Impacts of Logging on Storm Peak Flows, Flow Volumes and Suspended Sediment Loads in Caspar Creek, California

Jack Lewis, Sylvia R. Mori, Elizabeth T. Keppeler, Robert R. Ziemer

ABSTRACT

Models are fit to 11 years of storm peak flows, flow volumes, and suspended sediment loads on a network of 14 stream gaging stations in the North Fork Caspar Creek, a 473-ha coastal watershed bearing a second-growth forest of redwood and Douglas-fir. For the first 4 years of monitoring, the watershed was in a relatively undisturbed state, having last been logged prior to 1904, with only a county road traversing the ridgetops. Nearly half the watershed was clear-cut over a period of 3 years, and yarded primarily using uphill skyline cable systems to spur roads constructed high on the slopes. Three tributaries were maintained as controls and left undisturbed. Four years of data were collected after logging was completed. Exploratory analysis and model fitting permit characterization and quantification of the effects of watershed disturbances, watershed area, antecedent wetness, and time since disturbance on storm runoff and suspended sediment. Model interpretations provide insight into the nature of certain types of cumulative watershed effects.

INTRODUCTION

This paired-watershed study in the North Fork of Caspar Creek was motivated by a desire to understand how a particular logging system affects storm peak flows, flow volumes, and suspended sediment loads in a second-growth coastal redwood forest. The logging system consisted of clear-cutting with streamside buffers, and yarding primarily by skyline to spur roads located on upper slopes and ridges. Primary objectives were to quantify how impacts vary with different levels of disturbance and how the effects of a given disturbance vary downstream. Pursuant to these objectives, a statistical model was developed for a treatment-and-control experimental design involving multiple watersheds. The study was also an opportunity for testing new technologies, and demonstrates two new automated schemes for suspended sediment sampling. Techniques for estimating sediment loads from these samples are tested and applied.

Storm Peaks

Throughout much of the Pacific Northwest, a large soil moisture deficit develops during the dry summer. With the onset of the rainy season in the fall, the dry soil profile begins to be recharged with moisture. In the H.J. Andrews Experimental Forest in the Oregon Cascades, the first storms of the fall produced streamflow peaks from a 96-ha clear-cut watershed that ranged from 40% to 200% larger than those predicted from the pre-logging relationship [Rothacher, 1971; 1973]. In the Alsea watershed near the Oregon coast, Harris [1977] found no significant change in the mean peak flow after clear-cutting a 71-ha watershed or patch cutting 25% of an adjacent 303-ha watershed. However, when Harr [1976] added an additional 30 smaller early winter runoff events to the data, average fall peak flow was increased 122%. In Caspar Creek, Ziemer [1981] reported that selection cutting and tractor yarding of an 85-year-old second-growth redwood and Douglas-fir forest increased the first streamflow peaks in the fall about

300% after logging. The effect of logging on peak flow at Caspar Creek was best predicted by the percentage of area logged divided by the sequential storm number, beginning with the first storm in the fall. These first rains and consequent streamflow in the fall are usually small and geomorphically inconsequential in the Pacific Northwest. The large peak flows, which tend to modify stream channels and transport most of the sediment, usually occur during mid-winter after the soil moisture deficits have been satisfied in both the logged and unlogged watersheds.

Studies of large peak flows in the Pacific Northwest have not detected significant changes after logging. Rothacher [1971, 1973] found no appreciable increase in peak flows for the largest floods attributable to clear-cutting. Paired watershed studies in the Oregon Cascades [Harr et al., 1979], Oregon Coast Range [Harr et al., 1975; Harr, 1976; Harris, 1977], and at Caspar Creek [Ziemer, 1981; Wright et al., 1990] similarly suggested that logging did not significantly increase the size of the largest peak flows that occurred when the ground was saturated.

Using longer streamflow records of 34 to 55 years, Jones and Grant [1996] evaluated changes in peak flow from timber harvest and road building from a set of three small basins (0.6 to 1 km^2) and three pairs of large basins (60 to 600 km^2) in the Oregon Cascades. In the small basins, they reported that changes in small peak flows were greater than changes in large flows. In their category of "large" peaks (recurrence interval greater than 0.4 years), flows were significantly increased in one of the two treated small basins, but the 10 *largest* flows were apparently unaffected by treatment. Jones and Grant [1996] reported that forest harvesting increased peak discharges by as much as 100% in the large basins over the past 50 years, but they did not discuss whether the largest peak flows in the large basins were significantly affected by land management activities. Two subsequent analyses of the same data used by Jones and Grant concluded that a relationship could not be found between forest harvesting and peak discharge in the large basins [Beschta et al., 1997; Thomas and Megahan, 1998].

There are several explanations why relationships between land management activities and a change in storm peaks have been difficult to document. First, the land management activity may actually have no effect on the size of storm peaks. Second, because major storms are infrequent, the range of observations may not adequately cover the range of interest. Third, if the variability in response is large relative to the magnitude of change, it may be difficult to detect an effect without a large number of observations. Fourth, land-use changes in a large watershed are often gradual, occurring over several years or decades. The use of an untreated control watershed whose flows are well-correlated with the treated watershed can greatly increase statistical power, if both watersheds are monitored for an adequate number of years before and after the treatment is applied. The variability about the relation between the two watersheds can be critical. For example, when the South Fork (pre-treatment RMSE = 0.232) was used as the control, no change in peak streamflow was detected at the North Fork Caspar Creek weir after about 50% of the 473-ha watershed had been clear-cut logged. However, when the uncut tributaries within the North Fork (pre-treatment RMSE = 0.118) were used as the controls, an increase in peaks was detected [Ziemer, 1998]. In the analyses described in this paper, uncut tributaries in the North Fork will be used as controls for treated subwatersheds in the North Fork.

Sediment Loads

Paired watershed studies have been utilized to study the effects of logging activities on sediment loads as well as peak flows. Detecting changes in sediment loads is even more difficult than for peak flows, because sediment loads are more variable and more costly to measure. Studies are often dominated by a single extreme event [Grant and Wolff, 1991; Rice et al., 1979; Olive and Rieger, 1991], making the results more difficult to interpret. Most studies have utilized annual sediment loads [Harris, 1977; Rice et al., 1979; O'Loughlin et al., 1980; Grant and Wolff, 1991; Megahan et al., 1995], usually determined by surveys of settling basins behind impoundments. Sediment passing over a

spillway is typically determined using sediment rating curves that relate suspended sediment concentration and water discharge.

Only one of these studies has been conducted in the redwood region. Rice et al. [1979] reported the suspended sediment load was 270% above that predicted for 1 year following roading of the South Fork of Caspar Creek, and the debris basin deposit 50% above that predicted. Lewis [1998] estimated an increase of 212% in suspended load in the 6 years following logging of the South Fork, despite a 3300 m^3 landslide contributing directly to the stream in the control watershed.

In the Alsea watershed in coastal Oregon, Brown and Krygier [1971] found a doubling of sediment loads in the year after roading in two different watersheds. In the watershed that was completely clear-cut and burned to the mineral soil the next year, sediment loads increased more than 10-fold the first year, then gradually declined in 7 years to near pretreatment levels [Harris, 1977]. In the watershed that was 25% clear-cut in three small units and remained mostly unburned, the road effect diminished in the second year, and measured increases in loads were not statistically different from the pretreatment relationship. Differences between sediment yields from the two treated watersheds were attributed primarily to the burning.

Sample sizes are necessarily rather limited in analyses using annual loads, an unfortunate situation, considering the variability in response. It is rare to find studies with more than 5 years of pretreatment measurements of sediment on both control and treated watersheds. Exceptions are the experiments in the Alsea [Harris, 1977] and the Silver Creek [Megahan et al., 1995] watersheds, which had 7 and 11 years' pretreatment data, respectively. Many studies have used no pretreatment measurements at all [Plamondon, 1981; O'Loughlin et al., 1980; Leaf, 1970]. These must rely on unproven assumptions about the relation between control and treated watersheds. Post-treatment sample sizes are limited by the rapidly changing conditions that usually follow a disturbance. In analyses based on annual loads, conditions might return to pretreatment levels before enough data are available to demonstrate a change occurred. Even if a change can be detected, it is difficult to establish reasonable bounds on the magnitude of change in the face of such high variability and small sample sizes.

Some paired watershed studies have attempted to look at changes in sediment concentrations. In the Alsea watershed study, an analysis of changes in sediment rating curves was less effective than an analysis of annual loads [Brown and Krygier, 1971]. Such analyses will usually be limited by the inadequacy of models relating sediment concentration to flow. Olive and Rieger [1991] were unable to establish a useful calibration using sediment concentrations, attributing the failure to the highly variable hydrologic environment. Fredricksen [1963] used paired specimens (collected within 1 hour of each other) to analyze changes in the H.J. Andrews concentrations, but found it necessary to discard 8 of 83 data points that represented "unpredictable events" and "sudden movements of soil". Considering the episodic nature of sediment transport, it is not surprising that simul-taneous specimens from adjacent watersheds are poorly related. Such episodic events should probably be focused upon rather than discarded.

Utilizing storm sediment loads circumvents the problems of properly pairing concentration data and permits much larger sample sizes than are possible in analyses of annual loads. Larger sample sizes permit more powerful statistical analyses and construction of confidence limits and prediction limits for responses. Because of the cost of reliably estimating storm loads, studies based upon them are rare. Miller [1984] estimated storm loads from three control and three treated watersheds using pumped specimens triggered at regular time intervals. Although no pretreatment data were collected, the replication of both treatment and control permitted an analysis of variance on storm ranks each year following the treatment. But sampling at regular time intervals will tend to miss peak concentrations in flashy watersheds unless the intervals are very short, in which case more field and lab work is required. In our study we used schemes that increased the probability of sampling during high flows and turbidities.

Cumulative Effects

A great deal of concern has been focused on the cumulative watershed effects of forest harvesting activities. This study design includes multiple gaging stations in the same watershed in order to evaluate cumulative effects. According to the U.S. Council on Environmental Quality's interpretation of the National Environmental Policy Act, a "cumulative impact" is the impact on the environment which results from the incremental impact of the action when added to other past, present, and reasonably foreseeable future actions, regardless of what agency…or person undertakes such other actions [CEQ guidelines, 40 CFR 1508.7, issued 23 April 1971]. An activity's importance may depend heavily upon the context of historic and future land use. An infinite variety of interactions is imaginable. We attempt to answer three questions that arise with regard to cumulative watershed effects of logging activities :

1. How are impacts related to the total amount of disturbance? In particular, were the effects of multiple disturbances additive in a given watershed?
2. How do impacts propagate downstream? In particular, were downstream changes greater than would be expected from the proportion of area disturbed?
3. Can activities that produce acceptable local impacts result in impacts that are unacceptable by the same standard at downstream locations? In particular, were sediment loads in the lower watershed elevated to higher levels than in the tributaries?

The scope of these questions is limited here in order to permit scientific investigation. For example, question (2) does not consider that larger watersheds may experience different types of impacts than contributing watersheds upstream, and question (3) does not consider that different standards may be appropriate downstream because different resources may be at risk. Nevertheless, partial answers to these questions can be provided with regards to storm peak flows, flow volumes, and suspended sediment loads through watershed experiments and mathematical modelling.

Environment and History

The Caspar Creek Experimental Watersheds are a pair of rain-dominated forested catchments in the Jackson Demonstration State Forest on the coast of northern California. The 473-ha North Fork and the 424-ha South Fork are both located in the headwaters of the 2,167-ha Caspar Creek, which discharges into the Pacific Ocean near the town of Caspar. Uplifted marine terraces, to 320 m in elevation, are deeply incised by antecedent drainages resulting in a topography composed of steep slopes near the stream channel and broad rounded ridgetops. About one third of the basin's slopes are less than 17° and only 7% are greater than 35°. The watershed receives an average of 1200 mm of rainfall each year, 90% falling in the months of October through April. The forest is composed mainly of redwood (*Sequoia sempervirens* [D.Don.] Endl.), Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco), grand fir (*Abies grandis* [Dougl. ex D.Don] Lindl.), and western hemlock (*Tsuga heterophylla* [Raf.] Sarg.). The well-drained clay loam soils developed in sandstone and shale units of the Franciscan assemblage [Bailey et al., 1964] and are highly erodible.

Streamside landslides, gully erosion, and debris flows are the major erosional processes delivering sediment to the channel system. Soil pipes, common in the unchannelized swales, and steep ephemeral tributaries discharge to the Caspar Creek main stems. Based on debris basin surveys and suspended sediment measurements, the perennial, gravel-bed North Fork channel typically transports about 70% of its sediment load in suspension, and sand rarely exceeds 50% of the suspension. Gravel bars associated with woody debris jams and debris-induced bank erosion furnish the bulk of bedload transported during peak flows. Finer sediments cap the highest gravel bars and are stored in pools for transport during modest storm flows [Lisle and Napolitano, 1998].

Between 1860 and 1904, the old-growth forest in the Caspar Creek watersheds was clear-cut and burned. Log drives were triggered by opening the spillway gates of log crib along the main-stem reaches of both the North Fork and the South Fork, profoundly affecting channel morphology during the earliest logging effort [Napolitano, 1998]. These

gave way to semi-mechanized yarding of tributary catchments using railway inclines (tramways) and steam donkeys [Henry, 1998]. A historic stage coach route and a mid-1900's era forest road totaling 11.4 km in length follow the watershed divide along the north and east of the North Fork.

In 1962, Caspar Creek became the site of a paired watershed experiment. In 1968, the South Fork watershed was roaded, and from 1971-1973, it was selectively logged by tractor, while the North Fork watershed was maintained as an undisturbed control [Rice et al., 1979; Ziemer, 1981; Wright et al., 1990; Keppeler and Ziemer, 1990]. In 1985 and 1986, 59 ha of an ungaged tributary basin in the lower North Fork was clear-cut. The present study of cumulative impacts began in 1985 in the 384-ha Arfstein subwatershed (ARF), gaged on the North Fork's main stem just above the confluence with the ungaged tributary (Figure 1). When the stability of ARF's discharge rating equation recently came into question, we decided to use the larger North Fork watershed (NFC) in place of ARF for the analysis of storm peaks and flow volumes. ARF was retained, however, for the sediment analyses because roughly 40% of the suspended sediment settles in a debris basin immediately above the North Fork weir and thus is not measured at the NFC gaging station.

METHODS

Treatment

The treatment design was based on compliance with the California Forest Practice Rules in effect in the late 1980's, except that the proportion of the watershed cut in a 3-year period was atypically high for a watershed of that size. Streams bearing fish or aquatic habitat were protected with selectively logged buffer zones 15 to 46 m in width, depending on stream classification and slope steepness.

Logging began in the headwaters of the North Fork in May 1989 and ended in the lower watershed in January 1992 (Figure 1). Clear-cuts totalled 169 ha (43% of ARF) in blocks of 9 to 60 ha and occupied 30% to 98% of treated subwatersheds. Total logged areas, including timber selectively removed from stream buffer zones, are slightly larger (Table 1). The 60 ha cutblock was composed of two adjacent subwatersheds (CAR and GIB), and an exemption was required from the maximum clear-cut size permitted under California Forest Practice Rules in effect at the time. Of the clear-cut areas, 81% was skyline yarded to landings on spur roads built on the upper hillslopes away from the creeks. Logs only had to be suspended at one point, but in most cases full suspension was achieved by setting the chokers near the middle of the log. This prevented ground dragging except near landings and convex slope breaks. The remaining 19% of the clear-cut area was tractor yarded and was limited to ridgetop areas where slopes were generally less than 20°. In addition, about 34% of the timber was selectively removed from 19 ha of stream buffer zones. New roads, landings, skid trails, and firelines occupied from 1.9% to 8.5% of treated subwatersheds. Four cut units, totalling 92 ha, were broadcast burned following harvest.

