FOREST CANOPY EFFECTS ON SAMPLE SIZE REQUIREMENTS IN SNOW ACCUMULATION AND MELT COMPARISONS

D.L. Spittlehouse[1] and R.D. Winkler[2]

ABSTRACT

Forest management regulations in British Columbia attempt to minimise changes in the hydrologic regime by limiting the extent of forest removal. Considerable effort is currently being focused on the quantification of snow pack differences between a variety of forest treatments and stand types. This paper assesses what size of difference can reliably be detected by the methodologies used in these studies. We examine the sample sizes required to detect specified differences in April 1st. snow water equivalent and in the average melt rates resulting from harvesting an Engelmann spruce - subalpine fir forest. Snow water equivalent was measured at 64 points in a clearcut, in thinned and unthinned juvenile pine stands, and in a mature forest. A power analysis was used to determine sample sizes necessary to detect specified differences. The standard 10 point sampling approach was found to be adequate to detect differences of greater than 6 cm in snow water equivalent, and 1 cm d^{-1} in the mean daily snowmelt rate calculated over a six day period, for $\alpha = 0.05$ (Type I error) and power = 0.9 (Type II error). Increasing sample size to 30 points substantially decreased the size of the difference in snow water equivalent that could be reliably measured. However, 50 samples were required to determine differences in melt rate between the forest and clearcut of 0.5 mm d^{-1}.

INTRODUCTION

New forest practices regulations and guidelines in British Columbia (B.C.) attempt to protect water resources by specifying operational procedures aimed at minimising changes in the hydrograph of streams. In the southern interior, hydrographs are typically dominated by a single, snowmelt generated peak. Forest removal and regrowth of the forest are considered to have a significant effect on snow accumulation and melt, and thus on the hydrograph (Golding 1982, Toews and Gluns 1986). The effects of forest harvesting and regrowth on the snow pack are highly variable, both spatially and temporally. The within-site variability is often as great or larger than the mean difference between sites. Consequently, errors in the interpretation of study results is of concern. Winkler and Spittlehouse (1995) reviewed 40 papers evaluating the effects of harvesting on snow accumulation and melt. They found that none of the authors addressed the adequacy of their sample sizes, nor did they assess their data in terms of detecting effects that were hydrologically significant.

The B.C. Ministry of Forests has established a number of snow research projects in the southern interior of B.C. These projects are intended quantify differences in snow accumulation and melt under a range of canopy types resulting from forest management activities. The need to sample a large number of sites over a large area and cost considerations limit the number of sample points that can reasonably be taken at any one site. A common practice has been to follow the sampling procedure used by the B.C. Ministry of Environment for flood and water supply forecasting. Snow depth and water equivalent are measured at 10 points spaced 10 m apart along a level transect (Ministry of Environment 1981). Wyckoff (1957) expressed

---

[1] D.L. Spittlehouse, Forest Climatologist, B.C. Ministry of Forests, Victoria, B.C.
[2] R.D. Winkler, Regional Research Hydrologist, B.C. Ministry of Forests, Kamloops, B.C.

Presented at the 64th. Western Snow Conference 1996, Bend, Oregon

every snow surveyors concern "...is it necessary to collect more than 10 samples on one snow course? ... every sample can bring forth blood, sweat, tears and a profusion of profanity". Forty years later the question remains - what level of sampling is required to reliably determine the influence of forest management on snow accumulation and melt?

In a preliminary attempt to answer this question, Winkler and Spittlehouse (1995) used power analysis in a comparison of two mature forest types and adjacent clearcuts near Penticton, B.C. They found that the traditional 10 point snow sampling method was adequate to detect differences in April 1st. snow water equivalent (SWE) of 7 cm or greater, 90 times out of 100 at a significance level of 95%. A sample size of 50 was required to detect differences of 3 cm at the same statistical level.

In this paper we evaluate the sampling necessary to detect the effects of clearcut logging and forest regrowth on snow hydrology at a site near Kamloops, B.C. Power analysis is used to assess the ability to detect specified differences in the April 1 SWE, and in the maximum mean daily snowmelt rate over six days, between a mature forest, a clearcut, and two juvenile stands.

