectively; and Fig. 2c illustrates changes of method-derived efficiencies along transect from (0.0, 0.0) to (1.0, 1.0). The number of predicted unstable cell decreases from 262 in A1 case (Fig. 2a) to 136 in A2 case (Fig. 2b) whereas all three method-derived efficiencies increase. Both SR- and MSR-derived efficiencies increase smoothly along this direction (Fig. 2c) indicating their consistent sensitivity toward (1.0, 1.0). Yet, the kappa-derived efficiency shows a non-linear response. Kappa value increases rapidly after set (0.7, 0.7) is reached. Apparently, parameter set for A2 simulation is better than that of A1 simulation since better efficiency is achieved with less over-predicting. This is a correct direction conceptually for model calibration.

For comparison, Figs. 2d and e show spatial patterns for B2 (0.3, 0.7) and B1 (0.7, 0.3), respectively; and (f) illustrates the changing efficiency toward the wrong direction from B1 to B2. The two maps indicate that the number of predicted unstable cells increases from 119 in B1 case (Fig. 2e) to 279 in B2 case (Fig. 2d) with the SR-derived efficiency increases from 0.34 to 0.66 (Fig. 2f). In fact, this direction is regarded as random errors, so the kappa values are close to zero when the actual agreement is equivalent to the expected agreement. On the other hand, MSR-derived efficiency remains unchanged (around 0.5). Only the SR-derived efficiency keeps increasing along this direction (Fig. 2f). In such conditions, using SR to measure model efficiency might derive wrong parameters.

Too many over-predicted unstable cells in stable area imply the model does not ad-

equately grasp the landslide mechanism in the specific environmental setting. In fact, Casadei et al. (2003) has emphasized that one should avoid over-predicting unstable cells in using the slope stability model. Above notions support us to include stable cell prediction error into calculation. Moreover, previous studies (Carrara et al., 1995; Borga et al., 1998) have suggested that two errors should be included in assessing the accuracy of slope stability model: 1) landslide does not occur at a predicted unstable site (Type 3); and 2) slope failure takes place at predicted stable sites according to complex triggering processes (Type 2). Unfortunately, our results indicate the kappa, which holds complete logics in calculation, is too stern for landslide modeling. Among the three estimators, the MSR is optimal. It shows linear response to both stable and unstable cell errors, therefore, can produce reliable model efficiency guiding the model toward success at both axes.

6 Differences in measuring model efficiency by using SR and MSR: a real case

For purpose of comparison, the SR and the MSR were used separately to measure the efficiency of the same model. Here, we applied the SHALSTAB model onto the Chi-Jia-Wan watershed in Taiwan, where we have good landslide database and sufficient geology and vegetation information. We re-coded the SHALSTAB model developed by Dietrich et al. (2001) in C^{++} language. Governing equation in the SHALSTAB was published by Montgomery and Dietrich (1994):

$$\mathrm{FS} = \frac{\mathrm{C} + (1 - \frac{R}{T}\frac{a}{\sin\theta} \times \frac{\rho_w}{\rho_s})g\cos^2\theta \times \tan\phi}{\rho_s g \mathrm{Z}\sin\theta \times \cos\theta} \quad (4)$$

where FS is the factor of safety. Landslide occurs when FS<1. The term, ρ_w/ρ_s, is the bulk density ratio of water to soil, Z is the soil depth, θ is the slope gradient, a represents the specific contributing area (L), and $\tan\phi$ is the internal friction angle of the slope material. C is the effective cohesion (kpa), a combination of soil and root

cohesion. R is the rainfall intensity (L/T), and T is the soil transmissivity (L^2/T). The term, $\frac{R}{T} \times \frac{a}{\sin\theta}$, is the soil wetness related to pore water pressure.

We set the bulk density ratio (water to soil ratio) to be 0.4 and the soil depth (Z) to be 1.5 m according to field observations by Cheng (2003). Variables a and θ are determined according to DEMs. Thereby, only three process-related parameters (C, R/T, and ϕ) remained unknown. The GIS database and satellite image are used to construct spatial pattern of C, which is referred from the NDVI values from SPOT imagery (Rompaey et al., 2005). Randomly selected R and T values were used for the hydrological term (R/T ratio) for the entire area. The internal friction angles (ϕ) range from 30 to 45 degree according to the GIS and geological datasets. Details and ranges of variables are listed in Table 2. We focused on differences in model efficiency measured by the SR and MSR, therefore, details of parameter optimization will not be discussed in the following.

