Effect of Support Size on the Accuracy of Spatial Models: Findings of Rockfall Simulations on Forested Slopes

Luuk K.A. Dorren[1], Gerard B.M. Heuvelink[2] and Frédéric Berger[1]
[1] Cemagref Grenoble
2, rue de la Papeterie, B.P. 76, 38402, Saint Martin d'Hères, France
Tel: +33 4 76762806; Fax: +33 4 76513803
E-mail: luuk.dorren@cemagref.fr
[2] Laboratory of Soil Science and Geology
Wageningen University
P.O. Box 37, 6700 AA Wageningen, The Netherlands

Keywords: support; accuracy; rockfall simulation, spatial model

Abstract
The accuracy of model output increases with a decreasing support size of the input data, due to the increase of detail. This paper examines whether this is true for various spatial models developed for simulating rockfall. We analyze the effect of the support size on the accuracy of a set of models and their parameters. Both calibration and validation data were obtained from real-size rockfall experiments in France, where high-speed video cameras recorded the trajectories and velocities of more than 200 individual falling rocks with diameters between 0.8 and 1.5 meter. The second validation set was obtained in the Austrian Alps. Here we mapped rockfall impacts on trees to obtain the spatial distribution of rockfall impacts throughout the study site. These observed data are thoroughly compared with the output of a rockfall simulation model. One of the main findings is that a larger support size can be a more important cause of a larger model error than poor data quality.

1. Introduction
To sustain and protect today's livelihoods in the European Alps, protection forests are indispensable. Such forests cover the steep slopes of the main valleys and protect these developed and densely populated areas against snow avalanches, rockfall, debris flows and indirectly against flooding. Without these forests, the costs of building and maintaining technical protective constructions would be unaffordable. For these forests to provide optimal protection, adequate forest management is required (Motta and Haudemand 2000; Brang 2001. Spatial models that simulate natural hazards and integrate the protective function of forest stands can help to improve management decisions. The usefulness of such models however depends on the degree of model simplification. This in turn, depends on the availability and quality of input data, which are often determined by the feasibility of measuring the inputs with sufficient detail. Clearly, it is more difficult to obtain detailed terrain data for large catchments than for a small monitoring plot. Therefore, models and their input and output data are usually less detailed when moving up from a smaller to a larger spatial scale (Heuvelink 1998). Uncertainties or errors of model outcomes are thus related to spatial scale, of which one factor is the 'support' of the model input data. Here, support is defined as the largest area treated as homogenous such that an average value of the property of interest is known but not the variation within (Bierkens et al. 2000). If the model uses data on a large support, errors in the model outcomes could increase due to the loss of terrain information. The input data for a model should be of sufficient detail to capture the spatial variation that is essential to describe the process or pattern being modeled (Goodchild 2001). In some cases, however, the same model can be used for both a small and a large spatial scale, only the support of the input data may change.

In this paper we focus on a distributed model that simulates rockfall in forests at the slope scale, which was designed to use input data with a support of 2.5m × 2.5m. Our main objective was to analyze whether it is realistic and feasible to use this model for a larger scale (e.g. regional scale, for a study area covering 500 km2), using input data on a support of 25m × 25m. The larger support data is of poorer quality because it has less detail and involves a reduced sampling or mapping effort. We anticipated that the model using data with a support of 25m × 25m would produce a larger error than that used on a support of 2.5m × 2.5m. Our hypothesis was that poor data quality is the main cause of a larger model prediction error rather than the effect of simulating a similar process on a larger support. This hypothesis will be tested using a spatial rockfall simulation model and validation data coming from two typical forested slopes in the European Alps.

2. Rockfall modeling and validation

The rockfall model used in this study has been developed by Dorren et al. (2004). It is a spatial process based model developed for predicting 1) runout zones of rockfall events on both forested and non-forested slopes, 2) trajectories, velocities and energies of falling rocks and 3) impacts against tree stems and the accompanying energy loss. Model input data are raster based and include a Digital Elevation Model (DEM), a raster containing values for the tangential and one for the normal coefficient of restitution and rasters for the tree distribution (position and diameters). The DEM was used to determine the mean slope gradient and the fall direction and therefore it determines the acceleration and deceleration as well as the trajectory of a falling rock. The coefficients of restitution determined the amount of energy lost during a rebound on the slope surface. The tangential coefficient of restitution (r_t) determines energy loss parallel to the slope surface (due to surface roughness or vegetation) and the normal coefficient of restitution (r_n) determines energy loss perpendicular to the slope surface (due to elasticity of the material covering the slope surface). In the literature, a wide range of values exists for these parameters for many different types of surface cover material (Pfeiffer and Bowen 1989; Kobayashi et al. 1990; Giani 1992; Azzoni et al. 1995; Chau et al. 1998, 2002; Dorren and Seijmonsbergen, 2003).

