Observations from a Statistical Model for Maximum Avalanche Run-out Distances in Southwest Montana.

L.R. McKittrick and R.L. Brown
Civil Engineering Dept.
Montana State University
Bozeman, MT 59717

Abstract

An extreme value (Gumbel) distribution was evaluated as a model of maximum avalanche run-out distances in Southwest Montana. The model is defined in terms of dimensionless ratios that are defined relative to an arbitrarily defined reference point. The development and analysis of the statistical model was based on data from surveys of twenty-four avalanche paths. The technique bears promise as a user friendly tool; however, the model is quite sensitive to measurement uncertainty. The significance of this sensitivity depends on the definition of the reference point. Therefore, the reference point for this model was chosen to optimize the models sensitivity to field measurements. Comparison of the model developed here with models developed for other mountain ranges indicates that models developed for one mountain range cannot be used to accurately estimate maximum avalanche run-out distances in other ranges until natural sources of variation are better understood.

Introduction

The purpose of this research is to develop a more user friendly tool for estimating maximum avalanche run-out distances in Southwest Montana. An approach used by McClung, Mears, and Schaerer [3] offers a means of achieving this goal. This method is essentially a statistical evaluation of historical avalanche activity in any given mountain range. Maximum run-out distances are estimated as based on avalanche extremes that have occurred in the last "100 year" period. McClung and Mears [2] applied this approach to mountain ranges in Norway, the Sierras, the Colorado Rockies, the Canadian Rockies and Coastal Alaska. They noted substantial differences in the results between these ranges. These discrepancies may be due to natural sources of variation such as differences in terrain features, and weather conditions. The large majority of slide paths evaluated by McClung and Mears were large, over 350 meters vertical drop. In this study we tested the applicability of McClung's approach to the mountains in Southwest Montana where a large percentage of the slide paths are considerably smaller.

In a previous paper [4], we discussed the construction of a statistical model for extreme avalanche run-out distances in Southwest Montana. Here we repeat many of the details on the construction process but go into much more detail on the sensitivity of the model to errors and uncertainties.

Maximum avalanche run-out distances were measured on twenty-four avalanche paths in the Madison and Gallatin Ranges of Southwest Montana. Slope and distance measurements were taken with an inclinometer and tape measure. Maximum run-out distances were determined using records of past avalanches in terms of tree destruction and other vegetation damage. In several cases the return period of different avalanche run-out distances was clearly displayed as a series of steps in tree ages. In these cases only younger trees grew in the run-out zones of the more frequent avalanches while slightly older trees grew in the extended run-out zones of the larger but less frequent avalanches.

Since avalanche run-out distances are known to have a probabilistic nature, the data collected was used to construct a statistical model of maximum run-out distances. The model is based on the ratio of two distance parameters, which can be easily determined from field measurements.

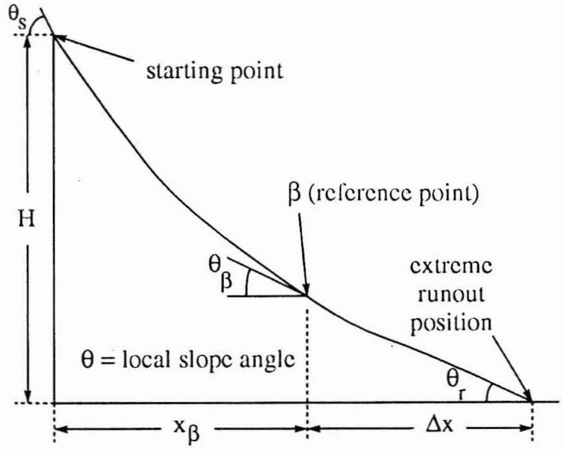

Figure 1: Parameters defined in terms of path geometry

The statistical approach taken by McClung, Mears, and Schaerer [3] is based on the definition of a dimensionless run-out ratio, which is then used as the random variable of a probability distribution. Before defining the dimensionless run-out ratio, a reference point must be selected along the avalanche path, as shown in Figure 1. McClung, Mears, and Schaerer define their reference ,β, point as the first point along the avalanche path that diminishes to a specified local slope angle, θ_β. This definition can be applied fairly consistently if the paths are continuously concave, but if the paths are interrupted by ledges or sloping benches then the reference point may change position as a function of the averaging or smoothing technique used to define local slope angles. Some of these ledges are so small that they presumably have an insignificant influence on the avalanche run-out distance so that they can be modeled by smoothing out the bump and choosing the next lower position which meets the definition of the reference point. If the slope is interrupted by a larger bench with a shallow slope and the reference point is positioned on this bench, the probabilistic model, developed using continuously concave slopes, appears to lead to an underestimation of the expected maximum run-out distance [1].