By using random number generator in assumption of uniform probability distribution (Table 2), we generated 4000 landslide susceptibility maps based on randomly selected parameter combinations. In all simulations, unstable cell is defined as FS<1. All landslide susceptibility maps are validated by actual landslide map taken from the database in the Industrial Technology Research Institute in Taiwan (Industrial Technology Research Institute, 1998). The SR and MSR are applied separately to measure the efficiency of individual simulation.

The SR-derived efficiency shows a wide range from <0.1 to as high as 0.9 (Fig. 3). By contrast, MSR-derived efficiency shows narrower range from ~0.5 to 0.75 peaking around 0.65–0.75. In Fig. 3, a dome shape scattering is observed with higher MSR-derived values at middle SR values. The MSR-derived efficiency increases as the increasing of SR-derived efficiency. This positive-correlation trend indicates a progressive prediction in unstable cells. However, the MSR-derived efficiencies start to decrease as the SR-derived efficiency keeps going higher than 0.65. Those cases with SR higher than 0.65 (vertical dashed line in Fig. 3), apparently, are resulted from over-prediction that has been sensed by the MSR method. Thus the best simulation

indicated by SR (marked by I) is significantly different from that of MSR (marked by II).

Here we present three susceptibility maps (marked by (I), (II), (III) in Fig. 3) in Fig. 4 for comparison. Related information for the three susceptibility maps is listed in Table 3. In Fig. 4a, the case (I), SR-derived efficiency is as high as 0.92 (see Table 3). This simulation hits 120 over 131 actual landslide sites. However, the coverage of predicted unstable area is as high as 75% of the total watershed area, which indicates a significant over-prediction. MSR gives this simulation an efficiency of 0.59 only. In Figure 4b, the case (II), 83 over 131 actual landslide sites are successfully predicted with predicted unstable area covering only 16% of the total watershed area (see Table 3). In this case, MSR-derived efficiency (0.75) is higher than the SR-derived efficiency (0.64). While in the case (III) (Fig. 4c), the same amount of landslide sites (83 over 131) is predicted when compared to the case (II). In this simulation unstable area covers 31% of the total watershed, which is much higher than that of the case (II). The MSR gives case (III) an efficiency of 0.67, which is lower than that given in case (II) with less over-prediction. However, old SR method gives the same efficiency (0.64) for both cases again indicating that SR is not aware of over-prediction.

This real case demonstrates that MSR can efficiently avoid over-prediction and guides the model calibration toward correct direction. Theoretically, model success should be characterized as the most number of actual landslides predicted with the least amount area predicted to be unstable (Casadei et al., 2003). Stand on this point; we suggest using the MSR to measure the landslide model performance. In addition, our example reveals that an improper estimator may retrieve questioned model results with biased parameters.

7 Conclusions

Performance measurements guiding the model calibration toward a correct direction is crucial in landslide modeling. Kappa, which gives very low model efficiencies due to its stern rules, is considered to be improper for landslide modeling. The old SR method

is insensitive to stable cell prediction; therefore, high performance can be achieved by increasing the failure in stable cell prediction without awareness. Consequently, researchers may retrieve improper parameter combination and generate questioned results when the old SR is applied. The case study in Chi-Jia Wan watershed demonstrates that MSR is able to overcome over-prediction and to provide much reliable and sensitive measure in model efficiency in comparison to the old SR. Therefore, MSR is an optimal indicator and may serve as likelihood measure in landslide modeling to confine the model calibration in a correct direction, subsequently, to retrieve the best simulation in abundant model outputs.

Acknowledgements. We acknowledge comments from M. L. Hsu at Natl. Taiwan Univ., Taiwan. The journal reviewers are gratefully acknowledged.

References

Ayalew, L. and Yamagishi, H.: The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, Geomorphology, 65, 15–31, 2005.

Borga, M., Fontana, G. D., and Marchi, L.: Shallow landslide hazard assessment using a physically based model and digital elevation data, Environ. Geol., 32, 81–88, 1998.