The study of Dorren et al. (2004), which was carried out in the Austrian Alps, showed that their model can quite accurately predict rockfall runout zones, as spatial patterns of rockfall accumulation zones mapped in the terrain were reproduced by the model with an R^2 of 0.74. Data on observed velocities and energies were not available for the site in Austria. Currently, however, these have been obtained from real-size rockfall experiments on a forested slope in the French Alps in the framework of the ROCKFOR project (ROCKFOR 2004). These experiments are visualized in figure 1. In short, we threw large, individual rocks (sphere type rocks with a mean diameter of 0.5 meter) down a forested slope with a mean gradient of 38 degrees. By using field measurements (laser vertex) and video cameras we captured the velocity and the trajectory of the rock in 3D. From these experiments we obtained the data presented in table 1. This table also presents the available simulation results obtained with our rockfall simulation model using data on a support of 2.5m × 2.5m, except for the forest input data (tree positions and diameters) which were on a support of 0.5m × 0.5m. The mean error (ME) and the mean squared error (MSE) of the output are calculated following:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2 \tag{2}$$

Where n is the number of observations, P_i is the simulated or predicted value and O_i is the observed value, which is obtained from the experimental data. The ME and the MSE are calculated on the basis of errors of different observations that are presented in table 1. The variation in the results of all the individual rockfall experiments is very large, which underlines that rockfall on forested slopes is a stochastic phenomenon. Moreover, we were able to carry out 102 individual rockfall experiments on a forested slope during the last two years. The model uses Monte Carlo simulations (Lewis and Orav 1989; Mowrer 1997) and consequently produces much more virtual rockfall trajectories than we ever could have observed during experiments. As a result, we cannot compare single observed trajectories with simulated ones and are thus forced to work with average values.