The run-out ratio is defined as the horizontal distance, Δx , between the β point and the extreme run-out position divided by the horizontal distance, x_β , between the β point and the starting position of the avalanche. In symbolic terms, we have

$$runout\ ratio\ (r) = \frac{\Delta x}{x_\beta} \qquad (1)$$

When we first began this study, we intended to simply extend or confirm the accuracy of the statistical model used by McClung, Mears, and Schaerer [3]; however, we are working with smaller avalanches paths and hence had to slightly alter their model. McClung, Mears, and Schaerer defined their β point where the local slope angle, θ, first diminished to ten degrees, but out of the twenty-four avalanche paths we surveyed, only five ran far enough to reach a local slope angle of ten degrees. In some cases, θ did not diminish to ten degrees for two or three hundred meters beyond the end of the extreme run-out position; therefore, we felt that a ten degree reference point had no physical significance relative to our paths and decided to search for a new definition for the β point.

To define the path profile, we surveyed the avalanche paths by taking slope and distance measurements every fifty meters in regions where the slope was relatively constant. We shortened our survey distance where the slope was highly variable. As mentioned previously, since local regions of the slope can display significant variability, the choice of the β point can be quite dependent on the averaging technique that is applied. To simplify the averaging process and keep it as consistent as possible, we decided to model each avalanche slope using the least squares fit second degree polynomial. This approach also allowed us to easily vary the definition of θ_β and solve for new β points on each avalanche path.

If we were to develop a probabilistic model with run-out ratios for all avalanches, we might expect a normal distribution to work quite well; however, since we are modeling only those avalanches we would expect to lie in the extreme tail of the normal distribution, we use an extreme value distribution, which, in a sense, models the right hand tail of the normal distribution. Here we are working with the extreme value (Gumbel) distribution initially used by McClung and Lied [1]. In this case, we have a probability density function of the form:

$$f(r) = \frac{1}{b} exp\left(-\frac{r-u}{b} - exp\left(-\frac{r-u}{b}\right)\right) \tag{2}$$

Since this distribution is not symmetric, the location parameter (u) and the scale parameter (b) do not represent the mean and standard deviation. Instead, the mean is of the form $u + \gamma b$, where $\gamma = 0.57721$ (Euler's constant), and the standard deviation is of the form $\pi b/\sqrt{6}$. Now if we let r_p represent the pth quantile, we have a cumulative distribution function of the form:

$$P(r \leq r_p) = \int_{-\infty}^{r_p} f(z)dz \tag{3}$$

$$= exp\left(-exp\left(-\frac{r_p - u}{b}\right)\right) \tag{4}$$

$P(r \leq r_p)$ is the non-exceedance probability divided by one hundred (i.e. $0 \leq P \leq 1$).

To obtain estimates for the parameters using our data, we let $r_p = \left(\frac{\Delta x}{x_\beta}\right)_i$, then rewrite Equation 4 as

$$\left(\frac{\Delta x}{x_\beta}\right)_i = u - bln(-ln(P_i)) \tag{5}$$

$$= u + bY_{P_i} \tag{6}$$

where $Y_{P_i} = -ln(-ln(P_i))$ is the called the reduced variate. Once we have defined appropriate values of P_i for each run-out ratio, we can solve for the parameters u and b using linear regression.

To define non-exceedance probabilities corresponding to each run-out ratio, we first arrange the N run-out ratios in increasing order so that

$$\left(\frac{\Delta x}{x_\beta}\right)_1 \leq \left(\frac{\Delta x}{x_\beta}\right)_i \leq \left(\frac{\Delta x}{x_\beta}\right)_N \tag{7}$$

For the corresponding non-exceedance probabilities, we use the empirical form obtained by

McClung and Mears [2],

$$P_i = \frac{(i - 0.4)}{N} \tag{8}$$

where $i = 1, 2, \ldots, N$.

