A Probabilistic Model to Evaluate the Optimal Density of Stations Measuring Snowfall

MARTIN SCHNEEBELI AND MARTIN LATERNSER

Swiss Federal Institute for Snow and Avalanche Research (SLF), WSL, Davos, Switzerland

(Manuscript received 12 February 2002, in final form 7 November 2003)

ABSTRACT

Daily new snow measurements are very important for avalanche forecasting and tourism. A dense network of manual or automatic stations measuring snowfall is necessary to have spatially reliable data. Snow stations in Switzerland were built at partially subjective locations. A probabilistic model based on the frequency and spatial extent of areas covered by heavy snowfalls was developed to quantify the probability that snowfall events are measured by the stations. Area–probability relations were calculated for different thresholds of daily accumulated snowfall. A probabilistic model, including autocorrelation, was used to calculate the optimal spacing of stations based on simulated triangular grids and to compare the capture probability of different networks and snowfall thresholds. The Swiss operational snow-stations network captured snowfall events with high probability, but the distribution of the stations could be optimized. The spatial variability increased with higher thresholds of daily accumulated snowfall, and the capture probability decreased with increasing thresholds. The method can be used for other areas where the area–probability relation for threshold values of snow or rain can be calculated.

1. Introduction

Snowfall in mountains is the major cause of the formation of avalanches. Avalanches are a major natural hazard for traffic and settlements in Switzerland and other alpine countries (Laternser and Schneebeli 2002; Mock and Birkeland 2000). Daily measurements of accumulated new snow are among the most important parameters to evaluate avalanche hazard (McClung and Schaerer 1993). New snow exceeding certain thresholds is well correlated with a certain degree of avalanche hazard. Wind and the stability of the already existing snowpack also contribute to the degree of avalanche hazard. The quality of hazard estimation depends strongly on how representative the snowfall measurements are. The variability of snowstorms is caused by storm type and orography (Spreitzhofer 1999). Serreze et al. (2001) discuss the distribution of large snowfall events but do not calculate the area of typical events. Weather radar provides high-spatial-resolution reflectivity observations (Germann and Joss 2001), but estimating precipitation accumulation or rates from such observations remains difficult (Gabellia et al. 2000); further, the intensity and amount of snowfall are difficult or impossible to determine with operational weather radars in alpine terrain (J. Joss 2003, personal communication). A dense network of gauges measuring snowfall either manually or automatically is therefore necessary to predict avalanche hazard. Such stations are costly, and their number should be optimized based on a rational criterion. Martinec and Sevruk (1992) list general criteria for station density regarding snow water equivalent, but not with regard to daily snowfall thresholds.

For model cases, Epstein (1966) calculates the optimal distribution of stations based on cell size area. Krzysztofowicz (1999) develops the method of the probabilistic quantitative precipitation forecast, which relates a point precipitation amount to a spatially averaged precipitation amount. A different approach to relating point precipitation to area precipitation is the different methods that use an areal reduction factor (ARF; Asquith and Famiglietti 2000; Rodriguez-Iturbe and Mejia 1974; Sivapalan and Blöschl 1998). ARFs are empirically derived functions of catchment area, storm duration, and return period. Bastin et al. (1984) determine the optimal placement of rainfall gauges for average area rainfall. These models differ from our needs. We are interested not in the optimization of the network with respect to the average area precipitation, but in the probability that a certain threshold of daily accumulated snowfall is measured.

We developed a statistical method that delivers a spatially distributed probability to observe a certain threshold of daily snowfall for a given network of stations. First, the area–probability relationship for different snowfall thresholds was calculated. Second, a probabilistic algorithm to calculate the capture probability...
(CP) of a station network was developed. This algorithm was applied 1) to artificial networks to construct graphs for an optimal spacing of stations, and 2) to different existing networks in the Swiss Alps and for different snowfall thresholds. Based on these results, rational arguments for the optimization of the existing networks measuring snowfall are proposed.