Three subwatersheds (HEN, IVE, and MUN) within the North Fork were retained in an unlogged condition for use as controls. In addition, the South Fork watershed, unlogged since 1973, was monitored for possible use as a control.

Gaging Stations

A total of 15 gaging stations were monitored: the North and South Fork weirs (NFC and SFC), four stations on the main stem of the North Fork, and nine on tributaries of the North Fork (Figure 1). The channel control structures at the North and South Fork gaging stations are 120° V-notch weirs with concrete upper rectangular sections. The lowest three main-stem stations (ARF, FLY, and LAN) are rectangular plywood sections, rated by discharge measurements. Each rated section has a natural bottom and a stable

downstream sill installed to control bed elevation within the rated section. Discharge at the upper main-stem station (JOH) and the nine tributaries is measured with Parshall flumes. Although the rated sections and flume installations were not designed to guarantee complete capture of subsurface intergravel flows, frequent inspections (before, during, and after storms) were made and regular maintenance was performed at these sites to ensure stable discharge estimates throughout the length of the study. Discharge ratings were validated with new measurements each year, and only station ARF required rating equation changes.

Suspended Sediment Data Collection

Selection At List Time (1986-1995). Selection At List Time (SALT) is a variable probability sampling method similar to PPS (probability proportional to size) sampling with replacement [Hansen and Hurwitz, 1943]. Their estimation formulas are identical. Both methods utilize an auxiliary variable, easily measurable for the entire population, to assign inclusion probabilities to each sampling unit of the population. (We have defined a sampling unit of the sediment population as the suspended sediment load passing a gaged cross-section in 10 min.) The variance is minimized for auxiliary variables that are proportional to the variable of interest. PPS requires enumerating the population and measuring the auxiliary variable on the whole population before sampling. SALT was developed as an alternative to PPS for populations which cannot be enumerated before sampling [Norick, 1969]. SALT inclusion probabilities are computed from an estimate of the auxiliary variable total. Immediately upon measuring each unit's auxiliary variable, a decision is made whether or not to select the sampling unit. The auxiliary variable might be a flow-based prediction of unit yield from a sediment rating curve [Thomas, 1985]. This results in a sampling rate that is proportional to predicted sediment yield. If the discharge and sediment rating curves are power functions of stage (water depth), the sampling rate will also be a power function of stage. In practice, we had to set an upper limit to the sampling rate and modify the parameters of the power function in order to sample small storms as well as large ones [Thomas, 1989].

To implement SALT, at each gaging station we interfaced an HP-71 calculator with an automatic pumping sampler and a transducer mounted in a stilling well. The calculator was programmed to "wake up" every 10 min, read the transducer stage height, calculate the auxiliary variable, and, using the SALT algorithm and a set of stored random numbers, decide whether to sample or not. If the decision was to sample, a signal was sent via an interface circuit board to the pumping sampler, which would then collect a specimen (to avoid ambiguity, the word "sample" is reserved to refer to a selected set of "specimens" or "bottles") from a fixed intake nozzle positioned in the center of the channel. Date, time, stage, and other bookkeeping details were recorded on the calculator for subsequent uploading.

Turbidity-controlled sampling (1996). After 10 years of monitoring, the number of gaging sites was reduced to eight: the North and South Fork weirs (NFC and SFC), two controls (HEN and IVE), one main-stem station (ARF), and three tributary stations (CAR, DOL, and EAG). At that time, SALT and the HP-71's were replaced by a turbidity-controlled sampling system utilizing programmable data loggers and *in situ* turbidity probes. Date, time, stage, turbidity, and sampling information are recorded at 10-min intervals. The nephelometric turbidimeters we are using emit infrared light and measure the amount scattered back to the probe. In lab tests, they respond linearly to sediments of a given size distribution. In the field, with mixed-size sediments present, departures from linearity are usually minor. During each storm event, when certain prespecified turbidity thresholds are reached, the data logger sends a signal to the pumping sampler to collect a concentration specimen. A separate set of thresholds is specified for falling and rising stage conditions. This system reduced sample sizes and field expenses considerably, while still permitting accurate estimation of sediment loads [Lewis, 1996].

Data quality control. Field crews typically visited each gaging station one to three times per 24-hour period during storms to check on flumes and equipment, record manual stage observations, measure discharge at rated cross-sections, and collect depth-integrated suspended sediment specimens. Chart recorders provided back-up data. When problems were encountered with the electronic stage record, they were corrected using observer records or digitized data from back-up chart recorders. In a few instances, portions of discharge records were corrected based on correlation with selected alternate gaging stations. All stage data were coded to indicate the quality of the data.

Storms with poor quality or reconstructed peak data were treated as missing data in the peaks analysis. Storms with 25% or more of the flow volume derived from poor quality stage data were treated as missing data in the flow volumes analysis.

In addition to the suspended sediment specimens collected by the SALT algorithm, auxiliary pumped specimens were manually initiated for comparison with simultaneous depth-integrated DH-48 specimens or to augment the sampling algorithm. On occasions when the HP-71/pumping sampler interface failed and could not be immediately repaired, the sampler was set to collect specimens at fixed time intervals. A total of 21,880 bottles were collected: 19,572 under SALT, 378 under the turbidity threshold algorithm, 1048 auxiliary, 686 depth-integrated, and 196 fixed-time specimens.

Suspended sediment concentration was determined in the laboratory using vacuum filtration. Specimens were coded to indicate such conditions as spillage, organic matter content, low volumes, and weighing errors. Those with serious errors were omitted from the analysis. Those with minor errors were re-examined in the context of the whole storm.

Field crews also noted conditions affecting discharge or sediment data including landslides, windthrow, and culvert blockages and diversions. Post-storm surveys of the watershed stream channels and roads were made to document erosion sources potentially affecting sediment loads.

Storm Definition and Feature Identification

A total of 59 storm events occurred during the 11-year study. Storm events were generally included in the study when the peak discharge at SFC exceeded 0.0016 $m^3s^{-1}ha^{-1}$ (recurrence interval about 7 times per year). A few smaller peaks were included in dry years. Multiple peak hydrographs were treated as multiple storms when more than 24 hours separated the peaks and the discharge dropped by at least 50% in the intervening period. When multiple peak hydrographs were treated as a single storm, the discharge for the peaks analysis was identified by selecting the feature corresponding to the highest peak at NFC. Thus the same feature was used at all stations, even if it were not the highest peak of the hydrograph at all stations. However, differences in peak discharge caused by this procedure were very small.

The start of a storm was chosen by seeking a point on the hydrograph, identifiable at all stations, where the discharge began to rise. The start times differed by no more than a few hours at the various stations. At the end of a storm, distinctive hydrograph features are more difficult to identify, unless a new start of rise is encountered. We therefore decided to use the same ending time for a given storm at all stations. The ending time was selected by observing the storm hydrograph for all stations and determining either the time of the next storm, the next significant rainfall, or a stable low-flow recession at all hydrographs, usually within about 3 days after the peak. The end of each storm was always well below the quickflow hydrograph separation point described by Hewlett and Hibbert [1967], except when the recession was interrupted by a new storm.

Dependent Variables

The response variables of interest in this study are storm runoff peak (instantaneous discharge), storm runoff volume (total discharge), and storm suspended sediment load (mass of particles greater than 1 micron in diameter). All are expressed on a unit area

(per hectare) basis. The runoff variables were derived from the 10-min electronic record of stage and rating equations relating discharge to stage at each station. The computation of sediment loads is more involved and is described in the next section.

Computation of Suspended Sediment Loads

Correction to obtain cross-sectional average concentration. The pumping sampler intakes were oriented downstream and centered in the inclined throat sections of the Parshall Flumes. In the rated sections (ARF, FLY, and LAN), the intakes were similarly oriented at a fixed position about 9 cm off the bed. To determine whether the specimens were starved or enriched because of sampler efficiency or nozzle orientation or position, simultaneous ISCO and DH-48 depth-integrated (equal transit rate) specimens were collected throughout the study. A log-log regression of depth integrated concentration versus fixed intake concentration was developed for each station. Although only six of thirteen regressions differed significantly from the line $y=x$ (experimentwise $\alpha=0.05$ with Bonferroni [Miller, 1981] adjustment), all fixed intake concentrations were adjusted using the back-transformed regression equations and corrected for bias [Baskerville, 1972] before storm loads were computed.

Load estimation in 1986-1995. Although sediment sampling followed SALT protocol in hydrologic years 1986-1995, we ultimately applied non-SALT methods of estimation to these samples for two reasons:

1. SALT does not provide a way to estimate sediment loads for periods when the sampling algorithm was inoperative due to equipment problems. Other methods can interpolate over such periods and utilize manually-initiated auxiliary specimens and those collected in fixed-time mode.
2. Using computer simulations on intensively collected storm data, other methods were found to have lower mean squared errors than SALT.

Although unbiased estimates of variance are not available for the alternate methodologies, the simulations strongly suggested that SALT variance estimates could be used as very conservative upper bounds on the variance. Two alternate methods were considered. In both of these methods the total load is computed by summing the products of water discharge and estimated concentration over all 10-min periods in the storm. The concentration, c, between adjacent sampled times t_1 and t_2 is modelled as either

1. a linear function of time: $c = c_1 + (t - t_1)(c_2 - c_1) / (t_2 - t_1)$, or
2. a power function of stage: $c = as^b$, where

$$b = \frac{\log c_2 - \log c_1}{\log s_2 - \log s_1}, \quad a = \frac{c_1}{s_1^b} \tag{1}$$

in which the subscripts identify concentrations and stages at times t_1 and t_2. These methods will be referred to as "time interpolation" and "stage interpolation" respectively. Stage interpolation has a better physical basis, but computational difficulties frequently arise when s_1 and s_2 are similar or equal, or when c_1 or c_2 is equal to zero. Therefore, time interpolation was substituted for stage interpolation when the power function defined by a pair of stages and sampled concentrations could not be computed or its exponent was not in the range between 1 and 10. If no specimens had been sampled within 10 hours prior to the start of the storm, the starting sediment concentration was assigned a value of zero and time interpolation was applied. An analogous procedure was followed for the end of the storm. The next section describes simulations leading to the decision that stage interpolation be used for estimating the sediment loads in 1986-1995.

Simulations comparing SALT and interpolation estimators. In addition to the usual SALT sampling, in 1994 and 1995 sediment concentration and turbidity at ARF was sampled at 10-min intervals for five storm events. This data, described in greater detail

by Lewis [1996], provided realistic populations with known sediment loads that could be used in simulations to evaluate the performance of different load estimation methods. In addition to these five populations, eight storm populations were available from previous studies on the North Fork of the Mad River in northwestern California: three storms from December 1982, January 1983 and December 1983 [Thomas and Lewis, 1995] and five storms from February 1983 [Thomas and Lewis, 1993]. The Mad River concentrations were derived from turbidity charts and form a smoother, less realistic, time series than the ARF measurements.

In the simulations, 5000 independent SALT samples were selected from each storm event using SALT sampling parameters that were in use at ARF in 1995 and parameters thought to be optimal at Mad River. The sediment load was estimated for each of the 5000 samples using SALT and time and stage interpolation. The simulation results are strictly applicable only to comparing these estimators under a specific SALT sampling protocol.

The simulation results are summarized in Table 2. While SALT was unbiased as expected, it consistently has much higher root mean square error (RMSE) than the interpolated estimators. This can be attributed to the interpolation methods that take advantage of local trends in concentration that SALT ignores. Because the Mad River storm populations were smoother than those from ARF, they indicate a somewhat greater advantage for the interpolated estimators.

While time interpolation appears to have slightly less bias than stage interpolation, the differences in both bias and RMSE are small relative to the loads. Real data differ from these simulated data in that unexpected time gaps are created during unavoidable equipment malfunctions. Stage interpolation is expected to mimic true concentrations better than time interpolation over large time gaps, so the latter method (with the exceptions noted earlier) was chosen for this study during the SALT years (1986-1995).

Quality control for load estimates (1986-1995). Determining which calculated sediment load data were of high enough quality to include in the analysis was a subjective process and involved an examination of plots showing the storm hydrograph, sediment concentrations, and quality codes. The primary considerations were the number of known concentrations (sample size) and their temporal distribution relative to the hydrograph. Out of 51 storms and 15 stations (765 combinations), 74% of the load estimates were judged acceptable. Because sample sizes were in proportion to the size of storm events, most of the discarded loads were from small events. In those events that were retained, the median sample size was 20 and the median standard error from SALT was 14% of the estimated load. Based on the simulations (Table 2) and the fact that SALT estimates did not utilize all the available concentrations, it is likely that the median error from the interpolated estimates is well under 10% of the estimated load.

Load estimation in 1996. With turbidity-controlled sediment sampling in place in 1996, sediment loads were computed using “turbidity rating curves”. Concentration was predicted by linear regressions of concentration on turbidity fit to each storm. This method was shown in simulations [Lewis, 1996], based on the same five ARF populations as shown in Table 2, to produce load estimates with RMSE of 8% or better while sample sizes were reduced to between 4 and 11, depending on storm size and sampling parameters. The interpolation methods used for 1986-1995 would not be as accurate for the generally smaller sample sizes obtained under turbidity-controlled sampling. However, because of intermittent fouling of the turbidity probes with debris and sediment, valid turbidities were not always available. During such periods, if enough concentration measurements were available (and extras were often triggered by false high turbidities), then time or stage interpolation was used. As a last resort, a sediment rating curve derived from nearby data was used to estimate concentrations. Out of 8 stations and 8 storms in 1996, a total of 46 sediment load estimates were judged to be of acceptable quality. The median sample size was 5 from these events.

Derivation of Independent Variables Used in the Analysis

The complete data set included both map-derived and field-derived variables. All disturbance variables were coded as proportions of watershed area. The basic watershed descriptors and variables that were useful in the analyses are shown in Table 1.

Topographic contours and streams were digitized from U.S. Geological Survey 7.5 min quadrangle maps. The mapped stream channels in harvest units were then extended to include all channels showing field evidence of annual scour and/or sediment transport before logging. Watershed boundaries were field-mapped using conventional tape-and-compass surveys, respecting diversions of surface runoff where road drainage structures directed flow into or out of the topographic watersheds. During road maintenance, efforts were made to limit changes in drainage due to ruts and berms. Harvest unit boundaries and roads were surveyed using differentially corrected GPS. All these lines were transferred to GIS coverages from which geographic variables were extracted. Burned areas, stratified into two severity classes, and herbicided areas were transferred to the GIS from field maps. For each variable measured, the area within 150 feet of a stream channel, and the length of channel within the affected area were extracted from the GIS.

The areal extent of ground disturbance from roads, landings, skid trails, firelines, and corridors created by dragging logs up the slope by cable were each determined from exhaustive field transects. The areas within 150 feet of a stream channel, and the number of stream crossings were also recorded for these variables.

Cutting age was calculated as the difference in hydrologic year of a given storm and the hydrologic year an area was logged. For watersheds with areas cut at different times, a weighted average cutting age was calculated using the cut unit areas as weights.

