RELIABILITY OF RESULTS AND CONCLUSIONS BASED ON STATISTICAL ANALYSIS OF HYDROLOGICAL DATA IN LAND-USE HYDROLOGY

By Anthony A. Salway[1]

INTRODUCTION

The British Columbia Ministries of Forests and Lands, Environment and Parks have identified water as a priority resource in consumptive use watersheds in the Nelson Forest Region. Forest managers require answers to basic hydrological questions. They need to know the extent to which timber harvesting can alter snow accumulation and melt rates in forest openings and how these effects combine with the reduction in evapotranspiration losses to produce water yield increases. The possibility of substantial increases in peak flows that could accelerate erosion and increase sedimentation levels in domestic water supplies, is a primary concern. However, an increase in low flows during late summer months when irrigation demands are high, would be quite beneficial.

The first section of this paper examines some of the relevant literature that addresses the question of increased snow accumulation in cutblocks and longevity of effect. In the second section, the question of water yield increases is probed via a selection of applicable "paired" watershed experiments. Due to space limitations, snow ablation study results and the secondary effects of differing aspects on accumulation and melt rates are not discussed in this paper. A large body of documented evidence leaves little doubt that forest cover removal results in increased snow accumulation in cutblocks (Golding, 1982) and causes water yield increases (Bosch and Hewlett, 1982), but the magnitude of these effects is highly variable. Operational planning of timber harvesting activities in sensitive watersheds would be facilitated by more accurate estimates of anticipated hydrological changes. This paper is concerned with the question of whether better predictions can be made from currently available experimental evidence.

I SNOWPACK WATER-EQUIVALENT EXPERIMENTS

Data Acquisition

Sampling Difficulties. The most common method of snowpack water-equivalent measurement employs "snow tubes". A variety of such snow samplers are used in North America. The most popular model is the standard Federal snow sampler. This is a small diameter cutter (11.2 cm^2). Systematic overmeasurement errors occur with all small diameter cutters (10-11.2 cm^2). The overmeasurement is 10 - 12% for cutters with blunt teeth such as the standard Federal (Western Snow Conference, 1983; p.13). Small diameter samplers can encounter cutter blocking, causing short cores, or snow collapse when ice layers are present. Shallow snowpacks in cold environments develop a bottom layer of depth hoar. Small diameter samplers do not accurately measure shallow snowpacks, with or without depth hoar. Proper snow sampling procedure requires that point measurements be rejected if core lengths are less than 90% of total snow depth. Besides ice layers and depth hoar, brush, slash, logs and boulders can also result in short cores.

Intercorrelated Variables. Snow accumulation on the ground depends upon numerous interrelated factors. In the absence of forest cover, topographic variables such as elevation, aspect and slope shape combine with wind velocity, air temperature and humidity to determine the quantity and distribution of solid precipitation falling on an open site, from each individual storm. The quantity, type and distribution of forest cover, initially determines how much snow reaches the ground at a forested site by intercepting snowfall directly and by modifying wind patterns, thereby indirectly affecting snow distribution. In the absence of forest cover, snow on the ground may be redistributed by wind and/or be lost by melting, evaporation or sublimation. Varying amounts of forest cover affect snow redistribution and sublimation rates by altering wind patterns and modify melt processes by increasing shading and longwave radiation. Snow on tree crowns may undergo sublimation, melting or evaporation. It may drip, slide or be blown off onto the ground over a short or a long period of time. Golding in his review of snow accumulation patterns in openings and adjacent forest (Golding, 1982) points out that the many "confounding factors" severely limit extrapolation of study results from one location to another.

Presented at the Western Snow Conference, April 1987, Vancouver, B.C.

[1] Integrated Hydropedology Ltd., R.R.#1, Winlaw, B.C., VOG 2J0

Experimental Design and Control. The previously discussed sampling difficulties associated with snow water-equivalent measurements can produce large and often systematic errors in the data. It is also extremely difficult to design an experiment that adequately controls all of the many variables, such that a few variables may be studied in isolation. Consequently, snow data often exhibit large dispersions in the statistical sense, thereby increasing the uncertainty of experimental results.

Most snow water-equivalent experimental designs employ a horizontal linear transect across a slope, a portion of which has been logged or is intended to be logged. Holding elevation constant eliminates the effect of this major factor on snow accumulation. Stations are arranged systematically along the transect. The sampling design is therefore a "systematic random" design, "stratified" at a particular elevation (Levin, 1981; pp.246-251).

Case Studies

Priest River, Various Opening Sizes. Haupt (1979a; pp.2-3) selected stations at 7.6, 15.2, 30.5, 121.9 and 304.8 m (25, 50, 100, 400, and 1000 ft) from the east edge of a forested (Western Hemlock, Red Cedar and White Pine) area, roughly along the 1310 m (4300 ft) contour on a north slope and the 1400 m (4600 ft) contour on a south slope in the Priest River Experimental Forest, Northern Idaho. The 7.6 and 30.5 m stations were, however, situated in the centre of 15.2 and 61 m wide clearcut strips, with long axes oriented north-south. The 121.9 and 304.8 m stations were located within two large clearcuts, a 26.3 ha (65 acre) cut on the north slope and a 10.1 ha (25 acre) cut on the south. The study design did not include pretreatment measurements to determine if inherent site differences existed. Two stations were selected 61 m inside the forest, one on each slope, to serve as controls. Two other stations within the forest were also used as controls. The prevailing wind direction was approximately south west. Six point measurements of snow water equivalent (SWE) were made at each station with a Federal snow sampler at estimated peak snow pack. Table 1 summarizes the results obtained by Haupt (1979a) for peak SWE on the north slope. South slope data were incomplete and sample means exhibited no significant differences.