2. Data and method
   a. Data

   Thresholds of accumulated daily new snow (NS) that are important for avalanche warning are 20, 30, or 50 cm of new snow (NS20, NS30, NS50) or for snowfall within 3 days to exceed 75 cm (NS3/75) (Schneebeli et al. 1997; Mock and Birkeland 2000). New snow is analyzed at 107 manual stations located between 1150 and 1850 m MSL in the Swiss Alps for the 30-yr period of 1969/70–1998/99 (Fig. 1, inset). The data were checked for consistency (for details see Laternser and Schneebeli 2003). The analyzed stations were restricted to an altitudinal sector because at lower altitudes part of the precipitation is often falling as rain and at higher altitudes the measurements are increasingly influenced by wind. For each day on which one or more of the threshold criteria were fulfilled, the NS data were linearly interpolated and resampled to a grid with a spacing of 5 km × 5 km (Fig. 1). The size of all areas exceeding the threshold was calculated. A single snowfall event, which can spread across all of Switzerland, can have regionally different NS, causing separated areas for 1 day (in Fig. 1 five areas). Each area of an event is considered independently. Because of the varying spacing between stations, an error estimate of a single area is difficult and depends on the size of the NS area. Small areas are probably underrepresented in the evaluation.

Fig. 1. Interpolated NS30 areas for an example day (black areas). The snowfall event on this day produced five areas of different size. White triangles show stations with NS > 30 cm, black dots are stations with NS ≤ 30 cm, and the gray area includes the entire interpolation area. Inset: Map of Switzerland with the snow stations used for the daily NS area calculation. Gray levels indicate the topography (white: <1000 m, light gray: 1000–2000 m, and dark gray: >2000 m). A dense station cover is primarily of interest in the areas where the topography is greater than 1000 m.
However, the large number of samples reduces uncertainties in the determination of the snowfall areas. Histograms for all thresholds were constructed from the resulting areas. Figure 2a shows an example for NS30.

### b. Network of stations

The existing snow station networks of Switzerland consist of manual and automatic stations, which are operated at variable spatial and temporal resolution (details in Laternser 2002). Four different network types were evaluated (Table 1). The automatic network is relatively young and increased from 7 stations in 1992 to more than 90 in 2002. Most of the manual network was established in the 1950s and is reporting daily. The manual and automatic networks are mutually exclusive. The operational network included all daily transmitting snow stations in Switzerland used for operational purposes (avalanche warning, weather forecasting) and is the union of the automatic and manual network. The climatological network includes additional manual stations that report only in monthly intervals. In contrast to the calculation of the NS area, the probabilistic model uses all stations, without altitudinal restriction.

### c. Calculation of the capture probability

The CP is defined as the probability that an event (snowfall exceeding a threshold) occurring at a location \( s \) is observed by a station. The capture probability for a specific threshold for each location in a station network is determined with the following steps: the calculation of the empirical histogram of the precipitation threshold in question, the transformation of the histogram to a two-dimensional probability distribution, and the calculation of the capture probability for each location in the network. As an example, the threshold of 30 cm of new snow is used. The empirical histogram for all 1786 NS30 areas of variable size is shown in Fig. 2a. The histogram was transformed to a probability distribution by dividing by the number of areas. The probability in relation to area (area–probability relation) for NS30 areas is shown in Fig. 2b. The calculation can proceed either with the empirical probability or with a fit to the distribution. In this work the probability distribution for NS20 and NS30 was fitted to a log-linear curve, and that for NS50 and NS3/75 was fitted to a local regression model. The areal shape (projection to the surface) of the two-dimensional probability distribution can be calculated directly from the data or modeled as a suitable geometrical shape. In the remainder of this work, the shape of the probability distribution of the snowfall thresholds was assumed to be circular. A circular area was most convenient for the calculations and was not too different from the observed areas. With this model, the radius \( r \) of an area \( A \) is \( r = (A/\pi)^{1/2} \).

The area–probability relation of the logarithmically transformed NS30 area \( A_{NS30} \) to the cumulative probability \( P \) was nearly linear for areas of up to 5000 km²:  

\[
P(A_{NS30}) = 0.31[\log_{10}(5000) - \log_{10}(A_{NS30})] 
\]

for \( 3 < A_{NS30} < 5000 \). (1)

### Table 1. List of examined combined snow-station networks.

<table>
<thead>
<tr>
<th>Type of network</th>
<th>No. of stations (active in 2001)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic</td>
<td>88</td>
<td>All automatic stations</td>
</tr>
<tr>
<td>Manual</td>
<td>135</td>
<td>All manual operational stations</td>
</tr>
<tr>
<td>Operational</td>
<td>223</td>
<td>All operationally reporting stations (automatic and manual)</td>
</tr>
<tr>
<td>Climatological</td>
<td>268</td>
<td>All stations combined for retrospective analyses</td>
</tr>
</tbody>
</table>
This log-linear relation was not valid for areas larger than about 5000 km$^2$, a circular area with a radius of 40 km. We suppose that this is an artifact of our area being limited to the size of the Swiss Alps. The probability for areas >5000 km$^2$ was set to 0, and therefore the station cannot see any NS30 outside. The area with a $P > 0$ is called area of visibility $A_n$, with the corresponding radius $r_n$ and diameter $d_n$. This simplified the modeling and was justified because in our application we were interested not in the extremely large and very rare events, but in the events with a high occurrence.