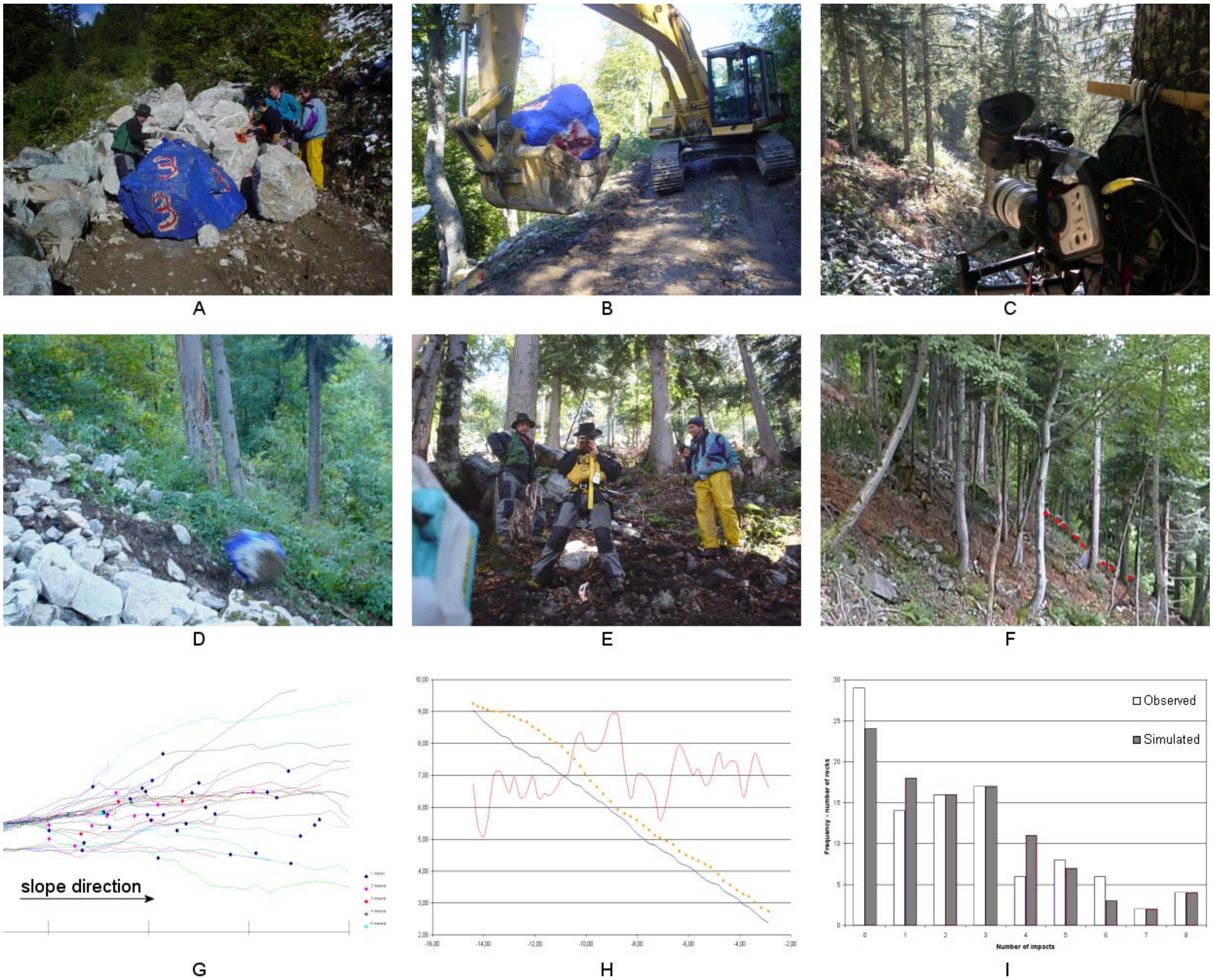

Figure 1. A) rocks are painted to leave marks after tree impacts and rebounds; B) a caterpillar throws the rocks down the slope; C) video cameras are installed along the slope to record the trajectory in 2D; D) a rock accelerating in the first meters; E) after each individual rockfall experiment, we capture its trajectory to obtain a 3D trajectory; F) Example of the camera view; G) forty observed rockfall trajectories; H) example of video image analysis providing the rockfall velocity and a 2D trajectory; I) observed (white) and simulated (gray) histograms of the number of tree impacts per rock.

The results presented in table 1 and figure 1-I show that our rockfall model produces acceptable results (ME = 0.01% and MSE = 171.93%). The larger errors are produced by the maximum velocity and by the residual risk of the forested area. The residual risk indicates the percentage of rocks that cannot be stopped by the forest. On the test site, this forest covered the upper 150 meters of the experimental slope. The model was capable to accurately simulate the percentage of rocks that stop on the experimental slope as a whole (observed 79% and simulated 79%). In addition, the model reproduced very well the average number of tree impacts as well as the spatial pattern of the observed rockfall trajectories shown in figure 1-G. Having obtained confidence in the developed model, our next step was to test it using input data on a support of 25m × 25m with different data qualities at a different site in the European Alps. Initial simulation with data on a support of 25m × 25m from the French site indicated that the same model can be used. Moreover, the runout zone was quite accurately reproduced, but simulated tree impacts and rockfall velocities differ significantly from our observations. Consequently, we would like to know which parameter is responsible for the larger errors and has to be improved in order to carry out rockfall assessment at a regional scale.

Table 1. Observed and simulated characteristics of rockfall on forested slopes.