TABLE 1
Peak Snow Water Equivalents for Various Openings and Forested Control Sites, North Slope, Priest River Experimental Forest

	Peak Snow Water Equivalents (cm)						
Year/Offset (m)	I (304.8)	II (121.9)	III (30.5)	IV (15.2)	V (7.6)	VI (0)1	VII (0)2
1969		62.5	58.2	54.9	48.3	43.2	45.7
1970		57.2	52.3	48.3	45.7	39.1	37.3
1971		51.3	49.0	47.0	40.9	34.3	27.4
1972	48.3	51.3	44.2	42.4	41.9	31.2	25.7
1973	47.2	36.8	31.8	29.0	26.9	13.0	11.9
Means	47.7	51.8	47.1	44.3	40.7	32.2	29.6

1 A station 61.0 m west of east edge of forest
2 A station 45.7 m north of north edge of small clearing
(After Haupt (1979a, p.28))

It is impossible to say whether the differences between the SWE values are significant for individual pairs of stations across the table, for any particular year, since sample standard deviations for each value were not quoted. Since six observations were taken at each station, estimates could be obtained. These estimates would be a measure of the individual point sampling errors caused by variations in the snowpack at the station, variations in measurement technique and possibly variations caused by the sampling equipment. Assuming that such errors are small, there does appear to be a trend of increasing SWE as a function of distance from the forest edge for each year.

Analysis of Variance (ANOVA) can be used to determine whether the differences between means occurred merely by chance or not (Levin, 1981; pp.415-424). Since the peak SWE values in Table 1 are dependent (or "paired" in the statistical sense) within rows, e.g., 1973 values are all lower than average, since 1973 was a low snowfall year, the data should be reduced to differences before ANOVA is applied. Table 2 is equivalent to the data subset employed by Haupt (1979a; pp.8-9) in his examination of north slope gains in peak SWE. Data for 1969 were dropped from the analysis to facilitate comparison with south slope date (1970-73). Differences were obtained by subtracting the row averages of columns VI and VII in Table 1 from the values for each offset. Since peak SWE for the 304.8 m offset was measured only in 1972 and 1973, a comparison of 1970-73 data for the 121.9, 30.5, 15.2 and 7.6 m offsets (columns II-V, Table 2) using ANOVA is most appropriate. A value for F of 9.86 was obtained, with 3 and 12 degrees of freedom, indicating that the average difference between the means of the columns (II-V, Table 2) is significant at the 99% level.

TABLE 2
Peak Snow Water Equivalent Differences for Various Openings Compared to Forested Control, North Slope, Priest River Experimental Forest

	Peak Snow Water Equivalent Differences (cm)				
Year/Offset (m)	I (304.8)	II (121.9)	III (30.5)	IV (15.2)	V (7.6)
1970		18.9	14.0	10.0	7.4
1971		20.3	18.0	16.0	9.9
1972	19.8	22.9	15.7	14.0	13.5
1973	34.8	24.4	19.3	16.5	14.5
Mean ($\bar{x}$)	27.3	21.6	16.8	14.1	11.3
Standard Deviation (s)	10.6	2.48	2.36	2.95	3.28

It would now be useful to test selected pairs of means for significant difference (Levin, 1981; pp.366-370). The results of these tests on the data in Table 2 are as follows. Due to its small sample size (n=2) and the disparity between the two differences for the 304.8 m offset, the mean difference of 27.3 is not significantly different from any of the other differences. Assigning 95% confidence limits to this difference indicates that it is not even significantly different from zero, i.e. 27.3 ± 95.3. The other differences are all significantly different from each other at the 95% level, except the adjacent 30.5/15.2 m and 15.2/7.6 m pairs.

In their study of snow accumulation in small forest openings in Alberta, Golding and Swanson (1978; p.383-4) similarly found insignificant differences between peak SWE differences for adjacent opening sizes, e.g. 3/4H versus 1H (similar to Haupt's 15.2/7.6 m pair) and 1H versus 2H versus 3H (similar to Haupt's 30.5/15.2 m pair). Their study was conducted over a short period, 1973-76, as was Haupt's study but their snowpacks were quite shallow (<14 cm for peak SWE over the 4 year period).

It is interesting to note from Table 2 that the dispersion of the data increases with decreasing offset, i.e. there is greater variability in peak SWE measurements close to the forest edge, as might be expected. Of course, a larger sample would produce more reliable conclusions. A good estimate of a population mean requires more than thirty observations (Sanders et al, 1976; p.149).