Carrara, A., Cardinali, M., Guzzetti, F., and Reichenbach, P.: GIS technology in mapping landslide hazard, in: Geographical Information Systems in Assessing Natural Hazards, edited by: Carrara, A. and Guzetti, F., Netherlands, 135–175, 1995.

Cheng, Y. L.: Study on Risk Analysis of Slopes Consideration Spatial Variability – A Case Study at the Lisan, Master Thesis, National Chunghsing University, 2003.

Casadei, M., Dietrich, W. E., and Miller, N. L.: Testing a model for predicting the timing and location of shallow landslide initiation in soil-mantled landscapes, Earth Surf. Processes Landf., 28, 925–950, 2003.

Dietrich, W. E., Reiss, R., Hsu, M. L., and Montgomery, D. R.: A process-based model for colluvial soil depth and shallow landsliding using digital elevation data, Hydrol. Processes, 9, 383–400, 1995.

Duan, J. and Grant, G. E.: Shallow landslide delineation for steep forest watersheds based on topographic attributes and probability analysis, in: Terrain Analysis – Principles and Applications, edited by: Wilson, J. P., and Gallant, J. C., John Wiley & Sons, New York, 311–329, 2000.

Industrial Technology Research Institute: The management of Dai-Jia Reservoir Watershed-the 4th technique report, The management committee of Dai-Jia Reservoir (in Chinese), 1998.

Lee, S.: Application and cross-validation of spatial logistic multiple regression for landslide susceptibility analysis, Geosciences, 9(1), 63–71, 2005.

Montgomery, D. R. and Dietrich, W. E.: A physically based model for the topographic control on the shallow landsliding, Water Resour. Res., 30, 1153–1171, 1994.

Pack, R. T., Tarboton, D. G., and Goodwin, C. N.: Assessing terrain stability in a GIS using SINMAP, in 15th annual GIS conference, Vancouver, British Columbia, 19–22 February 2001.

Rompaey, A. V., Bazzoffi, P., Jones, R. A. A., and Montanarella, L.: Modeling sediment yields in Italian catchments, Geomorphology, 65, 157–169, 2005.

Schuster, R. L. and Krizek, R. J.: Landslide: analysis and control, Transportation Research Board Special Report, 176, 1–10, 1978.

Sidle, R. C., Pearce, A. J., and O'Loughlin, C. L.: Hillslope stability and landuse, Washington, DC, American Geophysical Union, Water Resour. Monogr., 11, 140 pp., 1985.

Viera, A. J. and Garrett, J. M.: Understanding interobserver agreement: the Kappa statistics, Fam. Med., 37(5), 360–363, 2005.

Zhou, Q. and Liu, X.: Error assessment of grid-based flow routing algorithms used in hydrological models, Int. J. Geogr. Inf. Sci., 16(8), 819–842, 2002.

Zhou, G., Esaki, T., Mitani, Y., Xie, M., and Mori, J.: Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach, Eng. Geol., 68, 373–386, 2003.

Table 1. Basic information of artificial landslide maps in 20*20 matrix.

Landslide coverage	~5.0%	~10.0%	~15.0%
Pattern 1	16/23	21/41	25/62
Pattern 2	15/21	32/42	28/58
Pattern 3	15/17	29/38	34/59

* Denominator and numerator represent, respectively, numbers of landslide and total landslide cells. The rank of landslide aggregation: Pattern 3>Pattern2>Pattern 1.

Table 2. Model parameters, their ranges and assumption of probability distribution.

Parameter	Definitions	Range	Distribution
$C\,(x, y)$	The effective cohesion (in kpa). $C_{(x,y)} = C_{\min} + \frac{(C_{\max} - C_{\min}) \times \mathrm{NDVI}_{(x,y)}}{2}$	$C_{\min}$: 0.0~20.0 $C_{\max}$: 0.0~30.0	Uniform
R/T	A compound parameter of rainfall intensity and transmissivity. R (mm/hr), and T (m^2/hr)	R: 1.0 ~ 20.0 T: 0.001~10.0 R/T: 10^{-6}~10^{1}	Uniform
ϕ	The internal friction angle (in degree).	30~45	Uniform

NDVI $_{(x,y)}$ is calculated from SPOT imagery (Rompaey et al., 2005).

Table 3. Method-derived efficiencies and related information for three examples.