An antecedent wetness index intended to reflect seasonal differences in hydrograph response was derived using mean daily discharges from SFC. The daily discharges were accumulated and decayed using a 30-day half-life, i.e.

$$w_i = Aw_{i-1} + q_i \tag{2}$$

where w_i denotes the wetness index on day i, and q_i denotes the daily mean discharge at SFC on day i and the constant $A = 0.97716$ satisfies the relation A^{30}=0.5. The decision to use streamflow rather than precipitation to calculate antecedent conditions was based on the assumption that the history of the streamflow response would be a better predictor of streamflow than would the history of rainfall. The response of streamflow to precipitation is delayed as soil moisture deficit is recharged. A half-life of 30 days was selected to smooth the high frequency variation in streamflow, creating an index that would decline significantly only after lengthy dry periods. No optimization was done on the half-life, but it was found that $\log(w_i)$ made a slightly better predictor. The wetness index time series over the 11-year study period is displayed in Figure 2, with solid circles indicating the wetness level at the start of each storm. The wetness index varied from 13 to 150 at the onset of storms occurring in November and December, but assumed the full range from 13 to 562 at the onset of storms occurring in January, February and March. For two storms that occurred in May, the values of the index were 49 and 84.

Statistical Methods

Initially, simple log-log linear regressions were computed for each dependent watershed against selected control watersheds prior to treatment. The Chow test [Chow, 1960; Wilson, 1978] was used to test whether the post-treatment data differed in either intercept or slope from the pre-treatment regressions. Following Bonferroni's procedure [Miller, 1981] for these tests, an experimentwise error rate of 0.05 for 10 tests required setting the nominal α to 0.005 for each test. Because of their limited sample sizes, these tests, while easy to interpret, are not as powerful as models based on all of the data.

Models incorporating all of the watersheds were initially built up in a stepwise fashion using least squares estimation. At each step, residuals were plotted against candidate predictors to select the next variable and the appropriate transformation or form of inter-

action. Because a non-standard covariance model was employed, models were ultimately fitted using maximum likelihood estimation and selected using a combination of exploratory and diagnostic techniques.

Models for runoff (storm peaks and flow volumes). Consider the following pretreatment model:

$$\log(y_{ij}) = \beta_{0i} + \beta_{1i} \log(y_{Cj}) + \varepsilon_{ij} \tag{3}$$

where

y_{ij} = unit area response (peak or flow) at treated watershed i, storm j,
y_{Cj} = unit area response at control watershed in storm j,
ε_{ij} = non-independent normally distributed errors (see *Covariance Models* below),

and β_{0i} and β_{1i} are "location" parameters to be estimated for each watershed *i*. The log transformations are used in order that ε_{ij} appear to be normally distributed. The pretreatment model can be considered as a special case of the following model:

$$\begin{aligned} \log(y_{ij}) = &\left(\beta_{0i} + \beta_4 D_{ij} + \beta_6 D_{ij} \log(w_j) + \beta_7 D_{ij} a_i\right) \\ &+ \left(\beta_{1i} + \beta_5 D_{ij}\right) \log(y_{Cj}) + \varepsilon_{ij} \end{aligned} \tag{4}$$

where

D_{ij} = some measure of disturbance per unit area in watershed *i* at storm *j,*
w_j = wetness index at start of storm *j*,
a_i = drainage area of watershed *i,*

and β_4, β_5, β_6, and β_7 are parameters to be estimated. The log transformation of w_j is not critical, but was found to improve its explanatory value. Wetness enters the equation only as an interaction with D_{ij} because in the absence of disturbance wetness did not affect the relation between y_{ij} and y_{Cj}. As an interaction, it implies that the effect of disturbance on y_{ij} varies linearly with antecedent wetness. The $D_{ij}a_i$ term implies that the disturbance effect also varies linearly with watershed area and it is the key term in this model for detecting a cumulative effect. It describes how watershed impacts propagate downstream and we use it to test the null hypothesis that a unit area disturbance has the same unit area effect in watersheds of all sizes.

The first line of equation (4) permits the intercept of the relation between y_{ij} and y_{Cj} to change following disturbance. The second line, via the $D_{ij}\log(y_{Cj})$ term, permits the slope of that relation to change following disturbance. Equation (4) can be rearranged as

$$\begin{aligned} \log(y_{ij}) = &\beta_{0i} + \beta_{1i} \log(y_{Cj}) + \varepsilon_{ij} \\ &+ D_{ij}\left[\beta_4 + \beta_5 \log(y_{Cj}) + \beta_6 \log(w_j) + \beta_7 a_i\right] \end{aligned} \tag{5}$$

We now model the disturbance term using logged area and cutting age to represent loss of transpiration and interception following logging. Compacted areas such as roads, landings, skid trails, and firelines were not found to be useful predictors. Since relatively little transpiration occurs at Caspar Creek in the fall and winter, we treat areas logged in the fall or winter prior to the occurrence of a storm as special cases. Let

$$D_{ij} = f(t_{ij})(c_{ij}) + g(c'_{ij}) \tag{6}$$

where

t_{ij} = area-weighted mean cutting age (number of summers passed) in watershed *i* for areas logged in water years (defined as Aug.1 - July 31) preceding that of storm *j*
c_{ij} = proportion of watershed *i* logged in water years prior to that of storm *j*, and
c'_{ij} = proportion of watershed *i* logged prior to storm *j* but in the same water year

We model a linear recovery declining from a maximum of unity the year after cutting:

$$f(t_{ij}) = 1 - \beta_2 (t_{ij} - 1) \tag{7}$$

where β_2 is a parameter representing the recovery rate, and we assume the effect of newly cut areas depends only on the season they were cut:

$$g(c'_{ij}) = \beta_3^{(k)} c'_{ij} \tag{8}$$

where $\beta_3^{(k)}$ are parameters for the effect of cutting in the fall (k=1) and winter (k=2) immediately preceding storm j. Equation (6) becomes

$$D_{ij} = \left(1 - \beta_2 (t_{ij} - 1)\right) c_{ij} + \beta_3^{(k)} c'_{ij} \tag{9}$$

and the complete model is

$$\begin{aligned} \log(y_{ij}) = {} & \beta_{0i} + \beta_{1i} \log(y_{Cj}) + \varepsilon_{ij} \\ & + \left[\left(1 - \beta_2 (t_{ij} - 1)\right) c_{ij} + \beta_3^{(k)} c'_{ij}\right] \\ & \times \left[\beta_4 + \beta_5 \log(y_{Cj}) + \beta_6 \log(w_j) + \beta_7 a_i\right] \end{aligned} \tag{10}$$

To investigate whether unit area response increases downstream independently of disturbance, we can look for a relation between β_{0i} and a_i. Alternatively, we can replace β_{0i} with the linear expression $\beta_0^{(1)} + \beta_0^{(2)} a_i$ and test the hypothesis $H_0: \beta_0^{(2)} = 0$. If unit area responses tend to increase downstream, then cumulative impacts might occur where a response threshold of acceptability is exceeded only below some point in the stream network, even though unit area disturbance is no greater in that point's watershed than in watersheds further upstream.

Model (10) is not a linear model because it involves products of the parameters to be estimated. The non-linearity was introduced as a parsimonious way of modelling recovery with time since logging. It avoids introducing separate recovery parameters for each of the terms in equation (4) that involve D_{ij}.

Models for suspended sediment loads. Suspended sediment load from an untreated control watershed was found to be a much better predictor of sediment load at treated watersheds than water discharge at either location. However, the change in storm flow in the treated watershed, relative to that in the control, was found to be the next best predictor in a model for suspended sediment loads. The change in flow, Δq, was formulated two ways:

1. The residual from the flow model with D_{ij} set to zero

$$\Delta q_{ij}^{(1)} = \log(y_{ij}) - \left(b_{0i} + b_{1i} \log(y_{Cj})\right) \tag{11}$$

where b_{0i} and b_{1i} are estimates of the flow model parameters β_{0i} and β_{1i}.

2. The log of the ratio of the flows between the treated and control watersheds:

$$\Delta q_{ij}^{(2)} = \log\left(y_{ij} / y_{Cj}\right) = \log\left(y_{ij}\right) - \log\left(y_{Cj}\right) \tag{12}$$

The first form makes better sense hydrologically, but treating it as an independent variable may not be statistically legitimate later when estimating precision later on, because it involves parameter estimates from another model. Nevertheless, both forms of Δq were considered. These variables are not useful in a predictive setting because the flows are not known in advance, but the main purpose of these models is explanatory. If prediction is needed, then a third form might be substituted as an approximation to $\Delta q_{ij}^{(1)}$:

3. The predicted change in $\log(y_{ij})$ from equation (10):

$$\Delta q_{ij}^{(3)} = \left[\left(1 - b_2(t_{ij} - 1)\right)c_{ij} + b_3^{(k)} c'_{ij}\right] \times \left[b_4 + b_5 \log(y_{Cj}) + b_6 \log(w_j) + b_7 a_i\right] \quad (13)$$

where the b's are estimates of the β's in equation (10).

After Δq and one or two disturbance variables were included in the model, no further gains were realized in the sediment models by including factors such as antecedent wetness and cutting age. So, unlike the runoff models, the sediment models remain linear in their parameters:

$$\log(y_{ij}) = \beta_{0i} + \beta_{1i} \log(y_{Cj}) + \beta_2 \Delta q_{ij} + \left(\beta_3 + \beta_4 \log(y_{Cj}) + \beta_5 a_i\right) x_{ij}^{(1)} + \left(\beta_6 + \beta_7 \log(y_{Cj}) + \beta_8 a_i\right) x_{ij}^{(2)} + \varepsilon_{ij} \quad (14)$$

where

y_{ij} = unit area sediment load at treated watershed i, storm j,
y_{Cj} = unit area sediment load at control watershed in storm j,
Δq_{ij} = change in flow as defined by (11) or (12) in watershed i, storm j,
a_i = drainage area of watershed i,
$x_{ij}^{(1)}$ = a measure of unit area disturbance in watershed i, storm j,
$x_{ij}^{(2)}$ = a second measure of unit area disturbance in watershed i, storm j,
ε_{ij} = non-independent normally distributed errors (see *Covariance Models* below),

and the β's are parameters to be estimated. The logic behind the interaction terms involving $\log(y_{Cj})\, x_{ij}^{(k)}$ and $a_i x_{ij}^{(k)}$ is the same as in the runoff models. And, as with model (10), we can replace β_{0i} in (14) with the expression $\beta_0^{(1)} + \beta_0^{(2)} a_i$ to investigate whether unit area loads increase downstream independently of disturbance.

Covariance models. The residual covariance was found to depend upon watershed size and location. The correlations decreased with increasing distance between watershed centroids and the variance decreased with increasing watershed size. Serial autocorrelation in the residuals for most watersheds was weak or absent, so responses from different storms were considered independent. The errors were thus assumed to follow a multivariate normal distribution with a covariance matrix for each storm. The dimensions of this square matrix are equal to the number of treated watersheds having good data in that storm. The covariances in the matrix for storm j are modelled as:

$$Cov(\varepsilon_{i_1 j}, \varepsilon_{i_2 j}) = \sigma^2_{i_1 i_2} = \rho_{i_1 i_2} \sigma_{i_1} \sigma_{i_2} \quad (15)$$

where

$\rho_{i_1 i_2}$ = the correlation between $\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$,
σ_{i_1} and σ_{i_2} = the standard deviations of $\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$
$\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$ = errors for watersheds i_1 and i_2 in storm j

Subscripts j have been omitted from $\rho_{i_1 i_2}$, σ_{i_1} and σ_{i_2} because these terms are assumed to be independent of storm number and are, in fact, modelled upon the errors from all storms. Two models for the correlation $\rho_{i_1 i_2}$ were found to fit the runoff and sediment data.

1. Exponential decline with distance:

$$\rho_{i_1 i_2} = \frac{\exp(-\theta_1 d_{i_1 i_2}) + \theta_2}{1 + \theta_2} \tag{16}$$

where $d_{i_1 i_2}$ is the distance separating watersheds i_1 and i_2, and θ_1 and θ_2 are parameters to be estimated. In this model the correlations decline asymptotically from unity to the value $\theta_2/(1+\theta_2)$.

2. Linear decline with distance:

$$\rho_{i_1 i_2} = \begin{cases} 1, & d_{i_1 i_2} = 0 \\ \theta_1 - \theta_2 d_{i_1 i_2}, & d_{i_1 i_2} > 0 \end{cases} \tag{17}$$

The standard deviations σ_i were modelled as a declining power function of watershed area:

$$\sigma_i = \theta_3 a_i^{-\theta_4} \tag{18}$$

where θ_3 and θ_4 are parameters to be estimated. All peaks models discussed in this paper (other than the least squares fits) employed equations (15), (16), and (18). The flow and sediment models employed equations (15), (17), and (18)

Method of estimation. The parameters of the model were estimated using the method of maximum likelihood [Mood et al., 1974]. The likelihood function is assumed to be the multivariate normal density of the ε_{ij} treated as a function of the β and θ parameters. In practice we minimize the negative of the log likelihood. In this problem, the log-likelihood is equal to the sum of the independent storm-wise log-likelihoods. Thus, the dimension of the multivariate density function is the number of watersheds represented in a given storm, a maximum of 10. The log-likelihood functions and their gradients (derivative vectors) are shown in APPENDIX B. They were programmed in S-Plus [Statistical Sciences, 1995] and FORTRAN, and solved using the S-Plus function *nlminb* (nonlinear minimization subject to bound-constrained parameters). Least squares estimates of the parameters were used as starting guesses in these iterative numeric calculations.

Model size. The inclusion of up to 31 parameters in these models raises questions about overfitting. These questions were addressed by cross-validation (discussed below) after a model was selected, but the proper model size was selected with the objective of minimizing a variant of Akaike's information criterion [Burnham and Anderson, 1998],

$$\mathrm{AIC}_c = -2\log(L) + 2K\left(\frac{n}{n-K-1}\right) \tag{19}$$

where L is the maximum likelihood, K is the number of parameters estimated, and n is the sample size. This criterion is recommended over the unmodified AIC when the ratio n/K is small (less than about 40). The inclusion of the 20 location parameters β_{0i} and β_{1i} is strongly supported by AIC_c. Its value increased by 14 to 88 units in the various models when one or two parameters were substituted for either β_{0i} or β_{1i}. Increases of 10 or more AIC units indicate clearly inferior models [Burnham and Anderson, 1998]. Because of the computational time required to fit each model, it was impractical to obtain the likelihoods of all alternative models. For that reason, parameters other than β_{0i} and β_{1i} were evaluated using hypothesis tests based on the normal distribution, and AIC_c was computed only for the more promising candidate models.

Hypothesis testing. Maximum likelihood parameter estimates are approximately multivariate-normally distributed for large samples [Rao, 1973]. The estimated covariance

matrix of the estimates was obtained by inverting the observed information matrix, using a finite difference approximation to the Hessian, or matrix of second derivatives of the log-likelihood function [Bishop et al., 1975; McCullagh and Nelder, 1989]. (The observed information matrix is the negative of the Hessian, evaluated at the maximum likelihood estimates.) The standard errors, s_b, of the estimated parameters are the square roots of the diagonal of the covariance matrix. Since the parameter estimates are asymptotically normal, a simple test of the hypothesis H_0: $\beta_i = c$ is provided by observing whether or not the statistic $(b_i - c) / s_b$ is in the rejection zone of the standard normal distribution. The p-values from these hypothesis tests are identified as p_N in this paper. Tests with $p_N < 0.01$ are considered significant in this paper. Tests with $0.01 < p_N < 0.05$ are considered "suggestive" but not conclusive.