METHODS

Study area

The spatial and temporal variability in SWE and snowmelt was determined at Mayson Lake (51° 13' N, 120° 24' W), at 1300 m in gently undulating terrain, 55 km NNW of Kamloops, B.C. This site is about 200 km NNW of that described by Winkler and Spittlehouse (1995). The mature forest is a mixed stand of Engelmann spruce (*Picea engelmannii* Parry), subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.) and lodgepole pine (*Pinus contorta* Dougl.). Stand density is about 2000 stems ha^{-1} and mean tree height is about 23 m.

The clearcut was logged and planted with lodgepole pine which are now about 1 m tall. The two juvenile stands consist of 6-m-tall lodgepole pine. The unthinned juvenile stand density is about 4000 stems ha^{-1} with the base of the crown close to the ground. The thinned stand has approximately one quarter of the density of the unthinned stand and the pruned base of the canopy is at 3 m. These canopy types are typical of forests, clearcuts and regenerating stands found throughout the southern interior of B.C.

Sampling method

Results reported here relate to snow surveys done in 1995. The SWE was measured monthly from January through March, weekly in early April, and then every three days during the period of rapid melt. Measurements were made using a standard snow tube at 64 points, 15 m apart on a 105 m by 105 m grid. All measurements were made a minimum of 50 m from the edge of a treatment unit. These measurements are part of a larger project to assess the influence of forest canopies on the energy balance of a melting snow pack (Adams et al. 1996).

Data analysis

Statistical analyses in forest hydrology commonly involved testing the hypothesis that some forest treatment has had no effect on the hydrologic variable of concern. The t-test and analysis of variance (ANOVA) F-test are often used to compare treatment effects with the 'undisturbed' situation. These tests assume that the data are normally distributed, have equal error variances and independent error terms (Neter et al. 1990). It is unlikely that these basic assumptions are ever met with hydrologic data. However, the F-test is robust and Neter et al. (1990) note that problems can be minimised through the use of equal and large sample sizes, which is the case in our study.

Most analyses have concentrated on minimising the risk of concluding that a difference exists between sets of data when one does not occur (Type I error, denoted as α). Not usually considered, but of equal importance, is failing to conclude a difference exists when one actually does (Type II error, denoted as β). Avoiding Type II errors is particularly relevant to resource management issues when there are high economic and/or environmental costs associated with an incorrect decision (Peterman 1990).

Power analysis is used to determine the numbers of samples required to detect between-treatment differences. Power is defined as "the probability of detecting a difference from an experiment if in truth a real difference of a certain magnitude exists" (Sanders 1989). Power is commonly denoted as $1 - \beta$. The greater the probability of detecting departures from the null hypothesis, the greater the power of the test. Power is increased by decreasing the significance level (i.e. from 0.05 to 0.10), increasing the difference between means, and by increasing sample size (Nemec 1991). In our study, power analysis was done for each combination of treatment pairs to determine the number of samples which would be required to ensure that specific differences could be detected. The analysis used a program provided by the Biometrics Section, B.C. Ministry of Forests (V. Sit, pers. comm., 1994) written to run in SAS (SAS Institute Inc. 1988). All computations were done for a significance levels (α) of 0.05. The choice of the 0.05 significance level is arbitrary. Depending on the potential downstream impacts, the researcher might chose α=0.01 if the risk associated with conclusion error is high, or α=0.1 if the risk is low. If the latter is acceptable then the power for any sample size is greater than at α=0.05, particularly for small differences (Winkler and Spittlehouse 1995). We use a power of 0.9 as the level for minimising the risk of Type II errors.

Power was determined for differences in SWE and mean maximum snowmelt rates of 10%, 30%, 50% and 100% (melt rate only) of the average measured in the undisturbed forest, for sample sizes of 2 to 100 in increments of 2. The sample numbers cover survey sizes from the quick, index type survey to the upper, physically tolerable limit at which all measurements in a treatment pair can be made by a two person crew in one day.

SWE on April 1st. was chosen in this analysis because it is frequently used as an index of the maximum SWE at a site. This of course is not necessarily true because the value will vary with latitude, elevation and year. However, it is considered a useful index of year-to-year variation in snow accumulation. 64 values for each treatment were used in the power analysis.