Variable	Explanation	Observed	Simulated	Error (%)
Mean velocity	Average translation velocity in the forest [m/s]	8.2	8.8	7.3%
Max. velocity	Maximum translation velocity in the forest [m/s]	23.9	28.7	20.1%
Runout zone	Percentage of rocks stopping within 175 m [%]	79%	79%	0
Residual risk of forest	Percentage of rocks passing the forested zone [%]	34%	25%	-26.5%
Tree impacts	Mean number of tree impacts per falling rock [-]	2.3	2.4	4.3%
Mean impact height	Mean height of impacts on trees [m]	0.77	0.73	-5.2%
			Mean error:	0.01%
			Mean squared error:	171.93%

3. Model simulations with different support sizes

Both the input data and the validation data for the following model tests were obtained from a forested, active rockfall slope in the most western part of the Austrian Alps, located at 47°00´ latitude and 10°01´ longitude. This test slope could be divided in two areas. The rockfall source area, which is a steep cliff face dissected by large denudation niches and an accumulation area, a large post-glacially developed talus cone consisting mainly of rockfall scree, but also some debris flow material. The mean slope gradient in the source area is approx. 70 degrees and in the accumulation area 38 degrees. The slope length of the talus cone is 900 meters. An overview of the site is shown in Figure 2.

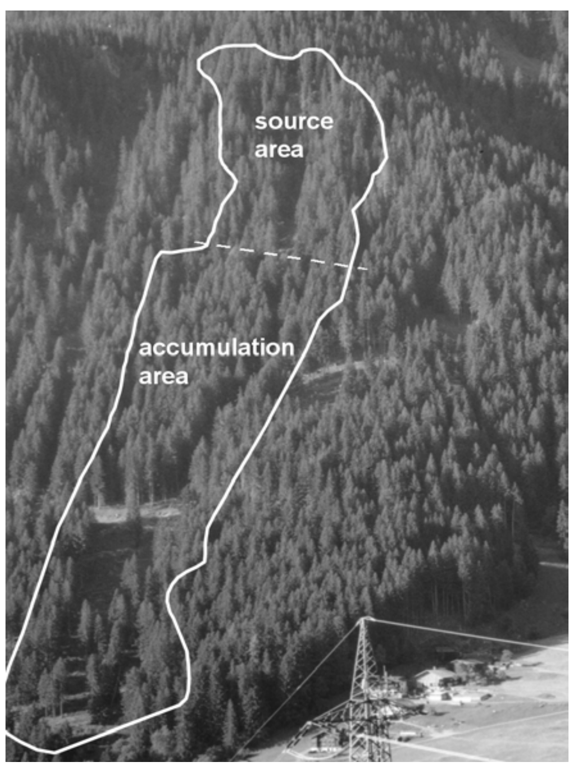

Figure 2. Photograph of the study site; the white outline represents the site used for simulation modeling.

Table 2. Input data for the three simulation schemes, and their origin.

Simulation scheme	Support	DEM	Tree distribution	r_t	r_n
1	2.5m × 2.5m HR	Contour lines	Orthophotos + detailed inventory	Detailed field map + literature	Detailed field map + literature
2	25m × 25m HR*	Aggregated from 1	Aggregated from 1	Aggregated from 1	Aggregated from 1
3	25m × 25m LR	Photogrammetry	Landsat TM + regional inventory	Landsat TM + literature	Landsat TM + literature

HR = High resolution, large support
HR* = HR data aggregated to 25m × 25m support
LR = Low resolution, small support

A set of different simulation schemes was defined to analyze the effect of input data with different aggregation levels on the accuracy of the output. In simulation scheme 1, input data with a support of 2.5m × 2.5m (high resolution – HR) were used. The output was aggregated to a support of 25m × 25m by averaging the values of the output on the 2.5m × 2.5m support. The output data were compared with validation data at the same support. These validation data were extracted from detailed forest inventory data. On the upper part of the accumulation area of our test site, 18 squares of 25m × 25m were selected randomly. For these 18 squares the number of rock impacts scars per tree volume were measured. These data were compared with the number of rock impacts on trees as simulated by the model. For standardization purposes, both the validation data (observed values) and the simulated data (predicted values) in the 18 squares were expressed as percentages of the summed values for all the randomly selected squares. In simulation scheme 2 the same input data were used as before, but they were aggregated to a support of 25m × 25m prior to running the rockfall model (high resolution data aggregated – HR*). In simulation scheme 3 the support of data was the same as in scheme 2, but the input data for simulation scheme 3 were obtained directly at a support of 25m × 25m (low resolution data – LR). Since the data used in simulation scheme 3 were obtained at the regional scale, and thus less detailed than the data used in simulation scheme 1 and 2, we can state that the quality of this data was considerably poorer. An overview of all the used input data is given in table 2. Further details on the data acquirement as well as on the used methods are described by Dorren and Heuvelink (2004). To test the effect of aggregation of individual parameters, we applied additional simulation schemes in which HR and LR data were mixed. These schemes and their produced errors are explained in table 3.