Haupt's conclusion that the north aspect showed an exponential increase in SWE with an expansion of the size of the clearing (Haupt, 1979a; p.9), must be viewed with some degree of skepticism in light of the foregoing discussion. Haupt's somewhat speculative final conclusion (Haupt, 1979a; p.24) that creating larger openings on north slopes "will develop a high potential for increasing total water yield" which, in combination with "higher melt rates ... can cause higher spring peak flows for many years after logging, thus imperiling water quality" is not well supported by his data and results.

Priest River, Strip Cut. Haupt (1979b) carried out another study of snow accumulation on a 100.6 m (330 ft) wide north-south oriented strip clearcut in Benton creek watershed (Priest River Experimental Forest), comparing SWE in the clearcut with that in the adjacent forest over a period of 33 years from 1942-74. The strip was harvested in 1940. After 20 years, natural regeneration had resulted in about 60% Western Red Cedar and 20% Western Hemlock some 3 to 4.5 m high. In 1966, the strip was thinned from about 8,150 to 990 stems per hectare. Point measurements of SWE were taken along horizontal transects at an elevation of approximately 1,065 m. Sample sizes ranged from 3 to 9 observations in the clearcut and 2 to 6 observations in the forest, spaced at 6 m (20 ft) to 15 m (50 ft) apart (Haupt, 1979b; p.4). As before (Haupt, 1979a) point sampling standard deviations were not available. Estimated peak SWE was obtained for each of the 13 years by averaging sample observations. Table 3 contains the results of this study for the north slope.

TABLE 3
Peak Snow Water Equivalents for 100 m wide Strip Cut and Forested Control Sites, North Slope, Priest River Experimental Forest

Year	Time Index	Peak SWE (cm)		Percent Difference
		Recovery Plot	Forest Plot	
1942	2	27.4	8.9	208
1949	9	56.1	29.2	92
1950	10	49.5	31.0	60
1951	11	33.5	16.8	99
1952	12	46.7	31.0	51
1967	27	33.0	11.2	195
1968	28	19.8	3.6	450
1969	29	45.0	37.1	21
1970	30	39.4	28.4	39
1971	31	31.8	16.3	95
1972	32	18.8	5.1	269
1973	33	13.2	2.8	371
1974	34	35.6	22.9	55
Means		34.6	18.8	

After Haupt (1979b)

The mean value of the differences between the paired forest and recovery plot values is 15.8 ± 3.0 cm of SWE at the 95% level. It is worthwhile, however, in this case, to regress SWE in the strip cut (y) upon SWE in the forest (x), see Haupt (1979b; p.9). Haupt produced the following two simple linear regressions for 1942-52:

$$\hat{y} = 17.0 + 1.10\,x; \quad R^2 = 0.867 \tag{1}$$

and for 1967-74:

$$\hat{y} = 16.4 + 0.828\,x; \quad R^2 = 0.888. \tag{2}$$

He claimed that the "level" of the 1967-74 regression was significantly below that of the 1942-52 regression, thereby indicating a "time-dependent decrease in snow W.E. gains" (Haupt, 1979b; p.8), despite the thinning in 1966. A good test of the difference between simple linear regressions has been described by Draper and Smith (1966, pp.67-69) and has been applied successfully by Harr and McCorison (1979) in their study of clearcut logging on peak flows in Western Oregon. The procedure uses the "extra sum of squares" principle. An F-value is calculated from:

$$F = ((SS_\omega - SS_\Omega)/2)/(SS_\Omega/(n_1+n_2-4)) \tag{3}$$

Where:

SS_ω is the residual of the overall regression,
SS_Ω is the sum of the residuals of the two separate regressions,
n_1 is the number of cases for the first regression,
n_2 is the number of cases for the second regression.

This value is then compared with F tab $(2,n_1+n_2-4)$ at the appropriate level of significance. For Haupt's data, the two regressions were not even significantly different at the 95% probability level. An overall regression for 1942-74:

$$\hat{y} = 16.0 + 0.988\ x; \quad R^2 = 0.845 \qquad (4)$$

can be fitted to the data, instead of the two separate regressions. This regression approximately reduces to:

$$(\hat{y}-x)/x = 16.0/x \qquad (5)$$

Hence Haupt's data shows that percent increase in SWE in the clearing relative to SWE in the forest is approximately inversely proportional to SWE in the forest. The constant of proportionality is 16.0.

Since there does appear to be a "time-dependent decrease in SWE gains" (Haupt, 1979b; p.8), it may be worthwhile to include a linear time index in the overall regression model. Re-analysing the data using a multi-linear regression approach gives:

$$\hat{y} = 23.8 + 0.920\ x - 0.291\ T; \quad R^2 = 0.908 \qquad (6)$$

where T = number of years since harvest. The T-term is just significant at the 95% level (partial F-value is 6.94 compared to F tab (1,10,0.95) of 4.96).

Substituting the average value for x of 18.8 cm (see Table 3) for both x and y into equation (6) and calculating T, provides an estimate for total time to complete recovery of 77 years. This period may be somewhat less, since the estimate is based on a linear fit to initial recovery data, during which growth is non-linear. Initial recovery was also interrupted by the thinning in 1966.