Observed change in response. "Observed change" in response was calculated by comparing the observed response, y_{ij}, with an estimate of the expected response, $E(y'_{ij})$, from the same storm and watershed in an undisturbed condition. We define the percentage change in response as

$$p_{ij} = 100\left(\frac{y_{ij} - E(y'_{ij})}{E(y'_{ij})}\right) = 100\left(\frac{y_{ij}}{E(y'_{ij})} - 1\right) \tag{20}$$

The expected undisturbed response, $E(y'_{ij})$, is a function of $E\left(\log(y'_{ij})\right)$:

$$E(y'_{ij}) = \exp\left[E\left(\log(y'_{ij})\right) + \tfrac{1}{2}\sigma_i^2\right] \tag{21}$$

Setting disturbance to zero in either model (10) or (14) above, we have $E\left(\log(y'_{ij})\right) = \beta_{0i} + \beta_{1i}\log(y_{Cj})$. The variances σ_i^2 are a function of θ_3 and θ_4 given by model (18). A nearly unbiased estimator of $E(y'_{ij})$ is given by

$$\hat{y}'_{ij} = \exp\left[b_{0i} + b_{1i}\log(y_{Cj}) + \tfrac{1}{2}\left(\hat{\theta}_3 a_i^{\hat{\theta}_4}\right)^2\right] \tag{22}$$

where b_{0i}, b_{1i}, $\hat{\theta}_3$, and $\hat{\theta}_4$ are the maximum likelihood estimates of β_{0i}, β_{1i}, θ_3, and θ_4, respectively. The term $\frac{1}{2}\hat{\sigma}_i^2 = \frac{1}{2}\left(\hat{\theta}_3 a_i^{\hat{\theta}_4}\right)^2$ is often called the Baskerville [1972] bias correction. An approximation for p_{ij} that we will call the "observed change in response" is obtained by substituting $\hat{y}'_{ij}$ for $E(y'_{ij})$ in (20):

$$\tilde{p}_{ij} = 100\left(\frac{y_{ij}}{\hat{y}'_{ij}} - 1\right). \tag{23}$$

Of course we are not just interested in the changes in response for the particular values of the explanatory variables encountered during the study. We would like to study the percentage change, p_0, for an arbitrary vector, $\mathbf{x}_0$, of explanatory variables. An unbiased estimator and confidence interval for $E(p_0)$ as well as a prediction interval for p_0 are derived in APPENDIX C. The confidence interval represents the uncertainty of the mean, $E(p_0)$, given $\mathbf{x}_0$. The prediction interval indicates the variability in the individual response p_0, given $\mathbf{x}_0$. Prediction intervals are wider than confidence intervals because they include the variability in the response about its mean value as well as the variability due to uncertainty in the mean itself.

Cross-validation of models. To investigate the possibility that the models were overfitted to the data, ten-fold cross-validation was used [Efron and Tibshirani, 1993]. The data are split into ten groups. Each observation is predicted from a model based on all of the data except that group to which the observation belongs. The RMSE of these predictions is called the cross-validation prediction error and it may be compared with the RMSE of the models fitted with all the data to assess overfitting.

A regression of the observed responses on the fitted values, known as the *calibration*, should have an intercept near zero and slope near unity. The regression of the observed responses on the cross-validated predictions is expected, in general, to have a slope less than one [Copas, 1983]. This phenomenon, known as *shrinkage*, implies that predictions of high or low response values tend to be too extreme. The degree of departure of the calibration slope from unity provides another measure of overfitting.

Because the data were not independent, the cross-validation was repeated using two different methods for splitting the data:

1. Data were randomly divided into groups of equal size.
2. Post-treatment data were omitted systematically, one station at a time.

The latter method does not provide cross-validated predictions for the pre-treatment data, but if all the data from a station, say watershed i, are omitted, it becomes impossible to estimate β_{0i} and β_{1i}, which are required to make predictions for that watershed. Nevertheless, the one-station-at-a-time method is probably a more rigorous validation for the inclusion of alternative disturbance variables because it will give higher error rates for models that include variables correlated with the response due to just one or two watersheds.

RESULTS

Storm Peaks

The analysis included 226 pre-treatment and 300 post-treatment observations representing 59 storms on the 10 treated watersheds. For the 226 pretreatment peaks, the control watersheds correlating best with watersheds to be treated were tributaries HEN and IVE, and MUN (Figure 3). The mean of the peaks at HEN and IVE (designated HI), or at HEN, IVE, and MUN (designated HIM), had higher correlations than did peaks from either HEN, IVE, or MUN individually. Because MUN was not monitored the last year of the study, HI was chosen as the control for the peaks analysis.

The Chow tests [Chow, 1960; Wilson, 1978], based on the HI control, revealed strong evidence that post-treatment data differed from pre-treatment regressions. Eight of the 10 watersheds departed ($p < 0.005$) from these regressions after logging commenced. The other two, FLY and LAN on the main stem, had p-values less than 0.05. Departures from the pre-logging regression were greatest in the clear-cut tributaries: BAN, CAR, EAG, GIB, and KJE (Figure 4).

Seasonal patterns in the departures from the pre-treatment regressions were evident in most of the treated watersheds. For example, Figure 5 shows the post-logging departures for watershed EAG plotted against storm number. The largest percentage departures occurred early in the season. These were usually, but not always, relatively small storms. Storms 28 and 29 did not show treatment effects, apparently because logging had just taken place the same winter, so insufficient time had elapsed for soil moisture differences to develop between the controls and the logged area. This exemplifies the situation that necessitated modelling of the disturbance term using equation (9).

To develop an overall model, an intercept and slope for each watershed (equation (3)) was initially fit by least squares. The residuals from this model show a strong interaction between proportion of area logged and antecedent wetness (Figure 6). Area logged includes clear-cut areas and a portion of each buffer zone corresponding to the proportion of timber removed (Table 1). The relation of the residuals with area logged is linear, the slope decreasing from strongly positive with increasing wetness (Figure 6, top row). The

relation with log(wetness) is linear, the slope becoming strongly negative with increasing logged area (Figure 6, bottom row). These relations imply a product term is an appropriate expression of the interaction, and the coefficient is expected to be negative. The fact that the average residual increases with different categories of area logged but not with wetness shows that a solo logged area term is needed in the model as well as the interaction product, but a solo wetness term is not. No variables related to roads, skid trails, landings, firelines, burning or herbicide application were found to improve the fit of the linear least squares model that includes logged area and its interaction with wetness. Adding logged area and the wetness interaction to the model, a plot of post-treatment residuals against time after logging (Figure 7) indicates an approximately linear recovery trend in the first 7 years.

When model (10) was fit to the data, the coefficient b_7 on the cumulative effect term did not differ significantly from zero (Table 3, p_N= 0.21). The coefficient b_5 was negative but not highly significant (p_N=0.047), weakly suggesting that the effect of logged area on peak flows tends to diminish in larger storms. The coefficient b_4 on logged area was positive as expected and its interaction with wetness, b_6, was negative as expected. The recovery coefficient, b_2, indicates an average recovery rate of about 8% per year. The null hypothesis for each of the parameters $\beta_3^{(k)}$ is H_0:$\beta_3^{(k)} = 1$, because the recovery model assumes a value of unity the year after logging. The coefficient $b_3^{(1)} = 0.59$ (standard error 0.10) indicates a reduced effect from fall logging on peaks in the following winter and $b_3^{(2)} = 0.00$ suggests that the effects of winter logging on peak flow are delayed until a growing season has passed.

There was no indication of a dependency on watershed area in either the coefficients b_{0i} or b_{1i} from model (10). When we replaced β_{0i} in model (10) with the expression $\beta_0^{(1)} + \beta_0^{(2)} a_i$, the coefficient $b_0^{(2)}$ was not significantly different from zero (p_N=0.58), indicating no trend of unit area storm peak with watershed area.

The exponentially declining correlation model (18) was used when solving model (10) for peak flows (with β_7 fixed at zero), and it can be seen to be a reasonable fit (Figure 8). The variance model (18) also seems reasonable (Figure 9). The Box-Pierce test [Shumway, 1988] did not indicate the presence of serial autocorrelation at any of the stations (minimum p=0.089). The residuals conform very well to the normal distribution (Figure 10), as do plots for individual stations (not shown), validating our choice of likelihood function. The lone outlier is from a storm at GIB that produced 2 peaks at all stations except GIB. (The first peak was selected for the storm but was identifiable only as a shoulder of the hydrograph at GIB.) The model fits the data very well (Figure 11). For the regression between observed and fitted values, $r^2 = 0.946$. This compares with $r^2 = 0.848$ for a model with no disturbance variables and $r^2 = 0.937$ for model (3) fit to only the pre-treatment data, so the model fits the post-treatment data as well as the pre-treatment data.

Magnitude of observed changes. Maximum peak flow increases based on equations (22) and (23) were about 300%, but most were less than 100% (Figure 12). The mean percentage increase declined with wetness but was still positive even under the wettest conditions of the study ($w_i > 500$), when it was 23% for clear-cuts but only 3% in partially cut watersheds. Increases more than 100% generally only occurred in clear-cuts under relatively dry conditions ($w_i < 50$) and when peaks in the control were less than 0.0025 $m^3s^{-1}ha^{-1}$ (return period 3-4 times per year). Large increases occurred less frequently as the winters progressed, but increases over 100% did occur in January and February. The mean percentage increase in peak flow declined with storm size and then levelled at an average increase of 35% in clear-cuts and 16% in partially cut watersheds for peaks greater than 0.004 $m^3s^{-1}ha^{-1}$ (return periods longer than 0.5 years) (Figure 13). For a storm size having a 2-year return period, the average peak-flow increase in 100% clear-cuts was 27% [Ziemer, 1998].

Figure 14 shows 95% confidence intervals for the modelled mean response in a 20-ha watershed that has been 50% clear-cut, for two wetness conditions and two cutting ages within the range of our data. The effect of antecedent wetness is a greater influence on

the response than time since cutting, although the recovery data only span 7 years. Prediction intervals are much wider than confidence intervals, revealing post-treatment variability that is greater than the treatment effect itself.

Storm Runoff Volume

The analysis included 527 observations representing 59 storms. For the same reasons as in the peaks analysis, HI (the mean of HEN and IVE) was chosen as the control. The modeling results are similar to the peaks analysis results, except that the watershed area interaction b_7 was marginally significant (Table 4, p_N=0.012) and watershed correlations were found to decline linearly with distance, so model (17) was used instead of (16) in the covariance model. For the sake of brevity, the modeling results for storm runoff volume are omitted, and we report only the coefficients (Table 4) and the magnitude of observed changes.

Magnitude of observed changes. The maximum storm runoff volume increase from equations (22) and (23) was 400%, but most were less than 100%. The mean percentage increase declined with wetness but was still positive even under the wettest conditions of the study ($w_i > 500$), when it was 27% for clear-cuts and 16% in partially cut watersheds. Increases more than 100% generally only occurred in clear-cuts under relatively dry conditions ($w_i < 100$) and when runoff volume in the control was less than 250 m^3ha^{-1}. Large increases occurred less frequently as the winters progressed, but in-creases over 100% did occur in January and February. The mean percentage increase in storm runoff volume declined with storm size and then leveled at an average increase of 30% in clear-cuts and 13% in partially cut watersheds for storm runoff greater than 250 m^3ha^{-1}.

Annual storm runoff volume (sum of storms) increased an average of 58% (1119 m^3ha^{-1}) in clear-cut watersheds and 23% (415 m^3ha^{-1}) in partly clear-cut watersheds (Table 5). Based on the complete discharge record at NFC, the runoff volume for the storms included in this analysis represents 41 to 49% of the total annual runoff volume in individual tributaries.

Figure 15 shows confidence intervals and prediction intervals for storm runoff volume in a 20-ha watershed that has been 50% clear-cut, under two wetness conditions and two cutting ages within the range of our data.

Suspended Sediment Loads

The relatively large number of missing observations resulting from quality control screening complicated the selection of controls for the sediment analysis. The use of synthetic controls such as HI and HIM permitted larger sample sizes because these means could be computed from any combination of non-missing controls. Thus the sample size was 376 with the HIM control, but only 333 with the HI control, and less than 300 with HEN or IVE alone. Although HIM control permitted the largest sample size, its correlations tended to be lower than those of HI (Figure 16). We therefore present the analysis twice, once with the HIM control and once with the HI control.

Chow tests [Chow, 1960; Wilson, 1978] for treatment effects at individual stations gave mixed results (Table 6). Only 2 of the tests were significant when HIM was used as the control and 3 were significant with the HI control. The tributaries all had more significant changes than the main-stem stations. Figure 17(top row) indicates that suspended sediment loads increased in all the clear-cut tributaries except KJE, where loads appear to have decreased after logging. The only partly clear-cut watershed on a tributary (DOL) also showed highly significant increases in sediment loads. The upper main-stem stations (JOH and LAN) showed no effect after logging, and the lower main-stem stations (FLY and ARF) experienced increases only in smaller storms. Summing suspended sediment over *all* storms, the four main-stem stations all showed little or no change (Table 7). Sediment loads at the North Fork weir, below ARF, increased by about 89%

per year, mainly as a result of a large landslide in the ungaged subwatershed that enters between ARF and NFC.

Models with HI control. The analysis included 333 observations representing 43 storms. In these models (14), the change in storm flow volume $\Delta q_{ij}^{(1)}$ was found to be the best explanatory variable after sediment load from the HI control, y_{HI}. Figure 18 shows the relation between the post-treatment sediment departures from pretreatment model (3)) and $\Delta q_{ij}^{(1)}$. Since both variables are differences in logarithms, it is convenient to express them as ratios of observed to predicted response, obtained by exponentiating the differences. The linear correlation between the sediment and flow departures is 0.54.

After $\Delta q_{ij}^{(1)}$ is in the model, disturbance variables explain only a very small part of the remaining variation (Figure 19). The length of unbuffered stream channel in clear-cut areas was one of the more useful disturbance variables in the sediment models. Under California Forest Practice Rules in effect during the North Fork logging, vegetation buffers were not required for stream channels that do not include aquatic habitat. The best models were found when this variable was separated into channels in burned clear-cuts and channels in unburned clear-cuts. The variable did not need to be separated, however, in the interaction terms. Thus the model (14) was modified to:

$$\begin{aligned}\log(y_{ij}) = {} & \beta_{0i} + \beta_{1i}\log(y_{(\mathrm{HI})j}) + \beta_2\Delta q_{ij}^{(1)} \\ & + \beta_3 x_{ij}^{(1)} + \beta_4 x_{ij}^{(2)} + \beta_5(x_{ij}^{(1)} + x_{ij}^{(2)})\log(y_{(\mathrm{HI})j}) \\ & + \beta_6(x_{ij}^{(1)} + x_{ij}^{(2)})a_i + \varepsilon_{ij}\end{aligned} \tag{24}$$

where

$x_{ij}^{(1)}$ = length of stream channel in burned clear-cuts, and

$x_{ij}^{(2)}$ = length of stream channel in unburned clear-cuts

To indicate the relative contribution of the various terms in model (24), the increase in residual sum of squares is shown for least squares models after dropping each explanatory variable (Table 8).

The maximum likelihood estimates for model (24) are shown in Table 9. The coefficient estimate b_3 is about 1.8 times b_4, suggesting that streams in burned clear-cuts contribute more sediment than those in unburned clear-cuts. The estimate, b_5, of the storm size interaction is negative, suggesting that the ratio between post-treatment and pretreatment sediment loads diminishes for larger events. The estimate, b_6, of the cumulative effect coefficient in this model was negative and was found marginally significant ($p_N = 0.044$). This interaction in the sediment model only partly offsets the small positive interaction that was noted in the runoff model and is hidden in the term $\Delta q_{ij}^{(1)}$. Other variables being equal, the model still predicts larger unit area sediment loads from larger watersheds (Figure 20). Because of its marginal significance, the β_6 term was dropped from the model for the remainder of this section.