4. Results of the simulation schemes

Comparison of the simulated impacts with the number of observed scars per unit tree volume provided the results shown in Table 3. Of the three initial simulation schemes, number 2 produced the largest errors and simulation scheme 1 the smallest errors. The latter is caused by a more accurate estimation of the larger observed values, which is shown by figure 3. Nevertheless, simulation scheme 1 also produced some considerable mismatches of the smaller observed values. The latter affects the mean squared error considerably for simulation scheme 1 (MSE1), as shown in Table 4 (MSE1 = 24.9).

Table 3. The MSE of 'intermediate' simulation schemes 4 to 7 and the initial schemes 1, 2 and 3.

Simulation scheme	Used data	ME	MSE
1	HRtree, HRr_n, HRr_t, HRDEM	0.0	24.9
2	HRtree*, HRr_n*, HRr_t*, HRDEM*	0.0	44.4
3	LRtree, LRr_n, LRr_t, LRDEM	0.0	31.3
4	LRtree, LRr_n, LRr_t, HRDEM*	0.0	47.8
5	LRtree, LRr_n, HRr_t*, LRDEM	0.0	35.7
6	LRtree, HRr_n*, LRr_t, LRDEM	0.0	31.4
7	HRtree*, LRr_n, LRr_t, LRDEM	0.0	29.3

* HR data aggregated to 25m × 25m support

Simulation scheme 2 produced the largest error (MSE3 = 44.4) and simulation scheme 3 produced an intermediate one (MSE2 = 31.3). The results presented in Table 3 show that the substitution of LRtree by the aggregated HRtree (simulation scheme 7) resulted in an MSE of 29.3, which is smaller than the initial MSE for scheme 3 (see also Figure 7a). Substitution of LRr_t by the aggregated HRr_t (simulation scheme 5) increased the MSE from 31.3 to 35.7. Table 2 also shows that the substitution of LRr_n by the aggregated HRr_n (simulation scheme 6) resulted in an increase of the MSE of 0.1, which indicates that the net effect of r_n on the simulation results was small. A remarkable result is that the substitution of the LRDEM by the aggregated HRDEM (simulation scheme 4) did not decrease MSE3. On the contrary, it resulted in a large increase of the MSE from 31.3 to 47.8. The scatter plot of this simulation result is shown in Figure 7b. This simulation scheme strongly overestimated the smaller observed values and strongly underestimated the larger observed values.