In summary, the data in Table 3 can be described by an overall linear regression, equation (4) or (6). Both regressions would be "satisfactory predictors" (Draper and Smith, 1966; p.64) since their F-values are at least 4 times tabulated F-values. Both equations imply that percent increase in SWE in the clearcut relative to the forest is inversely related to SWE in the forest (i.e. the amount of snow that falls per season). For Haupt's data (Table 3) percentages range from a low of 21% in 1969, a light snowfall year, to a high of 450% in 1968, a heavy snowfall year. Haupt's conclusion (Haupt, 1979b; p.13) that "the immediate response to clearcutting was a 56% increase in peak snow W.E." and that, "Thirty four years after clearcutting ... overall increase remained a substantial 37 percent" is a highly specific conclusion, depending on the particular amounts of snow that happened to fall during the period of study. Extrapolation of these percentages to other time periods and/or sites is not possible.

The T-term in equation (6) is just significant and implies that complete recovery could be expected in about 77 years, or less for that site.

Fraser Study. Troendle and Meiman (1984) reported results of a study originally described by Wilm and Dunford (1948). A 3 ha clearcut was created in a stand of lodgepole pine in the Fraser Experimental forest, Colorado in 1940. Peak SWE was measured at 25 sample points in the clearcut and forested control for 9 of the years within the period 1941-81, as the site regenerated to a 10-15 m high lodgepole stand. A few overmature spruce, fir and lodgepole pine had been left from the original harvest. In the summer of 1981, the 3 ha clearcut was reharvested along with an 8 ha area surrounding the clearcut, thereby creating a 22H wide opening, oriented east-west in the direction of the prevailing wind. After harvesting, the entire area of the clearcut was covered with a mat of slash, up to 60 cm high. Peak SWE was measured at the 25 sample points in the original clearcut block and forested control. Once again, sample standard deviations for individual stations and each year were unavailable and therefore assumed to be small compared with variations caused by climatic and site differences. The results are summarized in Table 4.

Unlike the previous studies (Haupt, 1979a and 1979b), this study included two sets of pretreatment measurements for 1938 and 1939 to determine if inherent site differences existed. Although the samples are very small (n=2), the values are so close that it is reasonable to infer that there is no inherent site difference. The mean value of the difference between the paired forest and clearcut values for 1941-81 is 5.7 ± 1.6 cm of SWE at the 95% level.

Regressing SWE in the clearcut (y) upon SWE in the forest (x) for 1941-81 and including a linear time index results in:

$$\hat{y} = 5.51 + 1.13\ x - 0.099\ T; \quad R^2 = 0.987 \qquad (7)$$

in which the T-term is quite significant at the 95% level (partial F-value is 17.08 compared to F tab (1,6,0.95) of 5.99). Substituting the average value for x of 16.8 cm (for 1941-81) for both x and y into equation (7) and calculating T, provides an estimate for total time to complete recovery of 77 years. This figure happens to be the same as that obtained for Haupt's data. Troendle and Meiman (1984, p.89) obtained a figure of 103 years using a similar regression but this was based on the period 1941-83, that included the two years following reharvest in 1981.

TABLE 4
Peak Snow Water Equivalent for 3 ha Clearcut and Forested Control Sites, Fraser Experimental Forest

Year	Time Index	Peak SWE (cm) Clearcut	Forest	Percent Difference
1938		18.0	18.0	0
1939		16.8	17.3	-3
Harvest				
1941	1	21.1	14.7	44
1942	2	23.1	15.7	47
1943	3	33.3	23.9	39
1956	16	31.5	24.9	26
1964	24	19.3	13.2	46
1968	28	20.1	16.3	23
1972	32	25.4	20.1	26
1976	36	19.3	15.7	23
1981	41	9.4	6.8	38
Harvest				
1982		25.4	19.8	28
1983		29.5	23.4	26

After Troendle and Meiman (1984)

In summary, Troendle and Meiman (1984, pp.88-90) obtained results that are quite similar to those obtained by Haupt (1979b; pp.6-9) for peak SWE in a clearcut versus the forest (equations (6) and (7)). Times to complete recovery were estimated from the two regressions at approximately 77 years.

It is also worth noting from Troendle and Meiman's study that, although the slash (up to 60 cm in depth) was effective at trapping most of the 25.4 and 29.5 cm of SWE that fell in the centre of the newly created 22H clearcut in 1982 and 1983, thereby reducing scour (Troendel and Meiman, 1984; p.96), the 5.6 and 6.1 cm increases, relative to the forest do not lend support to the snow retention function graph (Troendle and Meiman, 1984; p.88).

<u>West Kootenay Synoptic Study</u>. Toews and Gluns (1986) carried out a synoptic study of snow accumulation and ablation for a variety of paired forested and clearcut sites in south-eastern British Columbia over a three year period, 1983-85. SWE was measured along horizontal transects, at right angles to cutblock boundaries. Ten point measurements were made within the forest and ten within the clearcut, at 10 m spacings along each transect. "No measurements were taken within 50 m of the forest/clearcut boundary to avoid any possible edge effect that may arise from wind or the effects of differential melt caused by back radiation", (Toews and Gluns, 1986; p.102). As with Haupt's studies (Haupt 1979a,

1979b), Toews and Gluns' study design did not include pretreatment measurements to determine if inherent site differences existed.