The fitted intercepts b_{0i} from model (24), with β_6 fixed at zero, tend to increase with watershed area (Figure 21), with the exceptions of KJE (K) and JOH (J). This pattern in the intercepts is confirmed by substituting $\beta_0^{(1)} + \beta_0^{(2)} a_i$ for the term β_{0i}. The fitted coefficient $b_0^{(2)}$ is positive and differs significantly from zero (p_N=0.0031). The slope coefficients b_{1i}, are all between 0.8 and 1, except BAN (0.73) and EAG (1.06), and show no trend with area. Thus, ignoring the anomalous KJE and JOH for the moment, the unit area sediment loads from the watersheds prior to disturbance (Figure 22) tend to be highest in the four largest watersheds (ARF, FLY, LAN, and DOL), followed by the tributaries CAR, GIB, and EAG, and are lowest in the smallest watershed BAN.

Although there are signs of positive or negative trends in some individual watersheds, the residuals from model (24) display little if any trend with time (Figure 23). If the

anomalous JOH and KJE, which did not show treatment effects, are omitted, hints of a recovery trend disappear entirely.

The covariance model fit rather well for the sediment models based on HI. Correlations declined linearly with watershed separation (Figure 24) and variance declined as a power function of watershed area (Figure 25). The Box-Pierce test [Shumway, 1988] indicated (using an experimentwise error rate of 0.05) the presence of serial autocorrelation at four stations (ARF, BAN, GIB, and KJE) and suggests that we conservatively assess marginally significant terms in the model. The residuals again conform very well to the normal distribution and there is only one outlier (associated with stream bank collapses in EAG). The regression between observed and fitted values has $r^2 = 0.915$. This compares with $r^2 = 0.828$ for a model with no disturbance variables and $r^2 = 0.948$ for model (3) fit to only the pre-treatment data. So the complete model (without the cumulative effects term) explains (0.915 – 0.828) / (0.948 – 0.828) = 72% of the variation introduced by the post-treatment data.

Models with HIM control. This analysis included 376 observations representing 51 storms. In models developed with the HIM control, the log-ratio flow variable $\Delta q_{ij}^{(2)}$ was found to be a better explanatory variable than the flow model residual $\Delta q_{ij}^{(1)}$. The most important disturbance variable in these models is proportion of the watershed occupied by road cuts and fills. The length of stream channel in clear-cuts and the interaction terms in model (24) were not significant when tested in maximum likelihood models with the HIM control. This is partly explained by a high correlation (0.80) between road cut/fills and stream length in burned areas. A negative interaction between road cut/fills and watershed area was marginally significant (p_N=0.037). The maximum likelihood estimates for the model

$$\log(y_{ij}) = \beta_{0i} + \beta_{1i}\log(y_{(\mathrm{HIM})j}) + \beta_2\Delta q_{ij}^{(2)} + \beta_3 x_{ij} + \beta_4 x_{ij} a_i + \varepsilon_{ij} \qquad (25)$$

where x_{ij} is the proportion of the watershed occupied by road cuts and fills, are shown in Table 10. As with model (24), the interaction only serves to partly offset the positive interaction hidden in the $\Delta q_{ij}^{(2)}$ term, and we do not consider it significant. The trend in intercepts that was seen for model (24) is also present in model (25). Setting β_4 to zero, and substituting $\beta_0^{(1)} + \beta_0^{(2)} a_i$ for β_{0i}, we test $\beta_0^{(2)}$ and again find that it is positive and differs significantly from zero (p_N=0.0023). The residuals from model (25), with β_4 fixed at zero, do not display a significant trend with time since logging.

Magnitude of observed changes. Sediment load increases were calculated using equations (22) and (23) with the coefficients estimated from model (25). Median increases were 64% in partly clear-cut watersheds and 107% in clear-cut watersheds (Figure 26). Absolute increases were similar in clear-cut and partly clear-cut watersheds (Figure 27). Most of the larger percentage increases in clear-cuts were from small events and equated to relatively minor absolute increases in load. As one would expect, there is a tendency for percentage increases to decrease with storm size, and for absolute increases to increase with storm size. Figure 28 shows 95% confidence intervals and prediction intervals for the sediment model (25), with the area×disturbance interaction, β_4, set to zero. The watersheds are ranked by increasing proportion of road cuts and fills (x_{ij}). The uncertainty in the model and the variability in suspended sediment loads is much greater than for peak flow or storm runoff volume.

Summing storms by year, annual suspended sediment loads increased an average of 212% (262 kg·ha^{-1}yr^{-1}) in clear-cut watersheds and 73% (263 kg·ha^{-1}yr^{-1}) in partly clear-cut watersheds (Table 11). The absolute increases are heavily influenced by outlying data points that tend to occur in wet years (1993 and 1995), while the percentage increases weight all years approximately equally. If the extreme outlier in the partly clear-cut population (Figure 27) is omitted, the mean increase in that category drops to

67% (180 kg·ha^{-1}yr^{-1}). Because of the highly skewed distribution of sediment loads, median increases were much smaller: 109% (59 kg·ha^{-1}yr^{-1}) in clear-cut watersheds and 52% (46 kg·ha^{-1}yr^{-1}) in partly clear-cut watersheds. Based on the complete discharge record at NFC, the storms included in this analysis represent 36 to 43% of the total annual runoff in individual tributaries. However, these storms include roughly 90% of the annual suspended sediment load [Rice et al., 1979].

Cross-Validation of Models for Runoff Peaks, Volumes, and Sediment Loads

Predictions of storm runoff from random 10-fold cross-validation had RMSE only 2 to 3% (peaks) and 4% (volumes) higher than those from the original fitted models, for both pre-treatment and post-treatment responses (Table 12). The systematic cross-validation, omitting the post-treatment data one station at a time, gave RMSE 5% and 7% higher than the apparent post-treatment RMSE from the original runoff peaks and volume models, respectively. The systematically cross-validated RMSE values of 0.1739 and 0.1676 for logarithms of peaks and volumes correspond to prediction errors of about 20% for the untransformed responses. Calibration slopes (for regression of the observed versus predicted runoff) are very close to unity (Table 13) for both peaks and volumes. Both the random and systematic cross-validation calibrations are nearly indistinguishable from $y = x$ on 600 dpi letter-size plots. Both the RMSE and calibration results indicate the models for runoff peaks and volumes are not overfit. Remarkably, they appear to predict independent data nearly as well as the data to which the models were fit.

Predictions of suspended sediment loads from random cross-validation had RMSE 7% (HI control) and 4% (HIM control) higher than those from the original fitted models, for both pre-treatment and post-treatment responses (Table 12). On the other hand, the systematic cross-validation gave RMSE 32% (HIM control) and 50% (HI control) higher than the apparent post-treatment RMSE from the original sediment models. The systematically cross-validated RMSE values of 0.6724 and 0.6966 for logarithms of sediment loads correspond to prediction errors of about 100% for the untransformed responses. Calibration slopes for the sediment models are similar to the original models for the random cross-validation, but the systematic cross-validation has calibration slopes significantly smaller (Table 13), indicating substantial shrinkage in prediction of data from subwatersheds not used in model-fitting. The cross-validations indicate that the sediment models are not likely to predict future sediment loads well, and the associations identified between sediment loads and the disturbance variables in these models may be coincidental.

DISCUSSION

Storm Peaks

The effect of logging second-growth forests on streamflow peaks in Caspar Creek is consistent with the results from studies conducted over the past several decades throughout the Pacific Northwest. That is, the greatest effect of logging on streamflow peaks is to increase the size of the smallest peaks occurring during the driest antecedent conditions, with that effect declining as storm size and watershed wetness increases. However, increases were still apparent even in the largest storm of this study, which had a recurrence interval of 7 years at NFC.

Although the relative increases in peak flows tend to decline as storm size increases, the effects on large storms may still be important when recurrence intervals of a given size peak are considered. The curve for m=2, for example, in Figure 29 shows the increase in peak needed to reach a size that formerly had twice the recurrence interval, based on a curve fitted to the 28-year pre-logging partial duration series at NFC. Equivalently these are the increases necessary to halve the recurrence interval of the

peaks that would result from the increased flow regime. Under such a flow regime, the frequency of large peaks of a given size would double, roughly doubling the geomorphic work performed on the channel. For comparison, the increased peak flows observed in this study (Figure 13) have been included in Figure 29, assuming unit-area flow frequencies in the tributaries are the same as at NFC. Although the variability is very great, it appears that the average observed increases in clear-cuts are great enough to roughly halve the recurrence intervals for storm sizes greater than 0.004 $m^3s^{-1}ha^{-1}$ (return periods longer than 0.5 years). Average observed increases in partly cut watersheds were smaller.

Accounting for the amount of watershed disturbance, there was no evidence that either storm peaks or the logging effect on peaks was related to watershed size. Peaks in the smallest drainages tended to have greater responses to logging than in larger watersheds, but this was because the smaller watersheds had greater proportions disturbed. That is the typical pattern because Forest Practice Rules and economics usually limit the amount of intense activity occurring within any given watershed in any year. Therefore, it is possible for entire small first-order watersheds to be logged within a single year. However, as the size of the watershed increases, a smaller proportion of the watershed is likely to be logged in any given year. In the largest watersheds, harvesting may be spread over decades, within which time the earliest harvested areas will have revegetated.

The data from the streamflow, pipeflow [Ziemer, 1992; Keppeler and Brown, 1998], and soil moisture studies [Keppeler et al., 1994] at Caspar Creek all suggest that the peak flow response to logging is related to a reduction in vegetative cover. Reducing vegetative cover, in turn, reduces transpiration and rainfall interception. Since little soil moisture recharge occurs during the spring and summer growing season at Caspar Creek, large differences in soil moisture can develop between logged and unlogged watersheds by late summer because of differences in evapotranspiration. For example, by late summer, a single mature pine tree in the northern Sierra Nevada depleted soil moisture to a depth of about 6 m and to a distance of 12 m from the trunk [Ziemer, 1968]. This single tree transpired about 88 m^3 more water than the surrounding logged area, equivalent to about 180 mm of rainfall over the affected area. In the South Fork of Caspar Creek, the largest changes in peak streamflow after logging were found to be for the first storms after lengthy dry periods [Ziemer, 1981]. Similarly, after logging the North Fork, there was a strong interaction between the proportion of the area logged and watershed wetness that explained differences in streamflow peaks.

Evaporation of rainfall intercepted by the forest canopy can result in a substantial reduction in the amount of water that reaches the ground. Preliminary measurements at Caspar Creek suggest that average rainfall interception is about 20% of gross winter rainfall. Studies elsewhere have also reported that a large portion of annual rainfall is intercepted and evaporated from the forest canopy. For example, Rothacher [1963] reported that under dense Douglas-fir stands in the Oregon Cascades, canopy interception loss averaged 24% of gross summer precipitation and 14% gross winter precipitation. Percentage interception losses are greatest during low-intensity rainfall interspersed with periods of no rain. As with transpiration, rainfall interception can contribute to important differences in antecedent conditions between logged and unlogged watersheds. And during the large high-intensity storms that result in large streamflow peaks, rainfall interception is still important; about 18% of the rainfall from a 96-mm 24-hour storm was intercepted by the forest canopy at Caspar Creek. Differences in interception loss between logged and unlogged areas probably explain most of the observed increases in the larger winter peaks, when transpiration is at its annual minimum.

Road construction and logging were not applied as separate treatments in this study. And, because they are correlated, it is difficult to distinguish their effects statistically. However, soil compaction from roads and timber harvest represents only 3.2% of the North Fork watershed and ranges from 1.9% to 8.5% for the tributary watersheds. Further, roads, landings, and skid-trails in the North Fork are all located near the ridges and well away from any streams. Consequently, roads, soil compaction, and overland flow probably did not produce important changes in peak flow response of the North Fork watersheds. The recovery rate of about 8% per year for storm peaks supports the hypothesis that changes in peak flows are largely controlled by changes in vegetation.

Storm Runoff Volume

Analogous to the storm peaks model, the model for storm flow volumes showed that flow increases could be largely explained by the proportion of a watershed logged, an antecedent wetness index, and time since logging. Logging probably impacted both storm peaks and flow volumes via the same mechanisms: reduction of rainfall interception and transpiration.

Suspended Sediment Loads

The most important explanatory variable identified by the sediment models was increased volume of streamflow during storms after logging. This result is not unexpected because, after logging, increased storm flows in the treated watersheds provide additional energy to deliver and transport available sediment and perhaps to generate additional sediment through channel and bank erosion.

Whereas individual watersheds show trends indicating increasing or decreasing sediment loads, there is no overall pattern of recovery apparent in a trend analysis of the residuals from the model (Figure 23). This is in contrast with the parallel model for storm flow volume, and suggests that some of the sediment increases are unrelated to flow increases.

Other variables found to be significant, depending on the control watersheds used, were road cut and fill area and length of unbuffered stream channel, particularly in burned areas. One must be cautious about drawing conclusions about cause and effect when treatments are not randomly assigned to experimental units and replication is limited. Increases in sediment load in one or two watersheds can create associations with any variable that happens to have higher values in those watersheds, whether or not those variables are physically related to the increases. In this study, the contrast in response was primarily between watershed KJE, where sediment loads decreased, versus watersheds BAN, CAR, DOL, EAG, and GIB. Watershed KJE was unburned and also had the smallest amount of unbuffered stream of all the cut units. Watersheds EAG and GIB were burned and had the greatest amount of unbuffered stream in burned areas. Watershed EAG experienced the largest sediment increases and also had the greatest proportion of road cut and fill area. EAG was not unusually high in road surface area, and the larger road cut and fill area in EAG reflects roads that are on steeper terrain than in the other cut units.

Road systems would typically be expected to account for much of the sediment. During storm events frequent cutbank failures and culvert blockages along the pre-existing North Fork perimeter all-season road (dating back more than half a century) resulted in drainage diversions and sediment input to North Fork tributaries both before and after logging. But there is little field evidence of sediment delivery from the *new* spur roads in the North Fork watershed. In an inventory of failures greater than 7.6 m^3, only 8 of 96 failures, and 1,686 of 7,343 m^3 of erosion were related to roads and none were associated with the new roads. Based on 129 random erosion plots [Rice, 1996; Lewis, 1998] in the North Fork, the road erosion in EAG was 9.3 m^3ha^{-1}, compared to 34.5 m^3ha^{-1} for KJE and 16.6 m^3ha^{-1} for all roads in the North Fork. Thus it seems that the appearance of road cuts and fills in the model resulted from a spurious correlation. The *new* roads were relatively unimportant as a sediment source in the North Fork, probably because of their generally stable locations on upper hillslopes far from stream channels, the use of outsloping and frequent rolling-dips (drains), and negligible rainy season use.

Field evidence suggesting that unbuffered stream channels contributed to suspended sediment loads is more consistent. Channel reaches subjected to intense broadcast burns showed increased erosion from the loss of woody debris that stores sediment and enhances channel roughness. Annual surveys evaluating bank stability, vegetative cover, and sediment storage potential suggest the greatest sediment production and transport potential existed in the burned channel reaches. Bank disturbances from timber falling and yarding were evident in the unburned channels, but slash and residual woody debris provided both potential energy dissipation and sediment storage sites for moderating

sediment transport. Increased flows, accompanied by soil disruption and burning in headwater swales, may have accelerated channel headward expansion and soil pipe enlargements and collapses observed in watershed KJE [Ziemer, 1992] and in EAG, DOL, and LAN.

Based on 175 random 0.08-ha erosion plots in harvest areas [Rice, 1996; Lewis, 1998] in the North Fork, total erosion after logging in the burned watersheds EAG and GIB was 153 m^3ha^{-1} and 77 m^3ha^{-1}, respectively, higher than all other watersheds. Total erosion for the unburned clear-cut watersheds BAN, CAR, and KJE averaged 37 m^3ha^{-1}. These figures include estimates of sheet erosion, which is difficult to measure and may be biased towards burned areas because it was easier to see the ground where the slash had been burned. About 72% of EAG and 82% of GIB were judged to be thoroughly or intensely burned, and the remainder was burned lightly or incompletely. It is unknown how much of this hillslope erosion was delivered to stream channels, but the proportion of watershed burned was not a useful explanatory variable for suspended sediment transport. A plausible conclusion is that only burned areas in or adjacent to stream channels contributed appreciable amount of sediment to the streams.