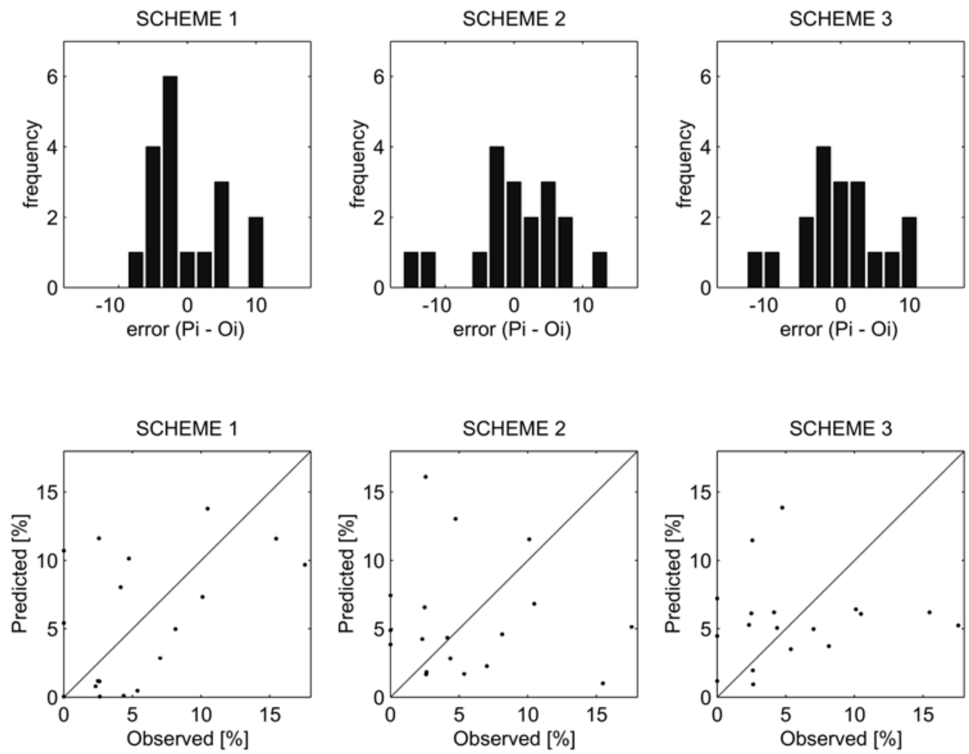

Figure 3. Histograms of the errors produced by simulation schemes 1, 2 and 3 and accompanying scatter plots with observed values versus predicted values.

5. Discussion

We did not expect simulation scheme 3 to give a smaller MSE than scheme 2. Rather, we anticipated that scheme 3 would perform the worst, because it uses input data of the poorest quality. The only difference between simulation schemes 2 and 3 is the values for four model parameters, which have all been changed simultaneously. The 'intermediate' simulation scheme 4 indicated that the aggregated HRDEM was mainly responsible for the increase in error. In the test site, the rockfall trajectories are for a large part determined by a preferred transport channels or gullies. These are represented in both the HRDEM and HRtree, but they are averaged-out to a certain extent in the aggregated HRDEM but still represented. As a consequence, the fall directions calculated on the basis of the aggregated HRDEM were generally towards this channel, which led to a concentration of falling rocks. Therefore, the number of impacts is higher in that part of the study area. This effect is reinforced by the fact that the transport channel is almost free of trees. Therefore, hardly any rock impacts against trees occur in the channel in reality. However, when using the aggregated HRtree, the forest structure in the channel as observed in the field and in HRtree is completely lost. Consequently, the number of trees in the fall track of the simulated rocks, as represented by the aggregated HRtree, is overestimated, although the number of trees in the channel is still smaller than in the surrounding areas. As a result, the number of impacts in the channel is more strongly overestimated than in the other simulation schemes. This error occurred to a lesser extent in simulation scheme 3, since in the LRDEM the channel was completely 'smoothed-out'. As a result a more uniform distribution of rock impacts was produced.

Overall, the results indicate that the distributed rockfall model used in this study, which was developed for rockfall assessment at a slope scale, can be used for rockfall assessment at a regional scale. As expected, the analyzed simulation schemes indicated that input data with a support of 25m × 25m increased the MSE compared to input data with a support of 2.5m × 2.5m. However, the simulated maximum extents of rockfall runout zones were similar for simulation schemes 1, 2 and 3. In addition, these simulated maxima also corresponded with those observed in reality, which shows that modeling rockfall runout zones at the regional scale is feasible and realistic, even for forested catchments. The simulated rockfall impacts on tree stems using data with a support of 25m × 25m were not accurate as the mean squared errors produced by simulation schemes 2 and 3 were much larger than the MSE of scheme 1. Using tree distribution data of higher quality could reduce the MSE of scheme 3 with about 2% as shown by the 'intermediate' simulation scheme 7. The accuracies of the results produced by scheme 2 and 3 indicate that simulating damage on tree stems caused by rockfall using data with a support of 25m × 25m is not realistic. To assess where and how much tree stem damage will occur, high quality forest data with a small support is required. It is however possible to simulate rockfall runout zones with a support of 25m × 25m.

6. Conclusions

This paper investigated the relationship between the aggregation level of the input data and the accuracy of the output of a rockfall simulation model. The results showed that the simulation of rockfall with a distributed model using data with a support of 25m × 25m is feasible and realistic to simulate rockfall runout zones, but not for the simulation of tree damage caused by rockfall. The latter arose because collisions of rocks against tree stems cannot be simulated accurately where the data are of poor quality and the support is large. As anticipated, the model using data with a support of 25m × 25m produces a larger error than that using data with a support of 2.5m × 2.5m. Our hypothesis was that poor data quality is a more important cause of a larger model prediction error than the effect of a larger support. This study showed that this is not necessarily true because the simulation scheme that used data of higher quality produced a larger error than the simulation scheme that used data of poorer quality. Here it was interesting to observe that the loss of important spatial structure in the input data (i.e. the rockfall channel represented in the slope map and in the tree distribution map), as caused by spatial aggregation, resulted in a larger model prediction error than the use of data that represented the landscape with less detail. The results of this study also indicate that the use of a regional DEM of high quality requires data on forest structure of higher quality than does a regional DEM of poorer quality in case of simulating rocks falling through mountain forests. It would be interesting to aim future rockfall modeling research determining the minimum support required to obtain realistic and trustworthy modeling results for the assessment of the degree of protection provided by mountain forests against rockfall hazards in the European Alps.