In the next step, the CP at each location $s(x, y)$ in an area of interest populated with stations was calculated. The CP around a single station is determined by the distance between station and location. In this case CP is identical to $P$. For two stations where the areas of the probability distributions are not intersecting, the CPs are the same as around a single station (in the example, if $d_n$ is larger than 80 km). If $d_n$ between the stations is smaller, the probability distributions interact in the overlapping areas. The joint probability $CP_j$ at a location $s(x, y)$ for a point influenced by $n$ stations is

$$CP_j(x, y) = 1 - \{[1 - CP_A(x, y)] \times [1 - CP_B(x, y)] \times \ldots \times [1 - CP_n(x, y)]\}, \quad (2)$$

where $CP_A$ is the independent CP around station $A$ at location $(x, y)$ (Fig. 3a). If the stations are in a statistical sense completely independent, then $CP_A$ is the correct probability.

Another solution is that the distance from the location $s$ to the nearest station determines the probability. The minimal distance to a station is calculated for each location $s$ in the region of interest. The minimal distance capture probability $CP_{\text{md}}$ is then determined by the distance $l$ to the nearest station $A$ and the corresponding capture probability $CP_A$ (Fig. 3b),

$$CP_{\text{md}}(x, y) = CP_A(x, y), \quad \min(l_A). \quad (3)$$

The local joint probabilities $CP_j$ are the maximal possible values, and the minimal distance probabilities $CP_{\text{md}}$ are the minimal possible values. Because there is spatial autocorrelation, $CP_j$ overestimates the true value; $CP_{\text{md}}$ underestimates the capture probabilities because the CP at a location $s(x, y)$ is not only determined by the nearest station. If the nearest station were to be eliminated, there is still the CP from the farther station.

The autocorrelation $\gamma$ for a location at a distance $l$ from the center of a station $A$ was calculated by integrating the probabilities of the intersecting areas of influence $CP_{\text{sect}}$ when the CP areas of two stations were moved closer and closer together,

$$\gamma(l) = \int CP_{\text{sect}}(l) \left/ \left( 2 \int CP_A \right) \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \righthand side of (5)}}. \quad (4)$$

The term $\left( 2 \int CP_A \right)$ normalizes the autocorrelation to a value between 0 and 1. The capture probability corrected for autocorrelation $CP_{\text{mod}}$ for two intersecting stations is (Fig. 4a)

$$CP_{\text{mod,}AB\{A\}} = 1 - \{1 - CP_A\} \{1 - CP_B(1 - \gamma)\} \quad (5a)$$

and

$$CP_{\text{mod,}AB\{B\}} = 1 - \{1 - CP_B\} \{1 - CP_A(1 - \gamma)\}. \quad (5b)$$

The calculation of $CP_{\text{mod}}$ at a location $s(x, y)$ that is influenced by several stations required an iterative procedure. First, $CP_{\text{mod}}$ for stations $A$ and $B$ was calculated [Eq. (5)]. Then $CP_{\text{mod,}AB\{A\}}$ was used as partial input for the calculation of $CP_{\text{mod,}AB\{B\}}$, and this procedure was continued until all intersecting areas were calculated. Last, all modified capture probabilities were superimposed,

$$CP_{\text{mod}} = \text{mean}(CP_{\text{mod}}, CP_{\text{mod}}, \ldots). \quad (6)$$

The results for $CP_j$, $CP_{\text{md}}$, and $CP_{\text{mod}}$ are compared in a cross section in Fig. 5. The $CP_{\text{mod}}$ is in between the
extreme solutions and cannot be distinguished from the average calculated from CPj and CPmd. In contrast to CPmod, CPj and CPmd are much easier to calculate. In addition, a variable shape of the capture probabilities around each station can, in principle, be taken into account to calculate CPj and CPmd.