Statistical results

Second degree polynomials worked reasonably well as models of the twenty-four avalanche paths we surveyed. All but two yielded fits with standard deviations less than six meters where the median was approximately 3.7 meters. Based on our data, we chose to define our reference point for $\theta_\beta = 18$ degrees. We did not develop a firm physical basis for this definition, but when higher values are used for θ_β, the model is much more sensitive to measurement errors. Using our data and definition of θ_β, we have the statistics shown in Table 1.

Variable	Mean	Standard Deviation	Range
H (m.)	248	123	68 - 553
θ_s (deg.)	38	4.0	31 - 46
θ_r (deg.)	14.5	6.1	1.5 - 25
Δx (m.)	66	102	-87 - 340
x_β (m.)	432	221	140 - 982

Table 1: Descriptive Statistics

Using regression on the linearized form of our extreme value distribution, we obtain the parameters and corresponding 95% confidence limits on the mean shown in Figure 2. Since we have less than thirty data points, we used the t distribution to calculate our confidence limits as discussed by Walpole and Myers [7, §8.3].

Though twenty-five avalanche paths were surveyed, one path, which had a large bench midway and was therefore inconsistent with the other paths, yielded a run-out ratio which far exceeded the run-out ratios of the other paths and was eliminated.

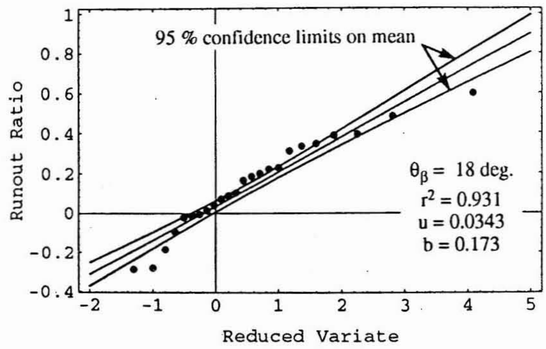

Figure 2: Regression line for extreme value distribution

Sources of Uncertainty

Since we expect a large range of weather patterns to occur over a 100 year period, we expect patterns which produce extreme avalanches to occur somewhere in that time frame. By measuring maximum avalanche run-out distances, based on vegetation damage in the last "100" years, we reduce the large variations in avalanche run-out distances due to year to year weather patterns. At the same time, without completing a detailed tree ring analysis at each path, we can not be certain of the true time period that the vegetation damage has recorded since the recovery rate for vegetation (trees) varies significantly between north and south aspects and also between higher and lower altitudes.

Of the avalanches recorded by vegetation damage, if we assume we have the maximum run-out distances for the last 100 years, then most likely we are actually working with avalanche run-out distances that have return periods greater than 100 years. For example, if we consider avalanches that have a return period of 1000 years, then for each path, based on an exponential distribution, there is roughly a 9.5 percent chance that we are working with a run-out distance that has a return period of 1000 years. Substituting into the binomial distribution, we estimate that approximately 2 of the 24 surveyed have a return period of 1000 years, as long as their vegetation records have not been erased by more extreme avalanches. Similarly, we would expect approximately 7 of the 24 surveyed to have a return period of 300 years. Though these occurrences may appear to be a problem at first, most likely they are not, since we would also expect similar occurrences in the next 100 years. Also, since these occurrences are included in our data, they are also factored into our model.

Avalanche paths have variations in flow resistance, path geometry, aspect, vertical drop, avalanche volume, etc. When we prepare to use the statistical model we must verify that the data set used to build the statistical model contained a sufficient number of avalanche paths with similar characteristics; otherwise the model may be biased and yield poor estimations of run-out distances. For example, if our model is developed primarily from confined avalanche paths, we would not feel comfortable using our model to estimate run-out distances for un-confined avalanches. Obviously to develop a more accurate model, it is necessary to work from a database of similar avalanche paths. On the other hand, to build a very general though conservative model, the data set must include a sufficient number of avalanche paths which bear all of the significant characteristics. With this observation in mind, we see that the data collectors must be clear about which sources of variation have been included in the development of the model.