Snow ablation rates were calculated for each point as the difference in SWE between two measurement periods divided by the number of days. Sublimation and evaporation rates are much less than 10% of the daily snowmelt rates during this period (Adams et al. 1996) and for the purposes of this analysis we will assume the melt rate equals the ablation rate. Snowmelt was calculated for April 22 to 28, the period of mean maximum daily snowmelt for all treatments (Figure 1). All points were included in this average, even if they were free of snow because the concern is how the whole treatment is influencing the hydrology. Sixty points per treatment were used in the analysis because several points in two treatments had extremely wet snow that could not be sampled, and the power analysis requires equal sample sizes. This mean maximum daily snowmelt rate is assessed because hydrologic problems often occur during periods with high rates of snowmelt.

There are four sources of error in the measurement technique. There is an error due to the fact that the same location cannot be sampled each measurement day. There is the problem that all measurements cannot be made simultaneously and that SWE will change during a day with a high melt rate. There is the potential for operator error, though all operators were experienced. The error due to the snow tube itself is the only one that can readily be quantified. Repeated weighings of the snow tube suggest a reliability of $\pm$ 1 cm for a SWE measurement. This results in an error of ± 0.3 cm d^{-1} for the average daily snowmelt rate over six days. The variability in the data being analysed reflect the contributions of all these errors.

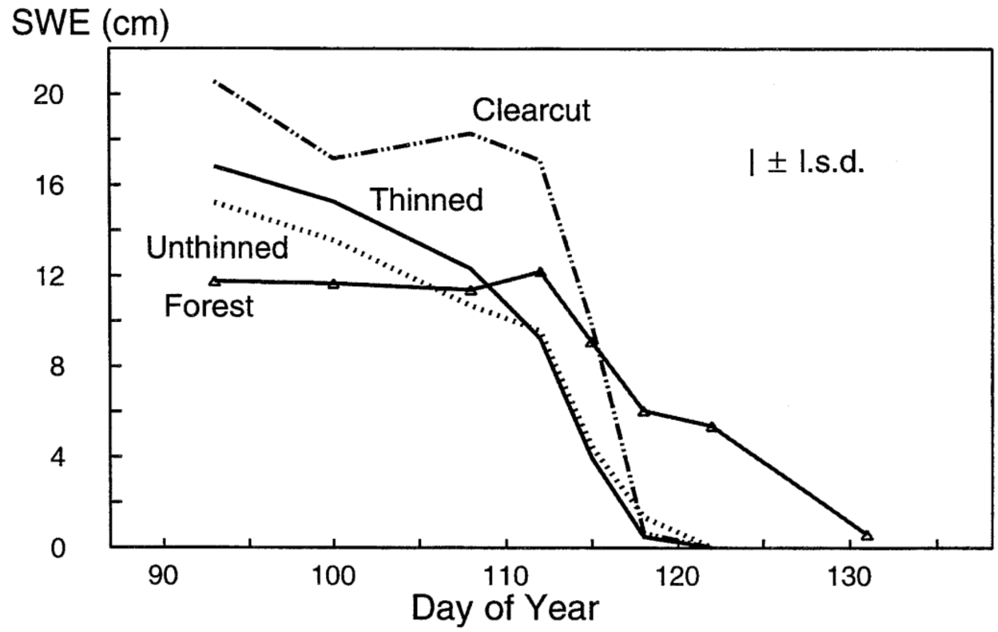

Figure 1. Snow water equivalent (SWE) from April to the end of snowmelt in 1995 for the clearcut (dashed and dotted line), thinned stand (solid line), unthinned stand (dotted line) and the mature forest (solid line with triangles). The bar represents ± least significant difference (l.s.d.) of the mean.

Table 1. A summary of April 1st. snow water equivalent (cm) for the four treatments for the 64 point grid. Maximum, mean, minimum and the standard deviation (s.d.) of the 64 values, and the standard error (s.e.) of the mean are presented.

Treatment	Max.	Mean	Min.	s.d.	s.e.
Clearcut	30	20.5	7	4.5	0.6
Thinned	26	16.8	8	3.9	0.5
Unthinned	24	15.2	4	5.3	0.7
Forest	22	11.8	2	4.2	0.5

RESULTS AND DISCUSSION

Figure 1 shows the changes in SWE from the beginning of April through to the end of the melt period. The data for the April 1 SWE are summarised in Table 1 and for the mean maximum daily snowmelt rate in Table 2. On average, the forest had 74% less snow than the clearcut, 42% less than the thinned stand and 29% less than the unthinned stand. The standard deviations were similar at all sites, varying from 22 to 36% of the mean. Differences in the mean daily melt rate (Table 2) between the forest and treatments are 170, 60 and 40%for the clearcut, thinned and unthinned sites, respectively. However, the standard deviations were larger than those for the SWE, varying from 26 to 70% of the mean.