After averaging each set of ten point measurements in the forest and clearcut for a particular transect, Toews and Gluns (1986) calculated percent differences. The results are summarized in Table 5.

TABLE 5
Percent Differences in Peak Snow Water Equivalent between Forested and Clearcut Sites in the West Kootenays

Year	Number of Sites	Percent Difference	Range of Difference
1983	27	35.3	10.7 - 68.0
1984	33	42.0	8.8 - 118.4
1985	29	33.4	3.7 - 88.9
	Average:	36.9	

After Toews and Gluns (1986)

Re-examination of Toews and Gluns' 1983 data, revealed that for 25 out of the 28 sites, differences between the forest and clearcut mean values of peak SWE were significant at the 95% level. After eliminating the worst sample pair, with the lowest level of significant difference, as Toews and Gluns did, due to its unreliability, a mean percent difference of 35.3 with a sample standard deviation of 16.4 was calculated for the 1983 data. Hence, at the 95% confidence level, $\mu = 35.3 \pm 6.5$.

Regressing peak SWE in the clearcut (y) against peak SWE in the forest (x) for all of the 1983 data results in the following regression:

$$\hat{y} = 18.4 + 0.976x; \quad R^2 = 0.802 \qquad (8)$$

which is remarkably similar to that obtained with Haupt's data (equation (4)). The F-value for this equation is 105.3 with 1 and 26 degrees of freedom, hence this would be a "good predictor" (Draper and Smith, 1966; p.64).

In summary, Toews and Gluns (1986) investigated peak SWE differences between forested and clearcut sites for a single effective average offset of 100 m or 4H on each side of the forest boundary. They concluded that "The data collected in this study showed no indication of decreasing accumulation with increased clearcut size in that the accumulations in the openings were in all cases greater than those in adjacent forested sites" (Toews and Gluns, 1986; p.109). Although opening diameters in their study varied between 6H and 42H, it should perhaps be mentioned that no measurements were made further out than 6H into the clearcut. In their summary, Toews and Gluns (1986; p.110) concluded that "Snow accumulation as measured on 60 sites over a 3-year period averaged 37% greater on clearcut sites than forested sites". They could also perhaps have noted that percent differences from about 4 to 118 were obtained (see Table 5), depending to a large extent on the amount of snow that fell at each pair of sites during the winter. Equation (8) for 1983 indicates that the percent difference is approximately inversely proportional to peak SWE in the forest, the constant of proportionality being 18.4. It would be worthwhile to develop similar equations for their entire data set.

II STREAMFLOW EXPERIMENTS

Data Acquisition

Sampling Difficulties. Stream discharge may be measured directly by metering using instruments such as the mechanical Price-Gurney flow meter or the electro-magnetic Marsh-McBirney. Discharge may also be monitored by observing stage heights and relating them to a previously established stage-discharge curve. Control structures such as weirs (broad or sharp-crested, rectangular, V-notched, trapezoidal, etc.), dams or flumes are generally required for accurate measurements. Formulae have been developed for simple geometric structures and these formulae may also be used in conjunction with stage height to estimate discharge. More sophisticated installations employ water level recorders for the estimation of stage height.

Many problems are associated with stream discharge measurement, not least of which is cost. The cost of installation, maintenance and observer's wages limit the number of creeks in a particular area that can be monitored on a regular basis. Winter observations are particularly difficult, when icing is a factor and heated installations are often required. Many snow melt dominated creeks exhibit peak flows in the spring that may be several orders of magnitude greater than late summer flows. Designing structures that measure accurately over such large ranges of flows is difficult. Instantaneous peak discharges that may only last for a few hours are often missed by manual observers in situations where water level recorders are not used.

Intercorrelated Variables. Snow melt dominated mountain creeks are often uniquely different in terms of response. At one extreme, flows are derived from the rapid drainage, after snowmelt, of soil water moving through shallow coarse-textured sandy material on steep slopes. Such creeks are quite "flashy". Other flows are derived from the slow drainage of less permeable, finer-textured material, that may be deeper, on possibly less steep slopes. These creeks rise more slowly and peak later than the more "flashy" creeks. Yet other flows are derived from deeper ground water storage and may exhibit little variation during peak snowmelt. Surface drainage divides may not coincide with subsurface divides, such that water may be transferred between adjacent drainages. A creek may gain or lose, rise or sink down a particular reach at certain times of year. Selection of a suitable reach for monitoring purposes is often critical. Besides these regime differences caused by geological and terrain factors, there are many others that influence streamflow. Drainage area, shape, orientation, range of aspects, elevation range, vegetative cover and climatic variables all influence streamflow to varying degrees.

Experimental Design and Control. The "paired watershed approach" is the only reliable way to isolate the effects of vegetation removal from other factors. The method firstly involves the selection of a similar watershed to the one scheduled for treatment. Streamflows are compared during a "calibration" period, prior to treatment and usually a simple linear regression relationship is established. After treatment, streamflows are again compared to see whether any change in their relationship has occurred.