The inventory of failures greater than 7.6 m^3 identified windthrow as another fairly important source of sediment. Of failures greater than 7.6 m^3, 68% were from windthrow. While these amounted to only 18% of the failure volume measured, 91% of them were within 15 m of a stream, and 49% were in or adjacent to a stream channel. Because of the proximity of windthrows to streams, sediment delivery from windthrow would be expected to be high. Windthrows are also important as contributors of woody debris to these channels, and play a key role in pool formation. Because woody debris traps sediment in transport, the net effect of windthrow on sediment transport can be either positive or negative. Woody debris inputs into the channel have been unusually high in the years since logging, partly because of a number of severe windstorms and partly because of the buffer strip design [Reid and Hilton, 1998]. While this has led to substantial bank cutting and channel reworking, the bulk of the increased sediment loads after logging watersheds BAN, CAR, EAG, and GIB has not yet reached the main stem stations FLY and ARF, much of it having been stored in reaches affected by blowdown [Lisle and Napoletano, 1998].

Cumulative effects. We have considered three types of information that the sediment models provide about the cumulative effects of logging activity on (unit area) suspended sediment loads. Keep in mind that the response being considered in all these questions is the suspended sediment load per unit watershed area for a given storm event and that watershed area was used in the model to represent distance downstream.

Question 1. **Were the effects of multiple disturbances additive in a given watershed?** This question may be answered partly by looking at the forms of the storm flow and sediment models. Analyses of residuals and covariance structures provide good evidence that the models are appropriate for the data, including the use of a logarithmic response variable. A logarithmic response implies a multiplicative effect for predictors that enter linearly and a power function for predictors that enter as logarithms. The flow response to logged area in model (10) is multiplicative, and the sediment response to flow increases in models (24) and (25) is a power function because Δq (equations (11), (12)) is equivalent to the log of a ratio. We next examine how much these relations differ from an additive relationship in the range of data we observed.

Consider $E(r_{ij})$, the expected value of the ratio between an observation and its expectation in an unlogged condition. From equations (9) and APPENDIX C, equations (35) and (36),

$$E(r_{ij}) = \exp\left[D_{ij}T_{ij}\right] \tag{26}$$

where $T_{ij} = \beta_4 + \beta_5 \log(y_{Cj}) + \beta_6 \log(w_j) + \beta_7 a_i$. The expected effect of combining two simultaneous disturbances D_1 and D_2 is

$$E(r_{1+2}) = \exp\left[(D_1 + D_2)T_{ij}\right] = E(r_1)E(r_2) \tag{27}$$

where $E(r_1) = \exp[D_1 T_{ij}]$ and $E(r_2) = \exp[D_2 T_{ij}]$ are the expected effects of the individual disturbances. The combined effect departs most from additive when $E(r_1) = E(r_2)$. For example, disturbances that individually would result in 10% and 30% increases in the response produce a combined increase of 43% ($1.10 \times 1.30 = 1.43$), while disturbances that individually would result in 20% increases, produce a greater combined increase of 44% ($1.20 \times 1.20 = 1.44$). If the disturbances were additive the combined increase would be a 40% increase in either case. For more than two disturbances, the departures from additivity can be somewhat greater. In general, multiple disturbances that have a combined effect of r on the response under a multiplicative model will result in a minimum increase of $\log(r)$ in the response under an additive model, where r is defined in the sense of r_{ij} above. (This results from a mathematical limit as the number of equal-magnitude disturbances contributing to the effect r becomes large.)

In the storm flow data, only the main-stem gaging stations received waters from multiple disturbances. The maximum observed increase in storm flow on any main stem gaging station was 118%, but 8 out of 10 increases were under 40% and the median increase was just 16%. Taking the logarithms of 2.18, 1.40, and 1.16, we find that multiple disturbances that could produce these increases in a multiplicative model would produce minimum increases of 78%, 34%, and 15%, respectively, under an additive model. Therefore, in the range of most of the data (increases less than 40%) the disturbance effect on storm flow is approximately additive.

Now we can evaluate the additivity of the disturbance effect on sediment load, since this is expressed mainly through Δq. For this evaluation we fit model {(25),(17),(18)}, but fixing the parameters involving road cuts and fills at zero. Under this model, analogously to equation (26) for the flow model, the expected value of the ratio between an observation and its expectation in an unlogged condition is given by

$$E(r_{ij}) = \exp\left[\beta_2 \Delta q_{ij}^{(2)}\right] = \exp\left[\beta_2 \log\left(\frac{y_{ij}}{y_{Cj}}\right)\right] = \left(\frac{y_{ij}}{y_{Cj}}\right)^{\beta_2} \tag{28}$$

The ratio of y_{ij} and y_{Cj}, the unit area flow volumes in storm j from the treated and control watersheds, is an expression of the increased flow related to tree removal. A plot of equation (28) using the maximum likelihood estimate of 1.514 for β_2 passes through (1,1) and is very nearly linear in the range $0.82 \le y_{ij}/y_{Cj} \le 1.92$, which includes 95% of the observations on the main-stem stations. It follows that the effect of flow on suspended sediment is approximately additive for stations which receive waters from multiple logging units. For example, a flow ratio of 1.40 corresponds to a 66% increase in sediment load, while a flow ratio of 1.80 corresponds to a 143% increase in sediment load. An additive flow effect would produce an increase of 66 + 66 = 132% in sediment load, not much less than 143%. Examples of smaller flow ratios deviate from additivity even less than this example.

So, in the range of data we observed, the effect of disturbance on flow is approximately additive, and the effect of flow on sediment loads is approximately additive. In summary, the mathematical approach indicates that the combined effect of multiple disturbances on sediment loads is very similar to the sum of the effects of the individual disturbances.

Question 2. Were downstream changes greater than would be expected from the proportion of area disturbed? This question was addressed by testing the coefficients of terms formed from the product of disturbance and watershed area. If the coefficient of this term were positive, it would imply that the effect of a given disturbance proportion increases with watershed size. The interactions of those disturbance measures that had explanatory utility in the sediment models were considered, including road cut and fill area and length of unbuffered stream channels. None of the product terms were found to have coefficients significantly greater than zero, indicating that suspended load increases were not disproportionately large in larger watersheds. To the contrary, the sum of the

observed sediment loads at the four main-stem stations were all within 25% of the sum of the loads predicted for undisturbed watersheds (Table 7). Channel cross-section measurements indicate 1040 metric tons of net filling in the main stem during the post-logging period [Lisle and Napolitano, 1998]. Much of the logging-related sediment from the tributaries has apparently been deposited in the main stem, especially in reaches affected by blowdowns and in alluvial bars near tributary confluences, and therefore has not reached downstream gages.

There is, however, one subwatershed where this second type of cumulative effect may be occurring. Watershed DOL, only 36% cut, includes the 100% cut watershed EAG, yet the percentage sediment increases have been similar (269% at DOL versus 238% at EAG). Several mechanisms appear to be responsible for the unexpectedly high loads at DOL. In the incised lower reach, bank failures and channel widening have occurred. In addition, a major stream diversion caused by a windthrow resulted in the formation of a major gully eroding 87 m^3 directly into the stream. Sediment is also being released from behind decaying logs that were placed in the channel for skidding by oxen during historic logging. Finally, all these processes would have been augmented by the increased storm flows that followed modern logging.

Question 3. Were sediment loads in the lower watershed elevated to higher levels than in the tributaries? Regardless of the control watersheds used, suspended sediment transport per unit watershed area tended to increase downstream before logging (Figure 21). This tendency may reflect a greater availability of fine sediment downstream in lower gradient channels. If unit area sediment loads increase downstream and result in water quality levels of concern with a smaller proportion of watershed disturbance than upstream locations, then cumulative effects may be said to have occurred, in the sense that activities producing acceptable local impacts resulted in impacts that are unacceptable by the same standard downstream.

To the extent that larger watersheds reflect average disturbance rates and therefore have smaller proportions of disturbance than the smallest disturbed watersheds upstream, one might expect sediment loads downstream to increase by less than those in the logged tributaries. In addition, as mentioned before, some of the sediment may be temporarily stored before reaching the lower stations. Indeed, in this study the post-treatment regression lines were much more similar among watersheds than the pretreatment lines, and the main-stem stations no longer transported the highest unit area sediment loads. However, larger watersheds will not necessarily behave the same way. For example, in geographically similar Redwood Creek in northwestern California, two main-stem gaging stations (175 km^2 and 720 km^2) yield higher sediment loads per unit area than three intensively logged tributaries [Lewis, 1998].

Cumulative effects considered in this paper were limited to a few hypotheses about water quality that could be statistically evaluated. But cumulative effects can occur in many ways. For example, resources at risk are often quite different in downstream areas, so an activity that has acceptable local impacts might have unacceptable offsite impacts if critical or sensitive habitat is found downstream. Different physical processes also tend to dominate upstream and downstream reaches. Channel aggradation may be the biggest problem downstream, while channel scour may be of concern upstream.

Subwatersheds and KJE anomaly. Analyses of the 5 clear-cut tributaries in the North Fork drainage show suspended load increases at all gaging stations located immediately below clear-cut units except at KJE, where loads have decreased. KJE had the highest pre-logging (1986-1989) unit area sediment loads of any of the tributaries (Figure 22), but, after logging, loads were similar to the other logged tributaries (Figure 17).

Prior to logging, the stream channel above KJE was unique. The KJE channel was an active gully with an abundant supply of sediment and the lowest gradient of any of the tributaries. After logging, the number of small debris jams doubled in the buffered channel above KJE, and further upstream the channel contained a large amount of logging debris and dense vegetative regrowth. Thus, opportunities for temporary sediment storage increased, and net energy available for sediment transport may have decreased, despite moderately increased flows, because of the increased channel roughness. The other tributaries were stable, vegetated, steep channels with limited sediment supplies and rela-

tively low unit area sediment loads prior to logging. In these tributaries the increased sediment introduced by logging was readily transported. While this explanation is speculative, response in sediment transport to a disturbance certainly will vary with channel morphology and the relative availability of sediment and energy.

CONCLUSIONS

The main conclusions from these analyses are:

- Models based upon the proportion of watershed area logged, an antecedent wetness index, time since logging, and the responses in unlogged control watersheds explained 95% of the variation in the logarithms of both storm discharge peaks and volumes. Goodness-of-fit is similar for pre-logging and post-logging data, and cross-validation indicates that the models were not overfit to the data.
- Storm discharge peaks and volumes after extended periods with little or no precipitation increased up to 300% and 400% respectively, but most increases were below 100%.
- The effect of logging on storm discharge peaks and volumes declines with increasing regional antecedent wetness, as indexed by a decay function of prior runoff at a control watershed. However, even under the wettest conditions of the study, increases in storm runoff from clear-cut watersheds averaged 23% for peaks and 27% for volumes.
- Relative increases in storm discharge peaks and volumes decline with storm size but were positive even in the largest storms of the study period.
- Average increases in annual storm runoff were 58% from 95-100% clear-cut watersheds and 23% from 30-50% clear-cut watersheds.
- Recovery rates in the first 4-7 years after logging are estimated to be 8% per year for peak flows and 9% per year for storm flow volumes.
- Effects of multiple disturbances on storm discharge peaks and volumes are approximately additive, and there is little evidence for magnification of effects downstream.
- Reduction in rainfall interception and transpiration by forest vegetation is the probable cause of increased storm discharge peaks and volumes following logging.
- Annual sediment loads increased 123-269% in the tributaries, but, at main-stem stations, increased loads were detected only in small storms and had little effect on annual sediment loads. At the North Fork weir, an increase of 89% was caused mainly by a landslide in an ungaged tributary that enters just above the weir.
- Much of the increased sediment load in North Fork tributaries was related to increased storm flow volumes. With flow volumes recovering as the forest grows back, flow-related increases in sediment load are expected to be short-lived.
- The effects of multiple disturbances on suspended loads in a watershed were approximately additive.
- In general, downstream suspended load increases were no greater than would be expected from the proportion of area disturbed. In one tributary, increased flows evidently impacted the channel in an uncut area downstream by mobilizing stored sediment and aggravating bank instabilities, but most of the increased sediment produced in the tributaries was apparently stored in the main stem and has not yet reached the main-stem stations.
- Before logging, sediment loads on the main stem were higher than on most tributaries. This was no longer the case after logging, apparently because sediment exported from tributaries was deposited at temporary storage sites, and smaller proportions of downstream watersheds were disturbed.
- Sediment increases in North Fork tributaries probably could have been reduced by avoiding activities that denude or reshape the banks of small drainage channels.
- Sediment loads are affected as much by channel conditions (e.g. organic debris, sediment storage sites, channel gradient, and width-to-depth ratio) as by sediment delivery from hillslopes.

APPENDIX A. Notation Used in the Text

a_i	Drainage area of watershed i
b_i	Estimate of parameter β_i
c_{ij}	Proportion of watershed i logged in water years prior to that of storm j, and
c'_{ij}	Proportion of watershed i logged prior to storm j but in the same water year
D_{ij}	Some measure of disturbance per unit area in watershed i at storm j
$d_{i_1 i_2}$	Distance between centroids of watersheds i_1 and i_2
K	Number of parameters estimated in a model
n	Number of observations used in an analysis
p_{ij}	True (unknown) percentage change in response of watershed i in storm j as a result of treatment
$\tilde{p}_{ij}$	"Observed" percentage change in response of watershed i in storm j based on a comparison of y_{ij} and $\hat{y}'_{ij}$
p_0	Percentage change in response, given an arbitrary vector $\mathbf{x}_0$
p_N	Significance level of a hypothesis test based on the normal distribution
$\Delta q_{ij}^{(1)}$	Residual from the flow model (3) containing only β_{0i} and β_{1i}
$\Delta q_{ij}^{(2)}$	Difference between the logarithms of flow in the treated and control watersheds
$\Delta q_{ij}^{(3)}$	Predicted change after logging in the logarithm of storm flow from eqn (10)
t_{ij}	Area-weighted mean cutting age (number of summers passed) in watershed i for areas logged in water years preceding that of storm j
w_j	Wetness index at start of storm j
$x_{ij}^{(1)}, x_{ij}^{(2)}$	Generic measures of unit area disturbance in watershed i at storm j
$\mathbf{x}_0$	Arbitrary vector of explanatory variables
y_{ij}	Unit area response at treated watershed i in storm j
y_{Cj}	Unit area response at control watershed in storm j
y'_{ij}	Unknown response at watershed i, if it had been left untreated, in storm j
$\hat{y}'_{ij}$	Estimate of y'_{ij}
β_{0i}, β_{1i}	Location parameters (slope and intercept) to be estimated for each watershed i
$\beta_0^{(1)}, \beta_0^{(2)}$	Parameters used to model β_{0i} as a function of a_i
$\rho_{i_1 i_2}$	Correlation between $\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$
$\sigma_{i_1}, \sigma_{i_2}$	Standard deviations of $\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$
ε_{ij}	Error or deviation of y_{ij} from model at treated watershed i in storm j
$\varepsilon_{i_1 j}, \varepsilon_{i_2 j}$	Errors for watersheds i_1 and i_2 in storm j
θ_i	Parameter in covariance model
$\hat{\theta}_i$	Estimate of parameter θ_i