Table 2. A summary of April 22 to 28 mean daily snowmelt (cm d^{-1}) for the four treatments for 60 sample points. Maximum, mean, minimum and the standard deviation (s.d.) of the 60 points and the standard error (s.e.) of the mean are presented.

Treatment	Max.	Mean	Min.	s.d.	s.e.
Clearcut	4.2	2.7	0.4	0.7	0.1
Thinned	7.5	1.6	0	1.4	0.2
Unthinned	3.5	1.4	0	1.1	0.1
Forest	3.5	1.0	0	0.7	0.1

The magnitudes of differences in SWE between the sites described here are consistent with differences observed at other locations in the dry, southern interior of B.C.. All April 1st differences recorded between 12 pairs of clearcut and forest sites over the past 4 years have been less than 10 cm and are frequently less than 5 cm (R. Winkler unpubl. data). Both above and below normal snow years are included in these data. Toews and Gluns (1986), using 10 point samples obtained differences ranging from 5 to 30 cm for south-eastern B.C., with the largest differences occurring in an above normal snow year.

A large increase in the power to detect a specified effect occurs with increasing sample numbers (Figure 2). Under our arbitrary limits of a 5% chance of making a Type I error ($\alpha = 0.05$) and a 10% chance of making a Type II error (power = 0.9), even 100 samples would not be adequate to detect a 10% (1.2 cm) difference between a clearcut and a forest in April 1st. SWE. Thirty samples are required to detect a difference of 3.6 cm in SWE. The standard 10 sample transect could reliably detect differences of greater than 50% (> 6 cm) in SWE. Winkler and Spittlehouse (1995) obtained an equivalent lower limit of 7 cm. Relationships similar to that in Figure 2 were obtained for the juvenile stands and the forest comparisons (Table 3). In these cases, detection of differences of 30% of the forest SWE requires 30 samples at $\alpha = 0.05$, and power = 0.9.

The power analysis for differences in mean daily snowmelt rate between the clearcut and forest (Figure 3 and Table 4) gave similar results to that for the April 1st. SWE. However, more samples are required to achieve the same reliability for any difference. A 10-point sampling procedure would be adequate to detect differences of 1 cm d^{-1}, or more, in melt rate (Table 4) between the clearcut and the forest, or the unthinned stand and the forest. Only a few more samples would be required for the thinned/forest comparison. Comparisons between the other treatments are not presented but have similar limits to those for the forest comparisons.

This analysis has only considered one set of treatment pairs. An increase in power can be achieved by having more of the same pairs (Nemec 1991), i.e., 3 clearcut/forest pairs of 10 point samples each would be more or less the same as 30 samples. However, this requires that each pair have similar stand composition, site physiography and probably climatic regime. An important question remains to be answered by further research - what is a hydrologically significant difference (change from the undisturbed forest situation) in SWE or in snowmelt rate? In other words, what is the change that would cause unacceptable effects on the hydrology of the watershed.

CONCLUSIONS

Large areal variation in SWE occurs below a forest canopy and in a clearcut. There is a need to determine the size of treatment effect that is hydrologically important and the acceptable level of risk associated with failing to detect such effects. Where SWE differences of less than 6 cm are expected to have an important