Case Studies

Salmon Arm Study. Cheng (1980) studied the hydrological effects of the Eden Forest Fire near Salmon Arm, B.C., by comparing streamflow data for Palmer Creek watershed, that suffered a 60% burn in 1973, with data from the Upper Salmon River. Table 6 summarizes the basic characteristics of the two watersheds.

TABLE 6

Physiographic and Biogeoclimatic Characteristics of the Study Watersheds

	Palmer	Upper Salmon
Area (km^2)	18.1	143.2
Elevation (m)--median	1,465	1,455
Elevation (m)--range	425 - 1,735	1,005 - 2,040
Slope Shape	convex	concave
Dominant Aspect	east	west
Soils	varied	varied
Parent Material	granitic	mainly volcanic
Forest Types	mixed	mixed
Mean Annual Precipitation (cm)	127	76
Mean Annual Runoff (cm)	36	17
% of April-August to Annual Runoff	81.6	84.0

After Cheng (1980)

According to Cheng (1980; p.240), the Upper Salmon River was the only other stream with concurrent streamflow data in the nearby area. Unfortunately, the two watersheds may not be sufficiently similar for reliable comparisons of flow data to be made using the "paired watershed approach". Cheng (1980; pp.242-243) did, however, attempt a comparison of seasonal water yields (April - August) for a pre-burn (1968-73) and post-burn (1974-77) period. The pre-burn regression for 1968-73 (omitting the 1972 data point due to its unreliability) was:

$$\hat{y} = 16.9 + 0.283\,x; \quad R^2 = 0.928 \qquad (9)$$

where $\hat{y}$ is the predicted mean daily flow (1/s) for Palmer Creek and x is mean daily flow (1/s) for the Upper Salmon River.

This regression is just significant at the 99% level (F-value is 38.7 compared with F tab (1,3,0.99) of 34.1). But, equation (9) cannot be regarded as a satisfactory prediction tool (Draper and Smith, 1966; p.64). Cheng (1980) constructed 95% confidence limits around the regression, i.e.:

$$\hat{y}_k \pm t(n-2,0.95)\; s(1/n + (x_k - \bar{x})^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2)^{1/2} \qquad (10)$$

(see Draper and Smith, 1966; pp.21-24), and obtained the result that only the seasonal water yield increases for 1975 and 1976 were significant at the 95% level. Actually, the limits indicate that the predicted values for y_k were 493 ± 79 and 453 ± 68 for 1975 and 1976, respectively. However, the estimated variance of a predicted value of an individual observation is more correctly given by:

$$s^2 + V(\hat{y}_k) = s^2\,(1 + 1/n + (x_k - \bar{x})^2 / \sum_{i=1}^{n} (x_i - \bar{x})^2) \qquad (11)$$

(see Draper and Smith, 1966; p.24). Hence, the predicted values for y_k were 493 ± 156 and 453 ± 150 for 1975 and 1976 respectively. Therefore 1976 was, in fact, the only year of the four post-burn years with a significant increase at the 95% level, see Table 7.

TABLE 7
Snowmelt Season Water Yield Changes in Palmer Creek Following the Eden Forest Fire

Year	y (1/s)	$\hat{y}$ (1/s)	$y - \hat{y}$ (1/s)	$(y - \hat{y})/\hat{y}$ (percent)
1974	725	668	57	8.5
1975	566	493	74	15.0
1976	782	453	329	72.5
1977	229	235	6	-2.4
Mean	575	462	113	24.5

After Cheng (1980)

It would be unwise to infer from this one point that the post-burn data show a significant seasonal water yield increase. The 95% limits mean that there is a 1 in 20 chance that a single point may lie outside them by pure chance. A second regression should be fitted to the post-burn data and tested for significant difference using the "extra sum of squares" approach as previously discussed. If this is done, the post-burn regression is not significant at the 95% level and there is certainly no significant difference between the two regressions.

Finally, averaging the percentage yield increases over the 4-year post-burn period (see Table 7) to obtain a figure of 24% is certainly not worthwhile, since 95% confidence limits are ± 51%, i.e. there is no significant difference between the actual and predicted means. Cheng (1980) also examined monthly water yields using the same analytical approach. Only the June calibration regression could be regarded as a satisfactory predictor (F-value of 246 compared with F tab (1,4,0.99) of 21.2). The calibration regressions for September and October were only just significant at the 99% level, those for April and August were just significant at the 95% level, May and November were barely significant at the 95% level and July was insignificant.

The constructed 95% confidence limits indicate that only a few post-burn points for each month show significant increases in water yield. Some, in fact, even exhibit significant decreases, which are difficult to rationalize. As Cheng (1980; p.244) quite rightly points out, these anomalous results could be due in part to the fact that for some years, the season may be more advanced or less advanced than normal for any one selected month. Analysis of monthly data on an arbitrary monthly basis, is therefore quite risky.

In summary, major differences in the basic characteristics of the treatment and control watersheds, coupled with extremely small data bases, produced statistical results that exhibit large dispersions and a general lack of significance, as Cheng (1980) has recognized in his summary.