Acknowledgements
The funding by the European Community for the ROCKFOR project (QLK5-CT-2000-01302) and the financial support provided by the Niels Stensen Stichting is gratefully acknowledged.

References

Azzoni, A., Barbera, G.L. and Zaninetti, A. 1995. Analysis and prediction of rockfalls using a mathematical model. International Journal of Rock Mechanics and Mining Science, 32, 709-724.

Bierkens, M.F.P., Finke, P.A., Willigen, P. De, 2000. Upscaling and downscaling methods for environmental research (Kluwer Academic Publishers, Dordrecht).

Brang P. 2001. Resistance and elasticity: promising concepts for the management of protection forests in the European Alps. Forest Ecology and Management 145, 107-117.

Chau, K.T., Wong, R.H.C. and Lee, C.F. 1998. Rockfall problems in Hong Kong and some new experimental results for coefficient of restitution. International Journal of Rock Mechanics and Mining Science, 35, 662-663.

Chau, K.T., Wong, R.H.C. and Wu, J.J., 2002. Coefficient of restitution and rotational motions of rockfall impacts. International Journal of Rock Mechanics and Mining Science, 39(1), 69-77.

Dorren, L.K.A. and Seijmonsbergen, A.C., 2003. Comparison of three GIS-based models for predicting rockfall runout zones at a regional scale. Geomorphology, 56(1-2), 49-64.
Dorren, L.K.A. and Heuvelink, G.B.M., 2004. Effect of support size on the accuracy of a distributed rockfall model. Int. Journal of Geographical Information Science, in press.
Dorren, L.K.A., Maier, B., Putters, U.S. and Seijmonsbergen, A.C., 2004. Combining field and modelling techniques to assess rockfall dynamics on a protection forest hillslope in the European Alps. Geomorphology, 57(3-4), 151-167.
Giani, G.P. 1992. Rock Slope Stability Analysis (Balkema, Rotterdam).
Goodchild, M.F., 2001. Models of scale and scales of modeling. In Modelling scale in Geographical Information Systems, edited by N.J. Tate and P.M. Atkinson (John Wiley & Sons, Ltd., Chichester), pp.3 – 10.
Heuvelink, G.B.M. 1998. Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems, 50, 255-264.
Kobayashi, Y., Harp, E.L. and Kagawa, T. 1990. Simulation of rockfalls triggered by earthquakes. Rock Mechanics and Rock Engineering, 23: 1-20.
Lewis, P.A.W. and Orav, E.J. 1989. Simulation methodology for statisticians, operations analysts, and engineers (Volume 1. Pacific Grove, Wadsworth & Brooks/Cole).
Motta, R., Haudemand JC. 2000. Protective forests and silvicultural stability. An example of planning in the Aosta valley. Mountain Research and Development 20, 74-81
Mowrer, H.T., 1997. Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. Ecological Modelling, 98, 73-86.
Pfeiffer, T.J. and Bowen, T.D. 1989. Computer simulation of rockfalls. Bull. Ass. Engineering Geologists, XXVI, 135-146.
ROCKFOR, 2004. Rockfall forest interrelation – EC funded project on the efficiency of the protective function of mountain forests against rockfall. [online] URL: http://rockfor.grenoble.cemagref.fr/

Citation:

Dorren, L.K.A., Berger, F. and Heuvelink, G.B.M., 2004. Effect of support size on the accuracy of spatial models: findings of rockfall simulations on forested slopes, Joint Meeting of The International Environmetrics Society and the Spatial Accuracy Symposium, 28 June - 1 July 2004, Portland, Maine, USA: 15 pp.