Unfortunately data contains not only variation due to path characteristics, but also possibly human errors and inherently variations due to measurement uncertainties. These errors and uncertainties should first be minimized through rigorous surveying techniques, but for uncertainties that cannot be easily removed, the objective is to choose a model that minimizes the uncertainty in the estimations it provides. To do this we must obtain a feel for the effect each type of measurement uncertainty might have on our set of possible models. To determine a model's sensitivity to these uncertainties we can use a Monte Carlo simulation [5]. In essence, we superimpose normally distributed noise (errors) on our data and then record the resulting distribution of model parameters. From the distribution of model parameters, we can determine confidence limits relative to our measurement uncertainties.

As mentioned earlier, to survey the avalanche paths we first estimate the extreme limit of the starting zone or run-out zone, whichever is more convenient. We then record points along the avalanche path in terms of polar coordinates. If the slope appears to be relatively constant we take steps with distances (radii) as large as fifty meters, and where there are large variations in slope we shorten the recording distance accordingly. We measure the distance with a fifty-meter tape and slope with an inclinometer (Brunton compass). This process is continued until we reach our estimate of the opposite extreme end of the avalanche path. Most of the paths have rocky cliff bands at the top which provide fairly definite starting points, but others begin in large meadows or on convex break-over points which leave the starting points open to judgement. The extreme run-out positions were usually better defined, but there were still cases where there was a choice between following a small off-shoot or stopping at the end of what appeared to be the bulk flow. Thus based on our measuring technique, we have the following four sources of uncertainty in our measurements:

1. estimation of starting zone position
2. measurement of distance
3. measurement of slope
4. estimation of extreme run-out position

As mentioned above, to estimate the effect each of these types of errors might have on our model we have superimposed normally distributed noise (errors) on top of the data we recorded. To do so, we have assumed the standard deviations listed in Table 2 for the respective error distributions.

Source of Uncertainty	Standard Deviation
Starting Zone Position	3 m.
Distance (radii)	1 m.
Slope	1 deg.
Extreme Run-out Position	3 m.

Table 2: Assumed Error Standard Deviations

If we consider the uncertainty due to the combined effect of these assumed errors we obtain the 95% confidence limits on the mean cumulative distribution function shown in Figure 3. These confidence limits imply that if we were to repeat our measurements on the same paths there is a 95% chance our new model would lie within these limits. To obtain a feel for how this uncertainty might affect the use of our model we have also plotted the corresponding 0.95 quantile corresponding to the 95% non-exceedance probability. By using the 0.95 quantile, we are saying that we expect 95% of all extreme avalanches to stop before they reach the corresponding run-out ratio. Therefore to give ourselves an edge against measurement errors we would probably use the upper 95% confidence limit due to measurement uncertainties and if we ignored uncertainty due to other factors, we could then state that we expect 95% of all extreme avalanches in the next 100 years to have a run-out ratio less than 0.571.

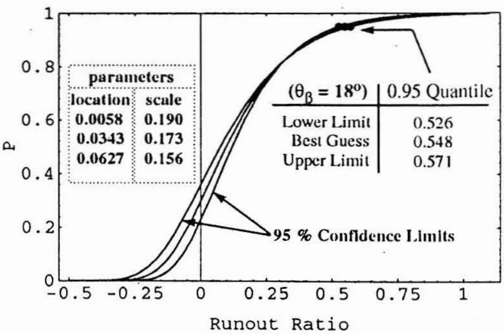

Figure 3: Confidence Limits for $\theta_\beta = 18°$ due to measurement uncertainties

Since we did not use McClung and Lied's definition for θ_β, we need to clarify why we chose to define $\theta_\beta = 18°$. We had hoped to establish a definition with a physical basis; however, we discovered that the model becomes much more sensitive to measurement errors as we increase the angle θ_β. Thus we decided to settle on a definition which, in a sense, optimizes the models sensitivity to these errors.

We begin by reminding ourselves exactly what we are trying to model. We would like to model extreme run-out distances; therefore our model should be sensitive to changes in run-out distances. Once we fix a reference point, we would like to determine the extreme run-out point relative to our reference point. In our geometrical model this distance is represented by the parameter Δx. Now suppose we change Δx by some fixed amount, δ. To consider the sensitivity of the run-out ratio to the change, consider the following equation:

$$\frac{\frac{\Delta x+\delta}{x_\beta} - \frac{\Delta x}{x_\beta}}{\frac{\Delta x}{x_\beta}} = \frac{\delta}{\Delta x} \qquad (9)$$

As Δx approaches zero the run-out ratio becomes increasingly sensitive to changes in the extreme run-out point. If Δx becomes too small then our run-out ratio becomes exceedingly sensitive to changes and consequently, errors. Based on this analysis, θ_β should be defined to place the reference point "near" the end of the avalanche paths.