Cheng's (1980; p.251) conclusions that the 1973 Eden Forest Fire caused "higher and earlier annual peak flows, the advancement in time of the major snowmelt runoff volume, increases in total April - August runoff volume and in monthly water yields during the August to November period", although apparently logical and in good agreement with results from studies elsewhere are not well supported by his own study results.

Carnation Creek Study. Hetherington (1982) compared two watersheds before and after timber harvesting with two controls, all within the main Carnation Creek watershed. Table 8 summarizes some of the basic characteristics of the four watersheds.

TABLE 8
Some Basic Physiographic Characteristics of the Study Watersheds

	Treatment Watersheds		Control Watersheds	
	B	H	C	E
Area (ha)	930	12	145	264
Elevation Range (m)	8 - 884	152 - 305	46 - 700	150 - 884
Dominant Aspect	west	east	north-west	west

After Hetherington (1982)

This multiple paired watershed study departs somewhat from the "norm" since, treatment watershed B actually includes treatment watershed H and the two controls C and E. Statistical comparison of B watershed with C or E watersheds, that are part of B watershed, is therefore not merely a simple comparison. Furthermore, B watershed was progressively logged from 1976 (1%) to 1981 (40%) (Hetherington, 1982; p.48), conflicting with the defined post-logging period (1978-81). H watershed was 90% logged in 1977-78 but it is somewhat different physiographically from the two controls as indicated in Table 8. Hetherington (1982) conducted a similar type of analysis to that of Cheng (1980) comparing pre-logging (1974-77) calibration regressions of seasonal water yield with individual post-logging (1978-81) points (Hetherington, 1982; p.50). Calibration regressions were all based on four pre-logging years and were significant at the 95% but not at the 99% level. Post-logging points all "fell within broad 95% confidence bands for all pre-logging regressions" (Hetherington, 1982; p.52), some above and some below the regression lines. Hetherington (1982; pp.58-59) also compared minimum daily summer flows and the number of low flow days. For H watershed, he reported an average increase of "at least 78% in 1978 and 1979" compared to control C, but "no change in 1980". For B watershed, there was an average decrease of "at least 47% in 3 of the 5 post-logging years" (1977, 79 and 81) but "little or no change in 1978 and 1980". Decreases in the number of low flow days were generally not statistically significant.

In summary, Hetherington's (1982) preliminary look at logging effects on the hydrologic regime of Carnation Creek suffers from similar problems to those encountered by Cheng (1980). Treatment watersheds were either an integral part of a control or not sufficiently similar to the control and periods of record were too short for meaningful statistical comparison. Spurious results were again obtained for seasonal water yield changes and changes in low flows. In consideration of the obvious limitations in the data and analysis, conclusions drawn from this study (Hetherington, 1982; pp.58-60) may be somewhat premature, as the author has indicated.

Hinton-Edson Synoptic Study. Swanson and Hillman (1977) conducted a synoptic study of 9 logged and 9 unlogged catchments in the Hinton-Edson area of Alberta, using 1974 hydrographic data. The 9 logged watersheds had percent cuts ranging from 35% to 84% in 1974 and the 9 unlogged controls ranged from zero to 21%. Drainage areas ranged from 7.0 to 23.9 km^2, elevation ranges were quite varied but dominant forest cover types were mainly lodgepole pine with some white spruce, (Swanson and Hillman, 1977; p.21). Data were converted to unit area form to allow for differences in area between catchments.

Unpaired comparisons of the means of logged versus unlogged runoff amounts for the spring freshet, one one-week recession period in June and three rainstorms in July, August and September were made. (Swanson and Hillman, 1977; p.26). Only the September difference was significant at the 95% level. Adjacent logged and unlogged pairs were also subjected to the same statistical comparison. The results were somewhat better with the differences for all periods slightly above the 95% level, except for the June recession. The overall difference for the entire gauged season (April 25 - September 15) was also just significant at the 95% level for the paired comparison. Individual hydrograph pairs display considerable variation in flow differences due to logging and in one case, Fox Creek with a 57% cut, exhibited consistently lower flows, than nearby Hendrickson Creek, which was unlogged (Swanson and Hillman, 1977; p.29). The authors felt that a design probability level of 80% was "acceptable for this evaluation test and for later use in making land-management decisions affecting water." (Swanson and Hillman, 1977; p.19). Most of the comparisons did, in fact, satisfy this modest statistical criterion.

In summary, this departure from the conventional "paired watershed approach", that assumes equal water yield (on a unit area basis) from both catchment types, prior to logging i.e., that there were no inherent site differences, is quite risky as Swanson and Hillman (1977; p.31) have pointed out. Absence of any pre-logging calibration data is a serious drawback in this type of analysis. Although the average yield increase was around 27% for unpaired logged versus unlogged catchments (April 25 - September 15) actual "results ranged from a slightly negative effect on one of the Berland catchments to a 100% increase on the McLeod catchments" (Swanson and Hillman, 1977; p.37).