APPENDIX B. Likelihood Function and Gradient

The model for the mean response can be written

$$\mathbf{u} = E(\mathbf{y}) = f(\boldsymbol{\beta}) \tag{29}$$

where $\mathbf{y}$ is an $n\times1$ response vector and $\boldsymbol{\beta}$ is a $p\times1$ vector of unknown parameters. The error, $\mathbf{e} = \mathbf{y} - \mathbf{u}$, is modelled as a multivariate normal variable depending on q parameters:

$$\begin{aligned} \mathbf{e} &\sim \mathbf{N}(0, \boldsymbol{\Sigma}) \\ \boldsymbol{\Sigma} &= \mathbf{G}(\boldsymbol{\theta}) \end{aligned} \tag{30}$$

where $\boldsymbol{\Sigma}$ is the $n\times n$ covariance matrix of $\mathbf{e}$ depending on $\boldsymbol{\theta}$, a $q\times1$ vector of unknown parameters. The elements of $\boldsymbol{\Sigma}$ are paramaterized by equations (15)-(18). The likelihood function and its logarithm are

$$\begin{aligned} L &= (2\pi)^{-n/2}\left|\boldsymbol{\Sigma}\right|^{-1/2}\exp\left[-\tfrac{1}{2}(\mathbf{y}-\mathbf{u})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{u})\right] \quad \text{and} \\ \ell &= \log(L) = -\tfrac{n}{2}\log(2\pi) - \tfrac{1}{2}\log\left|\boldsymbol{\Sigma}\right| - \tfrac{1}{2}(\mathbf{y}-\mathbf{u})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{u}) \end{aligned} \tag{31}$$

respectively, where $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$. The gradient consists of the partial derivatives of ℓ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$:

$$\begin{aligned} \mathbf{grad} &= \left(\frac{\partial\ell}{\partial\beta_1},\cdots,\frac{\partial\ell}{\partial\beta_p},\frac{\partial\ell}{\partial\theta_1},\cdots,\frac{\partial\ell}{\partial\theta_q}\right) \\ \frac{\partial\ell}{\partial\beta_i} &= \frac{\partial\mathbf{u}^T}{\partial\beta_i}\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{u}), \qquad i = 1,\cdots,p \\ \frac{\partial\ell}{\partial\theta_j} &= -\frac{1}{2}tr\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right) + \frac{1}{2}(\mathbf{y}-\mathbf{u})^T\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{u}), \qquad j = 1,\cdots,q \end{aligned} \tag{32}$$

in which $tr(\cdot)$ refers to the trace (sum of the diagonal elements) of the matrix. The partial derivatives, $\partial\mathbf{u}^T/\partial\beta_i$ and $\partial\boldsymbol{\Sigma}/\partial\theta_j$, are model-specific and can be derived from equations (10) and (14)-(18).

APPENDIX C. An Unbiased Estimator, and Confidence and Prediction Intervals for Percentage Change in Response

Let y_0 be the response given an arbitrary predictor vector $\mathbf{x}_0$ and let y_0' be the unknown response for the same storm assuming the watershed were undisturbed. A prediction interval is sought for $p_0 = 100\left[y_0/E(y_0') - 1\right]$, the percentage change in response, and an unbiased estimator and confidence interval are sought for its expectation, $E(p_0)$. It will be convenient to obtain the unbiased estimator and confidence interval first. Since $\log(y_0)$ and $\log(y_0')$ are assumed to be normally distributed,

$$\begin{aligned} E(y_0) &= \exp\left[E\left(\log(y_0)\right) + \tfrac{1}{2}\sigma^2\right] \qquad \text{and} \\ E(y_0') &= \exp\left[E\left(\log(y_0')\right) + \tfrac{1}{2}\sigma^2\right] \end{aligned} \tag{33}$$

Let us denote the ratio of the actual response to its expected undisturbed value by

$$r_0 = \frac{y_0}{E(y_0')} \tag{34}$$

Its expectation is

$$\begin{aligned} E(r_0) &= \frac{E(y_0)}{E(y_0')} \\ &= \frac{\exp\left[E\left(\log(y_0)\right) + \frac{1}{2}\sigma^2\right]}{\exp\left[E\left(\log(y_0')\right) + \frac{1}{2}\sigma^2\right]} \\ &= \exp\left[f_0(\boldsymbol{\beta})\right] \end{aligned} \tag{35}$$

where, for the runoff models (10),

$$f_0(\boldsymbol{\beta}) = \left[\left(1-\beta_2(t_0-1)\right)c_0 + \beta_3^{(k)}c_0'\right] \times \left[\ \beta_4 + \beta_5\log(y_{C0})\ + \beta_6\log(w_0) + \beta_7 a_0\right] \tag{36}$$

Since $\mathbf{b}$, the vector of estimates for $\boldsymbol{\beta}$, is asymptotically distributed normal, we have that $f_0(\mathbf{b})$ is asymptotically distributed normal with $E[f_0(\mathbf{b})] = f_0(\beta)$ and unknown variance σ_*^2 [Bishop et al., 1975]. In shorthand, $f_0(\mathbf{b}) \sim N\left(f_0(\boldsymbol{\beta}), \sigma_*^2\right)$ for large samples. The variance σ_*^2 may be approximated using the delta method [Bishop et al., 1975]:

$$\tilde{\sigma}_*^2 = \sum_{i=1}^{p}\sum_{j=1}^{p} \frac{\partial f_0}{\partial b_i}\frac{\partial f_0}{\partial b_j} Cov\left[b_i, b_j\right] \tag{37}$$

The covariances are estimated by the elements of the inverted information matrix [McCullagh and Nelder, 1989]. The information matrix is the negative of the matrix of second derivatives (Hessian) of ℓ with respect to the parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$.

Let us introduce an estimator $\hat{r}_0 = \exp\left[f_0(\mathbf{b}) - \frac{1}{2}\sigma_*^2\right]$. Its expected value is

$$\begin{aligned} E(\hat{r}_0) &= \exp\left(-\tfrac{1}{2}\sigma_*^2\right) E\left\{\exp\left[f_0(\mathbf{b})\right]\right\} \\ &= \exp\left(-\tfrac{1}{2}\sigma_*^2\right) \exp\left\{E\left[f_0(\mathbf{b})\right] + \tfrac{1}{2}\sigma_*^2\right\} \\ &= \exp\left\{E\left[f_0(\mathbf{b})\right]\right\} \\ &= \exp\left\{f_0(\boldsymbol{\beta})\right\} \\ &= E(r_0) \end{aligned} \tag{38}$$

Hence $\hat{r}_0$ is an asymptotically unbiased estimator for $E(r_0)$, and $100(\hat{r}_0 - 1)$ is an asymptotically unbiased estimator for $100(E(r_0) - 1) = E(p_0)$. In practice, because σ_* is unknown, we replace it with $\tilde{\sigma}_*$ in the expression for $\hat{r}_0$.

Next we will compute a confidence interval for $E(r_0)$, and convert it to a confidence interval for $E(p_0)$. A 100(1-α)% confidence interval for $f_0(\beta)$ is defined by the probability

$$\Pr\left[f_0(\mathbf{b}) - z_{\alpha/2}\sigma_* \le f_0(\boldsymbol{\beta}) \le f_0(\mathbf{b}) + z_{\alpha/2}\sigma_*\right] = 1 - \alpha \tag{39}$$

where $z_{\alpha/2}$ is the $\alpha/2$ cutoff point of the standard normal distribution. Applying the monotone transformation **exp** to all sides of the inequality yields a confidence interval for $E(r_0)$:

$$\Pr\left[\exp\left[f_0(\mathbf{b}) - z_{\alpha/2}\sigma_*\right] \le E(r_0) \le \exp\left[f_0(\mathbf{b}) + z_{\alpha/2}\sigma_*\right]\right] = 1 - \alpha \tag{40}$$

Noting that $E(p_0) = 100(E(r_0) - 1)$, the above confidence interval is readily transformed into a confidence interval for $E(p_0)$.

$$100(1-\alpha)\,C.I.\text{for } E(p_0)\text{: } 100\left[\exp\left(f_0(\mathbf{b}) \pm z_{\alpha/2}\sigma_*\right) - 1\right] \tag{41}$$

Since σ_* is unknown, we replace it with $\tilde{\sigma}_*$.

Finally, we will compute a prediction interval for r_0, and convert it to a prediction interval for p_0. Using model (10) and (33), we find

$$\begin{aligned} r_0 &= \frac{y_0}{E(y_0')} \\ &= \frac{\exp\left[\beta_0 + \beta_1 \log(y_{C0}) + f_0(\boldsymbol{\beta}) + \varepsilon_0\right]}{\exp\left[\beta_0 + \beta_1 \log(y_{C0}) + \frac{1}{2}\sigma^2\right]} \\ &= \exp\left[f_0(\boldsymbol{\beta}) + \varepsilon_0 - \tfrac{1}{2}\sigma^2\right] \end{aligned} \tag{42}$$

Since $\varepsilon_0 \sim N(0, \sigma^2)$ and, asymptotically, $f_0(\mathbf{b}) \sim N\left(f_0(\boldsymbol{\beta}), \sigma_*^2\right)$, and they are independent random variables, it follows that $f_0(\mathbf{b}) - \varepsilon_0 \sim N\left(f_0(\boldsymbol{\beta}), \sigma_*^2 + \sigma^2\right)$. Thus

$$\Pr\left[f_0(\mathbf{b}) - z_{\alpha/2}\left(\sigma_*^2 + \sigma^2\right)^{\frac{1}{2}} \le f_0(\boldsymbol{\beta}) + \varepsilon_0 \le f_0(\mathbf{b}) - z_{\alpha/2}\left(\sigma_*^2 + \sigma^2\right)^{\frac{1}{2}}\right] = 1 - \alpha \tag{43}$$

Subtracting $0.5\sigma^2$ and applying the monotone transformation **exp** to all parts of the inequality converts the middle term to r_0, yielding the following prediction interval:

$$100(1-\alpha)\,P.I.\text{for } r_0\text{: } \exp\left(f_0(\mathbf{b}) - \tfrac{1}{2}\sigma^2 \pm z_{\alpha/2}\left(\sigma_*^2 + \sigma^2\right)^{\frac{1}{2}}\right) \tag{44}$$

which is readily transformed to a prediction interval for p_0:

$$100(1-\alpha)\,P.I.\text{for } p_0\text{: } 100\left[\exp\left(f_0(\mathbf{b}) - \tfrac{1}{2}\sigma^2 \pm z_{\alpha/2}\left(\sigma_*^2 + \sigma^2\right)^{\frac{1}{2}}\right) - 1\right] \tag{45}$$

Since σ_* and σ are unknown, we replace them with $\tilde{\sigma}_*$ and $\hat{\sigma} = \hat{\theta}_3 a_0^{\hat{\theta}_4}$, where a_0 is the watershed area.

Confidence and prediction intervals for sediment models (24) and (25) are similar, but $f_0(\mathbf{b})$ is replaced by the linear functions $g_0(\mathbf{b})$ and $h_0(\mathbf{b})$, respectively, where

$$g_0(\mathbf{b}) = \beta_2 \Delta q_0^{(1)} + \beta_3 x_0^{(1)} + \beta_4 x_0^{(2)} + \beta_5 (x_0^{(1)} + x_0^{(2)}) \log(y_{(\mathrm{HI})0}) + \beta_6 (x_0^{(1)} + x_0^{(2)}) a_0 \tag{46}$$

$$\text{and} \quad h_0(\mathbf{b}) = \beta_2 \Delta q_0^{(2)} + \beta_3 x_0 + \beta_4 x_0 a_0 \tag{47}$$

Since these functions are linear, the delta method yields the exact variance, but, as before, the covariance matrix of **b** must be estimated from the observed information matrix, so σ_*^2 is still only known approximately.

Acknowledgments. This research was a result of a cooperative effort of the California Department of Forestry and Fire Protection and the USDA Forest Service, Pacific Southwest Research Station. The study design was a result primarily of the efforts of Raymond Rice, who was the Principal Investigator until his retirement in 1989. Robert Thomas designed the SALT sampling algorithm and Rand Eads designed the hardware/software interface used in the field implementation. In addition, the authors are grateful to Dave

Thornton and the many individuals, too numerous to mention here, who spent thousands of hours in the field (often during hazardous storm conditions), in the laboratory, and in the office to ensure the best possible data quality.

Detailed hydrologic and climatic data collected at Caspar Creek between 1963 and 1997 are available on compact disk from the authors.

REFERENCES

Bailey, E. H., W. P. Irwin, and D. L. Jones, Franciscan and related rocks, and their significance in the geology of western California, *Calif. Div. Mines Geol. Bull., 183*, 177 pp., 1964.

Baskerville, G. L., Use of logarithmic regression in the estimation of plant biomass, *Can. J. Forest Res., 2*, 49-53, 1972.

Beschta, R. L., M. R. Pyles, A. E. Skaugset, and C. G. Surfleet, Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon: An alternative analysis, unpublished report supplied by author, Department of Forest Engineering, Oregon State University, Corvallis, 1997.

Bishop, Y. M., S. E. Fienberg, and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, Mass., 557 pp., 1975.

Brown, G. W. and J. T. Krygier, Clear-cut logging and sediment production in the Oregon Coast Range, *Water Resour. Res. 7*(5):1189-1198, 1971.

Burnham, K. P. and D. R. Anderson, *Model Selection and Inference: A Practical Information-Theoretic Approach,* Springer, New York, 353 pp., 1998.

Chow, G. C., A test of equality between sets of observations in two linear regressions, *Econometrica, 28*, 591-605, 1960.

Cleveland, W. S., Robust locally weighted regression and smoothing scatterplots, *J. Amer. Stat. Assoc., 74*, 829-836, 1979.

Copas, J. B., Regression, prediction and shrinkage, *J. R. Statist. Soc. B, 45*(3): 311-354, 1983.

Efron, B. and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 436 pp., 1993.

Fredricksen, R. L., Sedimentation after logging road construction in a small western Oregon watershed, in *Proceedings of the Federal Inter-Agency Sedimentation Conference, 1963*, Agricultural Research Service Misc. Pub. No. 970, pp. 56-59, 1963.

Grant, G. E. and A. L. Wolff, Long-term patterns of sediment transport after timber harvest, western Cascade Mountains, Oregon, USA, *IAHS Publ. 203*, 31-40, 1991.

Hansen, M. H. and W. N. Hurwitz, On the theory of sampling from finite populations, *Ann. Math. Stat., 14*, 333-362, 1943.

Harr, R. D., Forest practices and streamflow in western Oregon, *Gen. Tech. Rep. PNW-49*, 18 pp., For. Serv., U.S. Dep. Of Agric., Portland, Oreg., 1976.

Harr, R. D., W. C. Harper, J. T. Krygier, and F. S. Hsieh, Changes in storm hydrographs after road building and clear-cutting in the Oregon Coast Range, *Water Resour. Res., 11*, 436-444, 1975.

Harr, R. D., R L. Fredriksen, and J. Rothacher, Changes in streamflow following timber harvest in southwestern Oregon, *Res. Paper PNW-249*, 22 pp., For. Serv., U.S. Dep. Of Agric., Portland, Oreg., 1979.

Harris, D. D., Hydrologic changes after logging in two small Oregon coastal watersheds, *U.S. Geol. Surv. Water Supply Paper 2037*, 31 pp., 1977.

Henry, N., Overview of the Caspar Creek watershed study, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 1-9, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

Hewlett, J. D. and A. R. Hibbert, Factors affecting the response of small watersheds to precipitation in humid areas, in *Forest Hydrology*, edited by W. E. Sopper and H. W. Lull, pp. 275-290, Pergamon, New York, 1967.