Now though we may define θ_β to optimize sensitivity to measurements of the extreme run-out position, we would also like to minimize the uncertainty in the model due to other measurements. We can apply the same technique to errors in the determination in the starting point that we used above. Let ϵ_s represent an error in the determination of the starting point. Then the following equation illustrates the fact that as x_β increases, errors relative to the estimation of the starting point become less significant.

$$\frac{\frac{\Delta x}{x_\beta} - \frac{\Delta x}{x_\beta+\epsilon_s}}{\frac{\Delta x}{x_\beta}} = \frac{\epsilon_s}{x_\beta + \epsilon_s} \qquad (10)$$

Hence this analysis also indicates θ_β should be defined to place the reference point "near" the end of the avalanche paths.

To determine the effect of errors in due to our survey techniques is not quite as straight forward. For this analysis, we used the Monte Carlo technique discussed earlier. Here we focus on the effect of the errors on the model parameters (u and b), since these parameters define our statistical model. To do this, we take normally distributed errors with the standard deviations given in Table 2 and superimpose them on our recorded measurements. By recording the distribution of model parameters we can obtain a feel for the uncertainty in our model due to possible measurement errors. With $\theta_\beta = 18°$, we obtain the standard deviations on our parameter distributions shown in Table 3.

	Standard Deviation	
Source of Error	u	b
Starting Position	0.000534	0.000724
Distance (radii)	0.00157	0.000805
Slope	0.0122	0.00715
Run-out Position	0.00254	0.00179

Table 3: Parameter Standard Deviations for $\theta_\beta = 18°$

To determine approximate 95% confidence limits relative to each source of error, we multiply the standard deviations in Table 3 by 2 and add or subtract from the mean parameter value. However, at the moment, we are still addressing the relationship between the definition of θ_β and uncertainty in the model due to slope and distance measurements. Here we will focus on slope measurements since they act as the dominant source of uncertainty. The standard deviations for model parameter distributions due to slope measurement errors are shown as a function of θ_β in Figure 4.

We feel that the sensitivity of the model parameters to errors in slope measurements is primarily due to our survey technique; since we work our way along the avalanche path measuring each point relative to the previous point, error tends to accumulate as we move up the slope. As θ_β increases, the reference point moves up the slope where accumulated errors can create significant changes in the local slope angle of our second degree polynomial model of the avalanche path. Changing the local slope angle leads to movement of the reference point and consequently affects the geometric parameters Δx and x_β.

To clarify how this uncertainty affects our model, we need to take our analysis one step

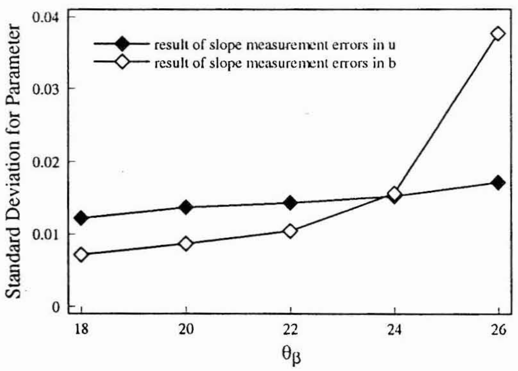

Figure 4: Standard deviations for model parameters due to uncertainty in slope measurements