	SR-derived	MSR-derived	Predicted unstable area (%)[a]	Landslide hit[b]
(I)	0.92	0.59	75	120/131
(II)	0.64	0.75	16	83/131
(III)	0.64	0.67	31	83/131

The ratio of predicted unstable area to total watershed area (in %).
Number of predicted landslide hits over the actual landslide sites.

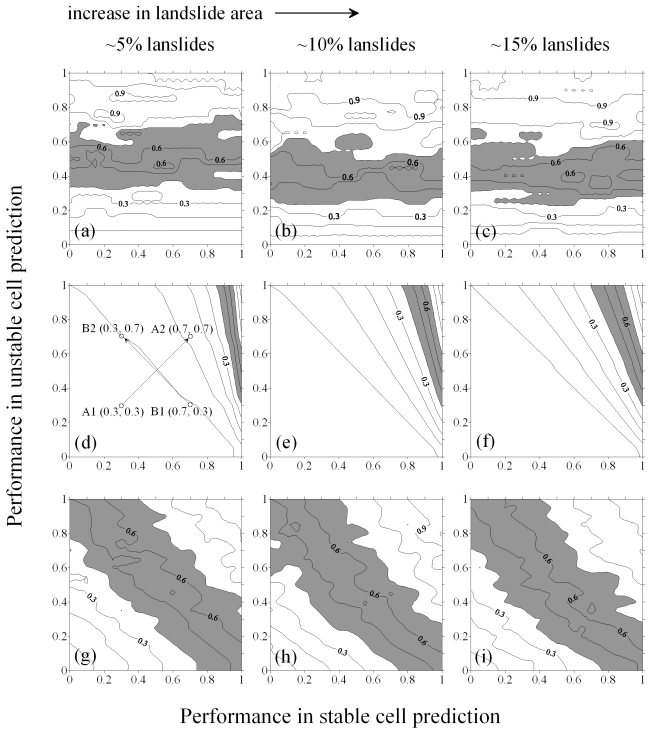

Fig. 1. The contour plots of model efficiencies derived from **(a)** SR, **(b)** kappa and **(c)** MSR methods. Landslide coverage increases rightward from 5% to 15%. Bands of moderate efficiency, 0.4–0.7, are shadowed for comparison. In **(d)**, arrow from A1 to A2 represents a correct direction (conceptually) for model calibration (see text). Arrow from B1 to B2 indicates a wrong direction.

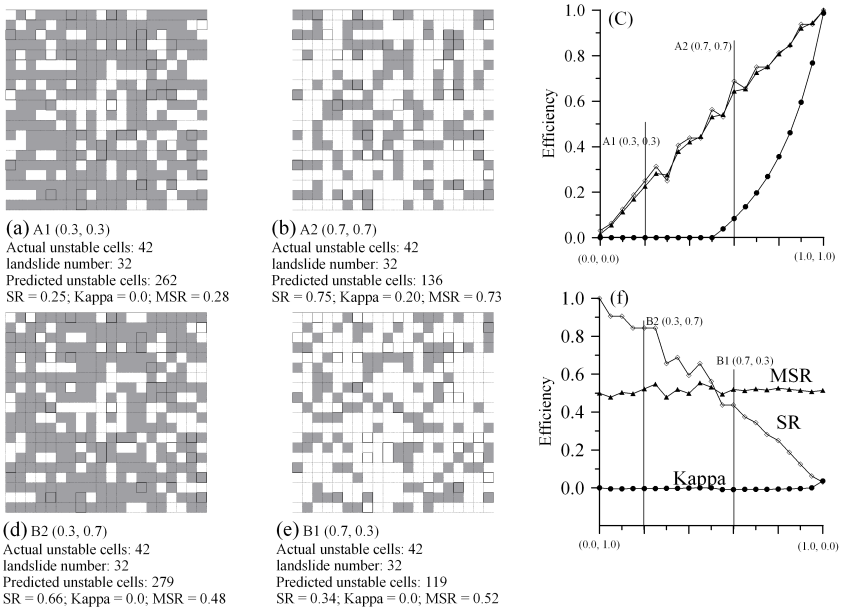

Fig. 2. Examples of actual landslide and overlaid susceptibility map (**a, b, d** and **e**) and method-derived efficiency (**c** and **f**) along two diagonal transects (see text). Coordination of the four examples is shown in contour plot in Fig. 1d. Relevant parameter and results are shown below each panel. Gray cells in maps represent predicted unstable cells. Actual landslide cells are in solid-line. Legends in (c) are the same as in (f).