3. Results

a. Area–probability relationship

The area–probability relationships for the different thresholds are shown in Fig. 6. The trend line for NS20 and NS30 was fitted well by a log-linear regression, and that for NS50 and NS3/75 by a local-regression model. The cumulative probability decreased from NS20 to NS50 for the same area; thus, the area of influence also decreased. NS3/75 followed NS20 for small areas below 200 km², but then the probability decreased as compared with NS20 and NS30. Up to an area of 500 km² (a radius of 12 km around the station) the relations were very similar for the different thresholds, except for NS50. The difference in the cumulative probability was at most 15% for the different thresholds. The autocorrelation functions (Fig. 7) were nearly identical for all thresholds except NS20.

b. Optimal design of networks

To find the optimal design of networks for a given area–probability relation, triangular equilateral grids of different spacing were constructed in an area of 150 km × 150 km. The triangles had side lengths of 5–50 km. The vertices were used as stations, and the capture probabilities for the different thresholds were simulated. The minimal capture probability CPminimum and the average capture probability CPmean within the central rectangular area of 50 km × 50 km were calculated for each grid. The relation between station spacing and the capture probabilities CPminimum and CPmean was different for the various snowfall thresholds (Fig. 8). A CPminimum = 80% for the ideal network was spaced for NS20 by 16 km, for NS30 by 13 km, for NS50 by 11 km, and for NS3/75 by about 13 km. The corresponding spacing for CPmean = 80% was about 1 km longer. Whereas for a high CP the optimal spacing for CPminimum and CPmean did not differ much, for a lower CP the difference between CPminimum and CPmean increased. The optimal station spacing for CPminimum and CPmean differed by 10 km for a CP of 50% for NS50.

c. Capture probabilities for different real networks

The calculated CP for the different snow station networks (Table 1) with regard to NS30 is shown in Fig. 9. Capture probabilities are most important in the mountainous area of Switzerland for avalanche forecasting.
but also in the flatter areas new snowfall sometimes 
causes serious disturbances of the traffic (Fig. 1, inset).
The distribution of automatic stations (Fig. 9a) was 
fairly good throughout the high mountains, resulting in CPs 
of greater than 70% in most areas. Major gaps existed 
in the central east (CP ≈ 50%), in the border areas, and 
in the southern valleys of the southeast (CP < 50%).
There were no stations at all in the middle and lower 
south, because the avalanche hazard is small in these 
regions. The manual stations (Fig. 9b) were evenly dis-
tributed across the Swiss Alps, with CPs of at least 50%.
Much higher CPs were only reached directly around 
stations and near station conglomerates; best covered 
were the central regions. In the lower areas north of the 
Alps (Swiss plateau, Jura Mountains) distances between 
stations were too large to result in spatially suf-
ficient CPs. In case the automatic network failed, the manual 
stations alone reached CPs of greater than 60% in most 
Alpine regions. The operational network resulted in a 
good coverage throughout the Alps (Fig. 9c). The CPs 
were greater than 90% for nearly one-half of the area, and 
most remaining areas had CPs of greater than 80%. 
Problematic regions with lower CPs were widespread 
in the south, in some parts in the east, and along the 
northern foothills. The sparse station coverage in the 
lower areas of the Swiss plateau and the Jura Mountains 
was not critical because these areas are not prone to 
avalanching. However, all border areas were insufficiently 
covered. The climatological network (Fig. 9d), 
which was only useful for retrospective climate anal-
yses, solved most of the deficiencies of the operational 
one. The histograms of the spacing between the stations 
(Fig. 10) indicate the homogeneity of the networks with 
regard to capturing heavy snowfall areas. Ideal histo-
grams had a dominant peak (mode) around the optimal 
distance (13 km for NS30 areas and CP_{minimum} = 80%) 
and strongly declining tails to the left and right. All 
networks had their mode near to the optimal spacing, 
but the tails were long, indicating a too-large or too-
small spacing. The automatic network was best in this 
respect, whereas the manual network had the broadest 
spacing, also caused by the stations on the plateau and 
the Jura. The histogram of the operational network had 
its mode between 5 and 10 km with a steep decrease 
for larger spacing and was, thus, near to an optimal 
configuration. However, the spacing of about 10% of 
the stations was too small to improve the capture prob-
ability. The climatological network was very similar to 
the operational network, but the number of stations with 
a spacing below 15 km increased further. The clima-
tological network was well suited to study the spatial 
distribution of snowfall because of this oversampling.
d. Effects of different thresholds on the capture probability for the operational network

The distributions of the capture probabilities for different thresholds of daily snowfall calculated for the operational network are shown in Fig. 11. The cumulative probability of NS20 and its autocorrelation $\gamma$ caused the highest capture probability (Fig. 11a), because the cumulative probability has the broadest tail. An increasing threshold (NS30, NS50) caused an increasingly poor capture probability (Figs. 11b,c). The initially steep probability function for NS50 led to considerably reduced CPs as compared with the NS30 situation, but the effect was soon reduced with growing distance away from the stations. The CP distribution for NS3/75 was in most areas above 0.8, except in the eastern part of the Swiss Alps (Fig. 11d). The comparison with the CP distribution of NS30 showed subtle differences caused by the different area–probability relationships and autocorrelations.