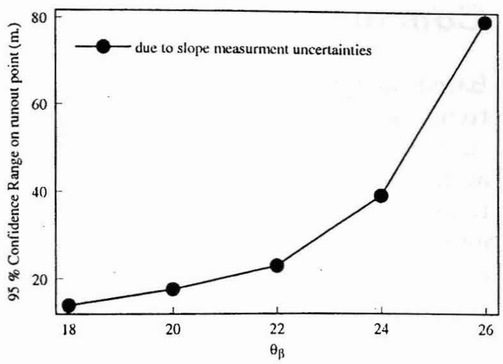

Figure 5: 95% confidence range on 0.95 quantile estimates of extreme run-out points

further. If we wanted to use this model to define an avalanche hazard zone, we would define our hazard zone in terms of avalanche run-out ratios with acceptable probability levels. For our example, we have chosen a 95% probability level. As mentioned previously, this implies that over the next 100 years, we expect 95% of all extreme avalanches to yield run-out ratios less than the run-out ratio calculated as the 0.95 quantile from our model. Now if we turn our attention to uncertainty in our model due to measurement errors. If we take into account the 95% confidence limits resulting solely from possible slope measurement errors, we have a confidence range on our 0.95 quantile run-out estimate as a function of θ_β , shown in Figure 5.

To calculate these data points, we took 2 times the parameter standard deviations then added and subtracted from the mean parameter values. This procedure gave us the 95% confidence limits on our cumulative distribution function and consequently on the 0.95 quantile estimates of the run-out ratios. To estimate the run-out point and the corresponding confidence limits, we took the median values for x_β for each definition of θ_β and multiplied by the run-out ratio estimates. The ranges shown in Figure 5 represent the difference between upper and lower 95% confidence limits. So for our 0.95 quantile estimate when $\theta_\beta = 18°$, we have confidence in our estimation of the extreme run-out point within ±7 meters; however, when $\theta_\beta = 26°$, we are only confident of our estimate within ±40 meters. Thus, to minimize uncertainty due to our surveying technique, these results also imply that θ_β should be defined so that the reference point lies near the end of the avalanche path.

In addition to the previously mentioned sources of uncertainty we also need to establish a quantitative measure of the quality of our model. In essence we'd like to know if there is a good chance that we've chosen the right statistical model. To do this we have chosen to use the W test as discussed in Wadsworth [6, §6.7]. Based on the size of our data set, with a 5% level of significance we should have a value for W between 0.369 and 0.979. Our data yields a W value of 0.939 which implies our model satisfies the requirements at a 5% level of significance; however, our model does not satisfy the requirements at a 1% level of significance.

Conclusions

Based on results obtained thus far, the extreme value distribution model bears promise as a user friendly tool for modeling even avalanches with small vertical drops; however, to be able to use this model more intelligently, natural sources of variation need to be further quantified.

The model developed here is based on a small data set and therefore cannot be considered reliable for estimating extreme avalanche run-out distances in the general case; however, the resulting extreme value distribution does help shed light on a few characteristics of the model. In particular, since most of our extreme run-out points never reached a local slope angle of ten degrees, our results indicate that the models constructed by McClung, Mears, and Schaerer are quite conservative when applied to avalanche paths in Southwest Montana.

To optimize the models sensitivity to measurement, the definition where $\theta_\beta = 10°$ used by McClung and Lied should work quite well, for large avalanche paths that run beyond local slope angles of ten degrees. For smaller avalanche paths which may not reach a ten degree slope angle, θ_β should be defined slightly greater than the slope of most of the run-out points, θ_r.

Acknowledgement

The work reported here was funded by the Army Research Office under Grant No. DAAL 03-92-G-0310. The Authors wish to express their appreciation for ARO's support.

The principal author is also grateful to Dave Macferran for many comments and discussions on modeling uncertainty and the corresponding statistical analysis techniques.

References

[1] D. McClung and K. Lied.
Statistical and geometrical definition of snow avalanche runout.
Cold Regions Science and Technology, 13:107–119, 1987.

[2] D. McClung and A. Mears.
Extreme value prediction of snow avalanche runout.
Cold Regions Science and Technology, 19:163–175, 1991.

[3] D. McClung, A. Mears, and P. Schaerer.
Extreme avalanche run-out: data from four mountain ranges.
Annals of Glaciology, 13:180–184, 1989.

[4] L. McKittrick and R. Brown.
A statistical model for maximum avalanche runout distances in Southwest Montana.
Annals of Glaciology, 17:In Press, 1992.

[5] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling.
Numerical Recipes.
Cambridge University Press, London, 1986.

[6] H. J. Wadsworth.
Handbook of Statistical Methods for Engineers and Scientists.
McGraw-Hill, New York, 1990.

[7] R. Walpole and R. Myers.
Probability and Statistics for Engineers and Scientists.
Macmillan, New York, 1972.