CONCLUSIONS AND RECOMMENDATIONS

Although peak SWE has been shown to be greater in clearcuts than the surrounding forests for offsets up to 5H (Haupt, 1979a), for a 1H offset in a 5H strip (Haupt, 1979b), in the centre of a 6H opening (Troendle and Meiman, 1984) and for an average effective offset of 4H (Toews and Gluns, 1986), results are highly variable. Percent increases relative to the forest range from about zero to several hundred. This enormous variation in percentages does not appear to be a reflection of site or climatic regime differences within and between studies but rather a reflection of the way in which past data have heretofore been analysed. Equations (4) and (8) indicate that percent increases are in fact, approximately inversely related to peak SWE, the constants of proportionality being a measure of the integrated effect of the forest cover in preventing snow from reaching the ground at the forested sites. Equations (6) and (7) show that complete hydrological recovery in terms of the snow interception/redistribution effect can be expected in 77 years or less. There is a tenuous indication that the centre of a newly created 22H clearcut exhibited an increase in peak SWE over a two year period but this was complicated by low snowpacks that barely covered the 60 cm mat of slash (Troendle and Meiman, 1984). The generally accepted snow retention function graph (Troendle and Meiman, 1984; p.88) implies that large clearcuts should have less snow than the surrounding forest. There is a need for controlled studies of snowpack accumulation in large openings (>6H), that compare measurements taken right across the clearcuts with those within the forest.

As generally recognized by the authors (Cheng (1980), Hetherington (1982) and Swanson and Hillman (1977)), the three "paired" watershed studies discussed in this paper suffered from a lack of sufficiently similar watershed pairs and pre-logging calibration and post-logging periods of record that were too short for reliable, statistical comparisons. Data exhibited large dispersions and in all cases a large range of results was obtained. Seasonal yield increases were highly variable (from about zero to one hundred percent) and monthly yields were either increased or decreased. Many of the observed changes were statistically insignificant. "Paired" watershed studies should probably not be conducted unless the control and treatment watersheds are sufficiently similar. Periods of record need to be longer, if classical statistical methods of comparison are used. Box and Jenkins (1970), and Chow (1964; 8, pp.78-89) have discussed ways in which short periods of record can be better utilized by producing more powerful models based on auto- and cross-correlations within and between time series, rather than trying to eliminate serial correlations by only using seasonal or monthly averages, thereby ignoring the fine structures of these series.

REFERENCES

Bosch J.M. and Hewlett J.D., 1982. A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. Journal of Hydrology, Vol.55, pp.3-23.

Box, G.E.P. and Jenkins, G.M., 1970. Time series analysis, forecasting and control. San Francisco, Holden-Day.

Cheng, J.D., 1980. Hydrologic effects of a severe forest fire. In proceedings of watershed management symposium, Amer.Soc. of Civil engineers, Idaho, Vol.1, pp.240-251.

Chow, V.T., 1964. Handbook of Applied Hydrology. McGraw-Hill.

Draper, N.R. and Smith, H., 1966. Applied regression analysis. Wiley and sons, New York.

Golding, D.L., 1982. Snow accumulation patterns in openings and adjacent forest. Proc.Can. Hydrology Symposium, Assoc.Comm. on Hydrology, Nat.Res. Council, Can. pp.91-112.

Golding, D.L. and Swanson, R.H., 1978. Snow accumulation and melt in small forest openings in Alberta. Can.Journ.For.Res. 8:380-388.

Harr, R.D. and McCorison, F.M., 1979. Initial effects of clearcut logging on size and timing of peak flows in a small watershed in Western Oregon. Am.Geophys. Union, Vol.15, No.1; pp.90-94.

Haupt, H.F., 1979a. Local climatic and hydrologic consequences of creating openings in climax timber of North Idaho. USDA For.Ser.Res. Paper INT-223.

Haupt, H.F., 1979b. Effects of timber cutting and revegetation on snow accumulation and melt in North Idaho. USDA For.Serv.Res. Paper INT-224.

Hetherington, E.D., 1982. A first look at logging effects on the hydrologic regime of Carnation Creek experimental watershed. Proc. Carnation Creek Workshop, Malaspina College, Nanaimo, B.C., pp.45-63.

Levin, R.I., 1981. Statistics for Management. Prentice Hall.

Sanders, D.H., Murph, A.F. and Eng, R.J., 1976. Statistics. McGraw-Hill.

Swanson, R.H. and Hillman, G.R., 1977. Predicted increased water yield after clearcutting verified in west-central Alberta. North.For.Res. Centre, Alberta. NOR-X-198.

Toews, D.A.A. and Gluns, D.R., 1986. Snow accumulation and ablation on adjacent forested and clearcut sites in southeastern B. C.. Proc. Western Snow Conf., pp.101-111.

Troendle, C.A. and Meiman, J.R., 1984. Options for harvesting timber to control snowpack accumulation. Proc. Western Snow Conf., pp.86-97.

Western Snow Conference, 1983. Metrication of manual snow sampling equipment. Western Snow Conference, Spokane, WA.

Wilm, H.G. and Dunford, E.G., 1948. Effect of timber cutting on water available for streamflow from a lodgepole pine forest. USDA Technical Bulletin 968.