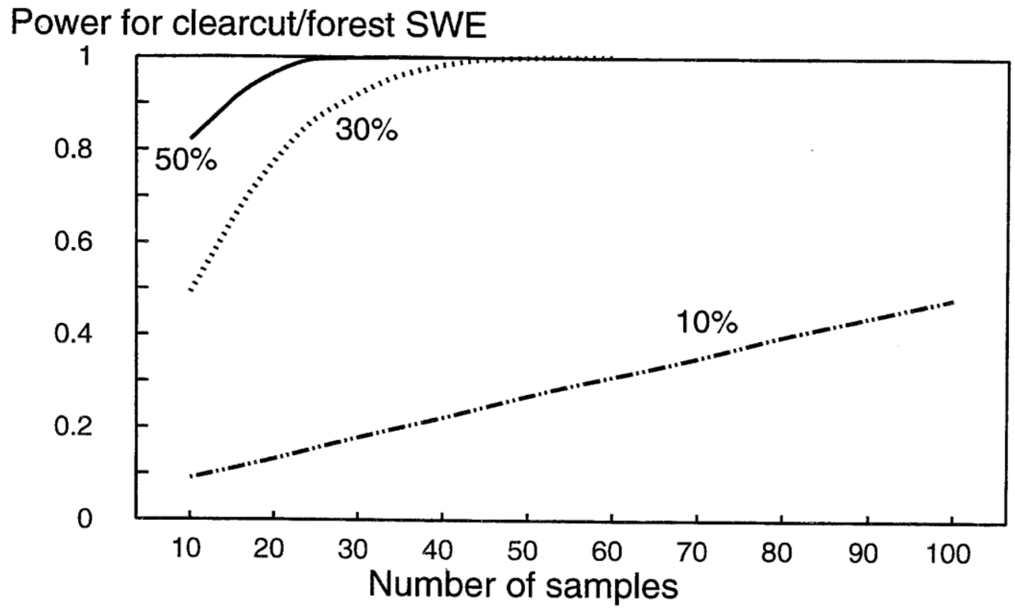

Figure 2. Power analysis for the clearcut/forest comparison for April 1st. SWE. Lines are for differences of 10, 30 and 50% of the forest value at α=0.05. Data are summarised in Table 3.

Table 3. Power analysis for comparisons of snow water equivalent between the treatments and the mature forest for sample sizes of 10, 20 and 50, and differences of 10, 30 and 50% of the forest value, i.e., 1.2, 3.6 and 6 cm, respectively.

		Power at differences of		
Sample size	Treatment pair	10%	30%	50%
10	Clearcut/Forest	0.09	0.49	0.82
10	Thinned/Forest	0.08	0.42	0.75
10	Unthinned/Forest	0.09	0.54	0.87
20	Clearcut/Forest	0.13	0.80	0.99
20	Thinned/Forest	0.12	0.73	0.97
20	Unthinned/Forest	0.14	0.85	0.99
50	Clearcut/Forest	0.27	0.99	1.00
50	Thinned/Forest	0.23	0.98	1.00
50	Unthinned/Forest	0.30	1.00	1.00

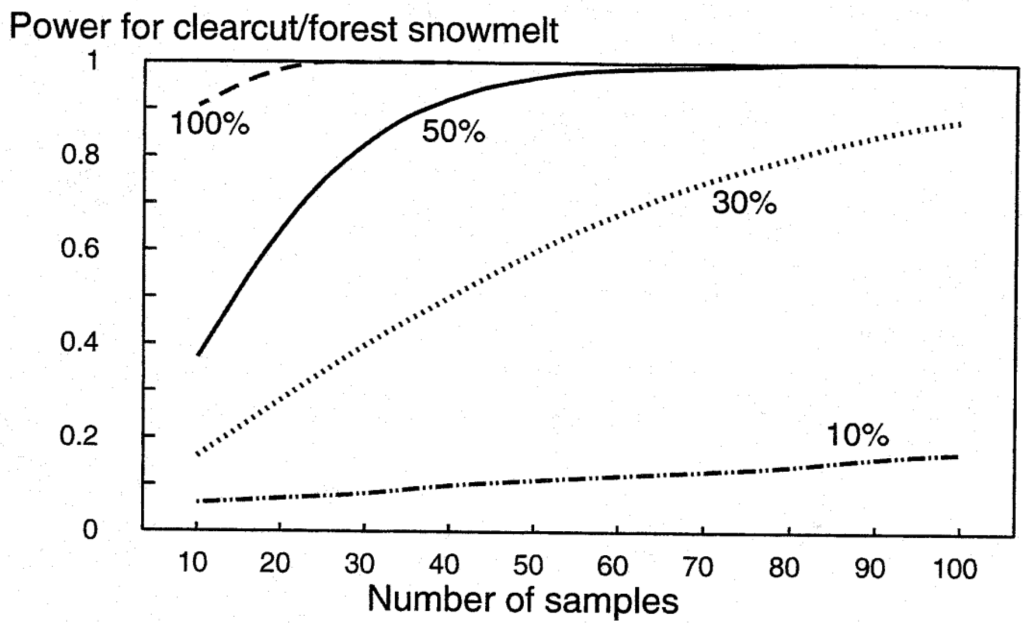

Figure 3. Power analysis for the clearcut/forest comparison of the April 22-28 mean daily snowmelt rate. Lines are for differences of 10, 30, 50 and 100% of the forest value at α=0.05. Data are summarised in Table 4.