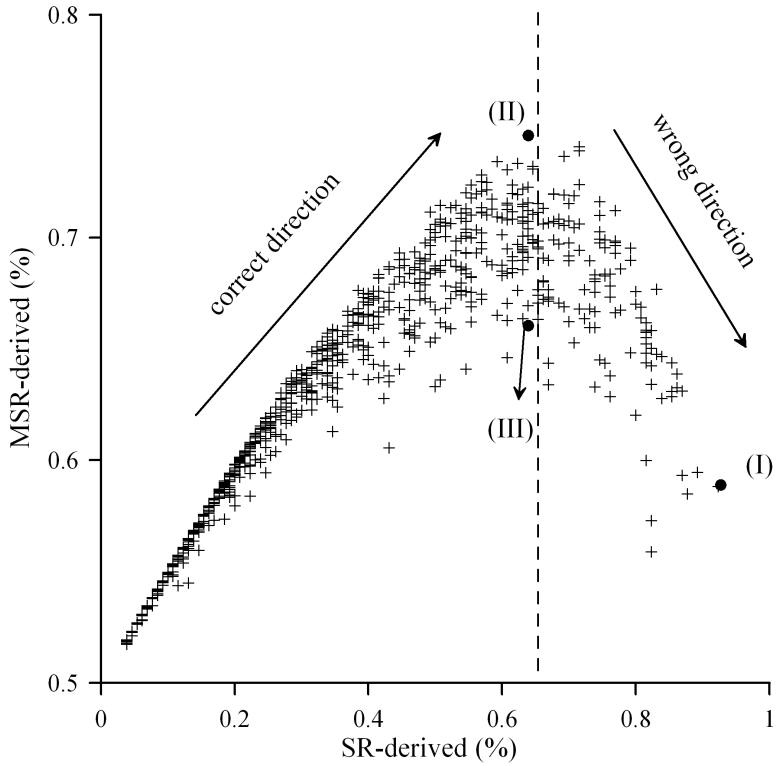

Fig. 3. Scatter plot of MSR-derived against SR-derived model efficiencies for 4000 model simulations. Landslide susceptibility maps of the three examples marked by (I), (II) and (III) are further illustrated in Fig. 4. The vertical dashed line of 0.65 of SR-derived efficiency is shown for reference.

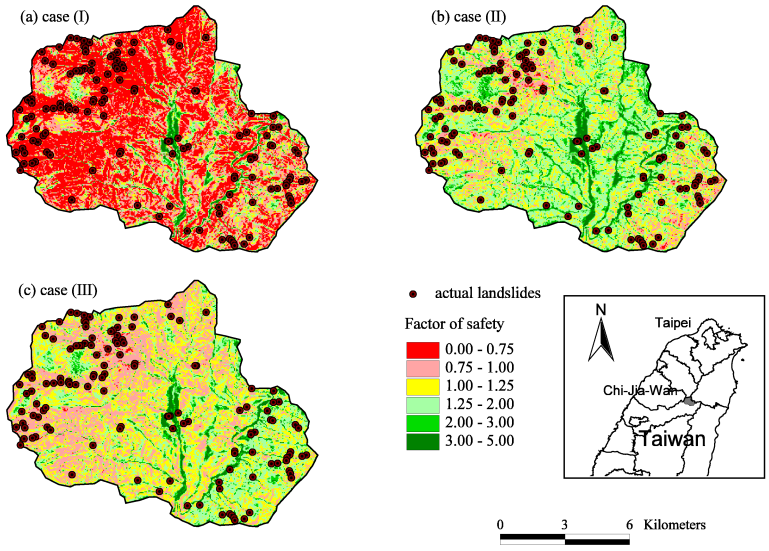

Fig. 4. Landslide susceptibility maps for the three simulation cases marked in Fig. 3. Location map of the watershed is shown at the lower right. Dots stand for actual landslide sites. Color scale for FS values is shown. The fraction of area coverage of FS<1, SR-derived efficiency, MSR-derived efficiency and landslide hits are listed in Table 3 for comparison.