4. Discussion

The proposed method required as input a precipitation area–probability relation. This relation can be calculated...
if spatially sufficiently dense and long-term records of snowfall exist. The required spatial density may vary depending on the weather system and the topography. The area of snowfall in complex terrain varies by geographic region (Serreze et al. 2001). The simulation of the capture probability assumes a probability distribution around a station that is not dependent on the direction (isotropic). The effective shape of a precipitation area is complex (Fig. 1) and no preferred direction could be detected visually within the several hundred days with snowfall. The calculated NS30 areas were only in a few cases elongated parallel to the main direction of the curved Alpine ridge. An isotropic probability distribution with a circular base area was therefore a reasonable approximation for our data.

The simulated network is an efficient way to determine the necessary spacing of new networks and to evaluate the effect between different thresholds of daily accumulated snowfall. The area—capture probability curves (Fig. 8) also show that the spacing is most sensitive between 10 and 20 km for $CP_{\text{minimum}}$ and between 10 and 30 km for $CP_{\text{mean}}$. For NS50, a spacing of 20 km between stations results in a $CP_{\text{minimum}}$ of only 0.42, and a spacing of 10 km results in a $CP_{\text{minimum}}$ of 0.85. The area curves could also be used as one input parameter to calculate the marginal costs for an optimal network. In this case, also a quantification of the cost and probability of avalanches caused by an NS50 area would be necessary.

The spatial distribution of the capture probabilities for different networks (Fig. 9) is an essential tool to visualize the quality of the network. Regions with either too-high or too-low station densities can be identified. Based on the spacing distribution (Fig. 10), stations with a too-small spacing could be eliminated, this is the case for about 10% of the operational network. If the same number of stations is kept, the distribution can be optimized toward more homogenous coverage. The existing operational network has the lowest capture probability along the national border. This could be improved by creating an international network including all stations within the European Alps.

The capture probabilities of the operational network for different thresholds (Fig. 11) show an increasing uncertainty with increasing threshold (from NS20 to NS50). This is relevant to avalanche forecasting because it must be expected that with higher daily accumulated snowfall the spatial resolution gets increasingly smaller. A map on which the capture probability is visualized will increase the reliance of the forecaster on the inter-
pretation of maps of new snowfall. In practice, the avalanche forecaster must be aware that NS50 (and larger thresholds) will often go unobserved by the stations and that isolated spots with larger-than-observed NS are likely.

The proposed method could be used with little or no adaptation for other European countries (Austria, France, Germany, Italy) for which other snow-measuring networks exist. Although the derivation of the area–probability relationship was developed with data only from Switzerland, it is very probable that this relation can be directly used elsewhere in the Alps, because the storm systems are similar. Area–probability relations for other storm systems (e.g., Rocky Mountains, Scandinavia) will likely be different. The method then requires an adapted area–probability relation.

Additional studies, including orographic precipitation effects and redistribution of snow by wind, could further improve knowledge of the spatial distribution of daily accumulated snowfall. However, such studies would require a spatially much higher resolution of the station network, probably below 5-km spacing (Doorschot et al. 2001). Weather radar could be extremely useful (Rasmussen et al. 2003), but it is not clear whether the actual problems of determining the fall velocity of the snow crystals can be resolved for operational radars in alpine terrain. The extension of the proposed method to non-circular and variably oriented base areas for the capture terrain. The extension of the proposed method to non-circular and variably oriented base areas for the capture probability could take into account orographic effects.

Additional studies, including orographic precipitation effects and redistribution of snow by wind, could further improve knowledge of the spatial distribution of daily accumulated snowfall. However, such studies would require a spatially much higher resolution of the station network, probably below 5-km spacing (Doorschot et al. 2001). Weather radar could be extremely useful (Rasmussen et al. 2003), but it is not clear whether the actual problems of determining the fall velocity of the snow crystals can be resolved for operational radars in alpine terrain. The extension of the proposed method to non-circular and variably oriented base areas for the capture probability could take into account orographic effects. This extension could be realized by modifying the empirical CP by taking into account topography and wind directions prevailing during snowfall.

Acknowledgments. We thank the anonymous reviewers for their detailed comments to improve the manuscript, Dr. J. Joss for comments on the use of weather radar for snow precipitation, and MeteoSwiss for providing access to its climate database.

REFERENCES