Table 4. Power analysis for comparisons of April 22-28 mean daily snowmelt rate between the treatments and the mature forest for sample sizes of 10, 20 and 50, and differences of 10, 30, 50 and 100% of the forest value, i.e., 0.1, 0.3, 0.6 and 1.1 cm d^{-1}.

		Power at differences of			
Sample size	Treatment pair	10%	30%	50%	100%
10	Clearcut/Forest	0.06	0.16	0.37	0.99
10	Thinned/Forest	0.05	0.09	0.18	0.85
10	Unthinned/Forest	0.06	0.11	0.23	0.94
20	Clearcut/Forest	0.07	0.28	0.66	1.0
20	Thinned/Forest	0.06	0.14	0.32	0.99
20	Unthinned/Forest	0.06	0.17	0.42	1.0
50	Clearcut/Forest	0.11	0.60	0.97	1.0
50	Thinned/Forest	0.07	0.29	0.68	1.0
50	Unthinned/Forest	0.08	0.37	0.81	1.0

hydrologic effect, the 10 point snow sampling method is generally inadequate to detect this difference more than 90 times out of 100 at $\alpha = 0.05$. A similar conclusion can be drawn for comparisons between forests and young second growth stands. If sample size is increased to 30, then differences of 4 cm can be detected. The sample requirements for determining differences in the maximum mean rate of snowmelt are similar to those for the SWE. In our case, 10 samples are adequate where differences greater than 1 cm d^{-1} (at $\alpha = 0.05$ and power = 0.9) are considered important. If there is an environmental or economic need to reduce Type I and Type II errors below 0.05 and 0.10, respectively, or the size of the hydrologically significant difference is small, then the number of samples taken must be substantially increased.

ACKNOWLEDGEMENTS

We thank Dr. Val. Lemay, Faculty of Forestry, University of British Columbia, Vancouver, and Vera Sit and Wendy Bergerud, Ministry of Forests, Victoria for statistical advice. Many thanks also go to Laurie Lyons and Ralph Adams for tireless collection of countless snow pack measurements.

REFERENCES

Adams, R.S., D.L. Spittlehouse and R.D. Winkler, 1996. "The effect of a canopy on the snowmelt energy balance", Proceedings 64th. Annual Western Snow Conference, April 15-18, 1996, Bend OR.

Golding, D.L., 1982. "Snow accumulation patterns in openings and adjacent forest", Canadian Hydrology Symposium:82, Associate Committee on Hydrology, National Research Council of Canada, June 14-15, 1982, Fredericton, NB., p.91-112.

Ministry of Environment, 1981. "Snow survey sampling guide", Water Management Branch, Surface Water Section, Victoria, B.C., 28 pp.

Nemec, A.F.L., 1991. "Power analysis handbook for the design and analysis of forestry trials", Biometrics Information Handbook No. 2, Ministry of Forests, Victoria, B.C. 26 pp.

Neter, J., W. Wasserman and M.H. Kutner, 1990. "Applied linear statistical models - regression, analysis of variance, and experimental designs, 3rd. edition", R.D. Irwin Inc., Homewood, IL, 1181 pp.

Peterman, R.M., 1990. "Statistical power analysis can improve fisheries research and management", Canadian Journal of Fisheries and Aquatic Science 47:2-15.

Sanders, W.L., 1989. "Use of PROC GLM of SAS (or a similar linear model computing tool) in research planning", HortScience 24:40-45.

SAS Institute Inc., 1988. "SAS/STAT user's guide", Release 6.03 edition, SAS Institute Inc., Cary, NC.

Toews, D.A.A. and D.R. Gluns, 1986. "Snow accumulation and ablation on adjacent forested and clearcut sites in southeastern British Columbia", Proceedings, 54th. Western Snow Conference, Phoenix, AZ, April 15-17, 1986, pp. 71-80.

Winkler, R.D. and D.L. Spittlehouse, 1995. "The importance of sample size in forest/clearcut snow accumulation comparisons", Mountain Hydrology, Peaks and Valleys in Research and Application, B.T. Guy and J. Barnard (eds.), Canadian Water Resources Association, Cambrigde ON., pp. 39-45.

Wycoff, P., 1957. "Snow surveys from the snow surveyors' side", Proceedings, 8th.Annual Western Snow Conference, pp. 57-59.