Jones, J. A. and G. E. Grant, Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon, *Water Resour. Res., 32*, 959-974, 1996.

Keppeler, E. T., R. R. Ziemer, and P. H. Cafferata, Changes in soil moisture and pore pressure after harvesting a forested hillslope in northern California, in *Effects of Human-induced Changes on Hydrologic Systems*, edited by R. A. Marston, and V. R. Hasfurther, pp. 205-214, American Water Resources Association, Herndon, Virginia, 1994.

Keppeler, E. T. and D. Brown, Subsurface drainage processes and management impacts, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 25-34, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

Keppeler, E. T. and R. R. Ziemer, Logging effects on streamflow: water yields and summer flows at Caspar Creek in northwestern California. *Water Resour. Res., 26*(7)*,* 1669-1679, 1990.

Leaf, C. F., Sediment yields from central Colorado snow zone, *J. Hydraulics Division, ASCE Proc. 96*(HY1): 87-93, 1970.

Lewis, J., Turbidity-controlled suspended sediment sampling for runoff-event load estimation, *Water Resour. Res., 32*(7), 2299-2310, 1996.

Lewis, J., Evaluating the impacts of logging activities on erosion and suspended sediment transport in the Caspar Creek watersheds, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 55-69, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

Lisle, T. E. and M. B. Napolitano, Effects of recent logging on the main channel of North Fork Caspar Creek, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 81-85, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

McCullagh, P. and J. A. Nelder, *Generalized Linear Models*, 2nd edn., Chapman & Hall, London, 1989.

Megahan, W. F., J. G. King, and K. A. Seyedbagheri, Hydrologic and erosional responses of a granitic watershed to helicopter logging and broadcast burning, *Forest Science, 41*(4), 777-795, 1995.

Miller, R. G., *Simultaneous Statistical Inference*, 2nd edn., Springer-Verlag, New York, 299 pp., 1981.

Miller, E. L., Sediment yield and storm flow response to clear-cut harvest and site preparation in the Ouachita Mountains, *Water Resour. Res. 20*(4): 471-475, 1984.

Mood, A. M., F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd edn., McGraw-Hill, New York, 564 pp., 1974.

Napolitano, M. B., Persistence of historical logging impacts on channel form in mainstem North Fork Caspar Creek, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 97-101, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

Norick, N. X., SALT (selection at list time) sampling: a new method and its application in forestry. M.A. Thesis, 50 pp., Univ. of California, Berkeley, 1969.

Olive, L. J. and W. A. Rieger, Assessing the impact of land use change on stream sediment transport in a variable environment, *IAHS Publ. 203*, 73-81, 1991.

O'Loughlin, C. L., L. K. Rowe, and A. J. Pearce, Sediment yield and water quality responses to clearfelling of evergreen mixed forests in western New Zealand, *IAHS Publ. 130*, 285-292, 1980.

Plamondon, A. P., Logging and suspended sediments input in small streams: concentration, origin, and duration, presented at *XVII IUFRO World Congress, Kyoto, Japan*, Sept. 1981.

Rao, C. R. *Linear Statistical Inference and its Applications*, 2nd edn., Wiley and Sons, New York, 625 pp., 1973.

Reid, L. M. and S. Hilton, Buffering the buffer, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 71-80, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

Rice, R. M., F. B. Tilley, and P. A. Datzman, A watershed's response to logging and roads: South Fork of Caspar Creek, California, 1967-1976. *Res. Pap. PSW-146*, 12 pp., For. Serv., U.S. Dep. Of Agric., Berkeley, Calif.,1979.

Rice, R. M., Sediment delivery in the North Fork of Caspar Creek, unpublished final report, Agreement CA94077, Calif. Dep. of For. and Fire Prot., Sacramento, 1996.

Rothacher, J., Net precipitation under a Douglas-fir forest. *Forest Science*, *4*, 423-429, 1963.

Rothacher, J., Regimes of streamflow and their modification by logging, in *Proceedings of the Symposium on Forest Land Use and Stream Environment*, edited by J. T. Krygier and J. D. Hall, pp. 55-63, Oregon State Univ., Corvallis, 1971.

Rothacher, J., Does harvest in west slope Douglas-fir increase peak flow in small streams? Pacific Northwest Forest and Range Experiment Station, *Res. Pap. PNW-163*, 13 pp., For. Serv., U.S. Dep. of Agric., Portland, Oreg., 1973.

Shumway, R. H., *Applied Statistical Time Series Analysis*, Prentice Hall, Englewood Cliffs, N. J., 379 pp., 1988.

Statistical Sciences, S-Plus Version 3.3 for Unix, StatSci, a division of MathSoft, Inc., Seattle, 1995.

Thomas, R. B., Estimating total suspended sediment yield with probability sampling, *Water Resour. Res*., *21*(9), 1381-1388, 1985.

Thomas, R. B., Piecewise SALT sampling for estimating suspended sediment yields, *Gen. Tech. Rep., PSW-114*, 11 pp., For. Serv., U.S. Dep. Of Agric., Berkeley, Calif., 1989.

Thomas, R. B. and J. Lewis, A comparison of selection at list time and time-stratified sampling for estimating suspended sediment load, *Water Resour. Res. 29*(4), 1247-1256, 1993.

Thomas, R. B. and J. Lewis, An evaluation of flow-stratified sampling for estimating suspended sediment loads, *J. Hydrol.*, *170*, 27-45, 1995.

Thomas, R. B. and W. F. Megahan., Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon: A second opinion, *Water Resour. Res., 34*(12):3393-3403, 1998.

Wilson, A. L., When is the chow test UMP? *The American Statistician, 32*(2), 66-68, 1978.

Wright, K. A., K. H. Sendek, R. M. Rice, and R. B. Thomas, Logging effects on streamflow: storm runoff at Caspar Creek in northwestern California, *Water Resour. Res., 26*(7), 1657-1667, 1990.

Ziemer, R. R., Soil moisture depletion patterns around scattered trees, *Res. Note PSW-166*, 13 pp., For. Serv., U.S. Dep. of Agric., Berkeley, Calif., 1968.

Ziemer, R. R., Stormflow response to roadbuilding and partial cutting in small streams of northern California, *Water Resour. Res., 17*(4), 907-917, 1981.

Ziemer, R. R., Effect of logging on subsurface pipeflow and erosion: coastal northern California, USA., *IAHS Publ. 209*, 187-197, 1992.

Ziemer, R. R., Flooding and stormflows, in Proceedings, Conference on Coastal Watersheds: The Caspar Creek Story, *Gen. Tech. Rep. PSW-168,* pp. 15-24, For. Serv., U.S. Dep. Of Agric., Albany, Calif., 1998.

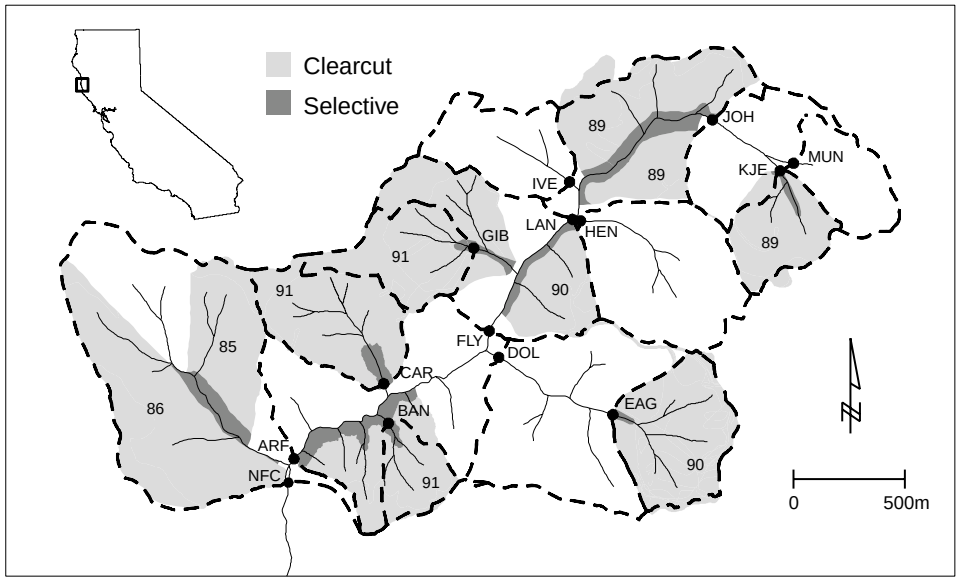

Figure 1. North Fork Caspar Creek. Gaging stations are identified by 3-letter abbreviations and dots, subwatershed boundaries by dashed lines, and logged areas by shading. Inset locates Caspar Creek within California.

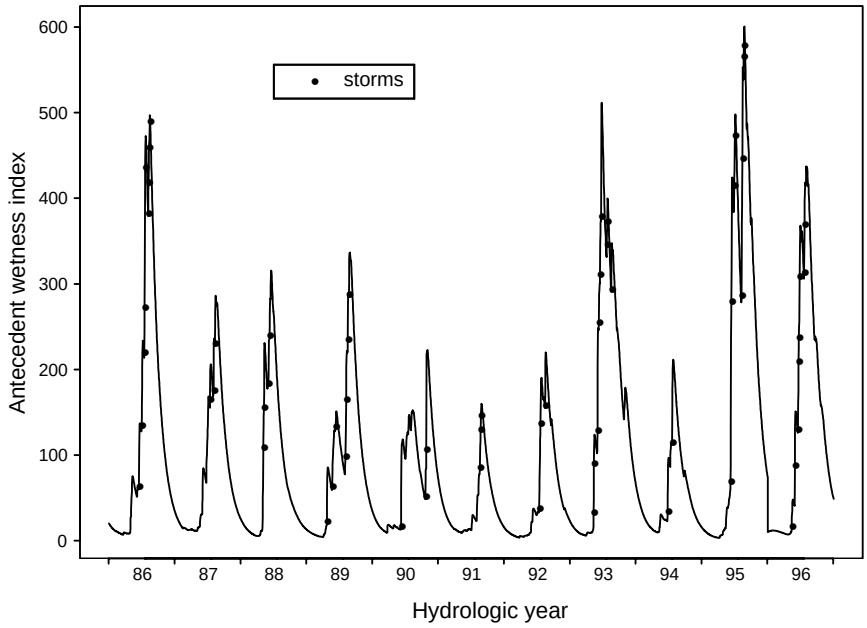

Figure 2. Antecedent wetness index (equation (2)) and temporal distribution of storms for the period of study (1986-1996). Solid circles indicate the wetness level at the start of each storm.

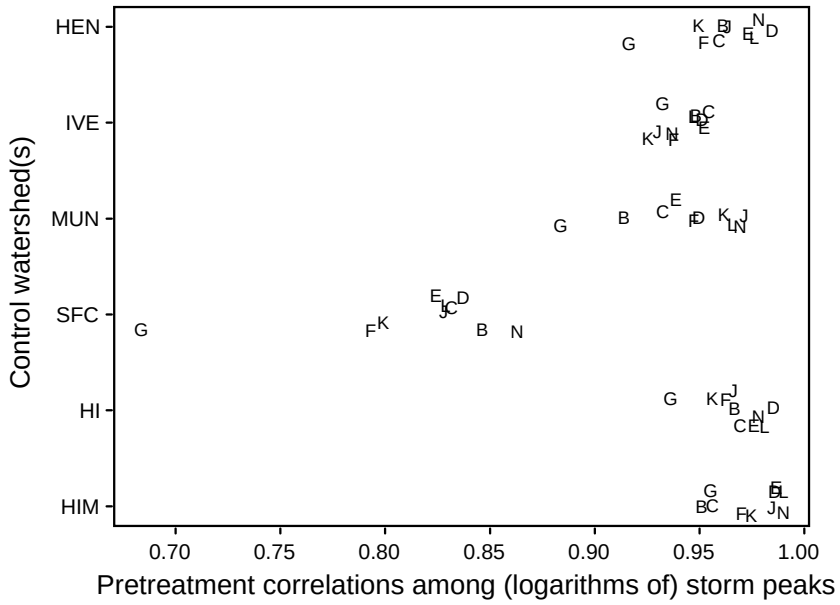

Figure 3. Pretreatment correlations between logarithms of storm peak at treated watersheds and alternative control watersheds. Letters designate watersheds (e.g. G is watershed GIB). Random noise has been added to the vertical plotting positions to improve readability.

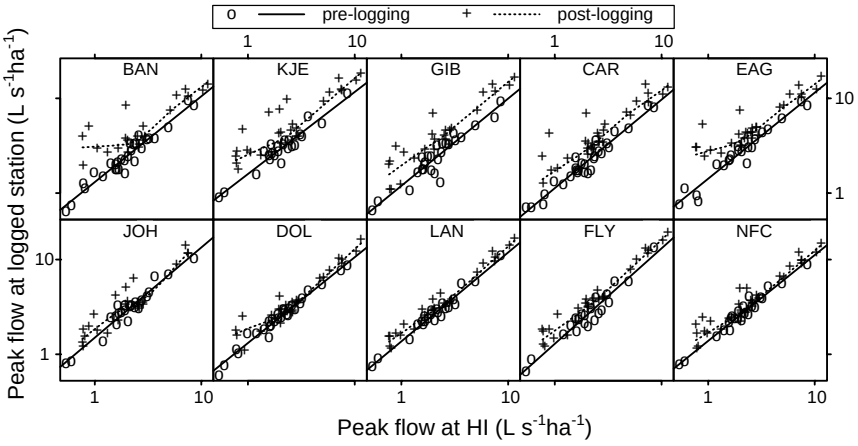

Figure 4. Relation between peak streamflow in the 10 treated tributaries in the North Fork of Caspar Creek, and that of the HI control. Post-logging relations were fitted by locally weighted regression [Cleveland, 1979]. The top row represents 95-100% clear-cut watersheds.

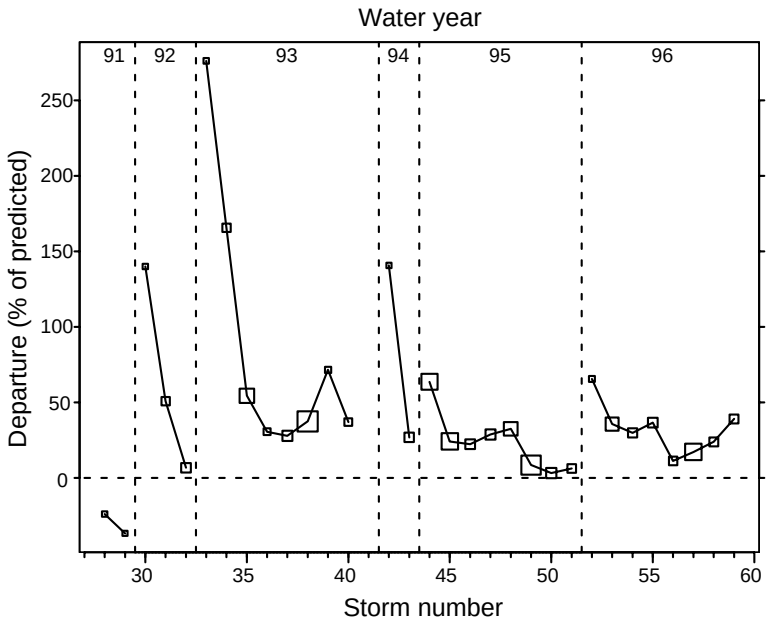

Figure 5. Post-logging departures of storm peaks (as percentage of predicted) at watershed EAG from those predicted from pretreatment regression on HI control. Axes are logarithmic. Symbol sizes indicate relative size of storm peak at HI control. Vertical dotted lines separate water years. About half the watershed was winter-logged before storm 28 and logging was completed by storm 30.

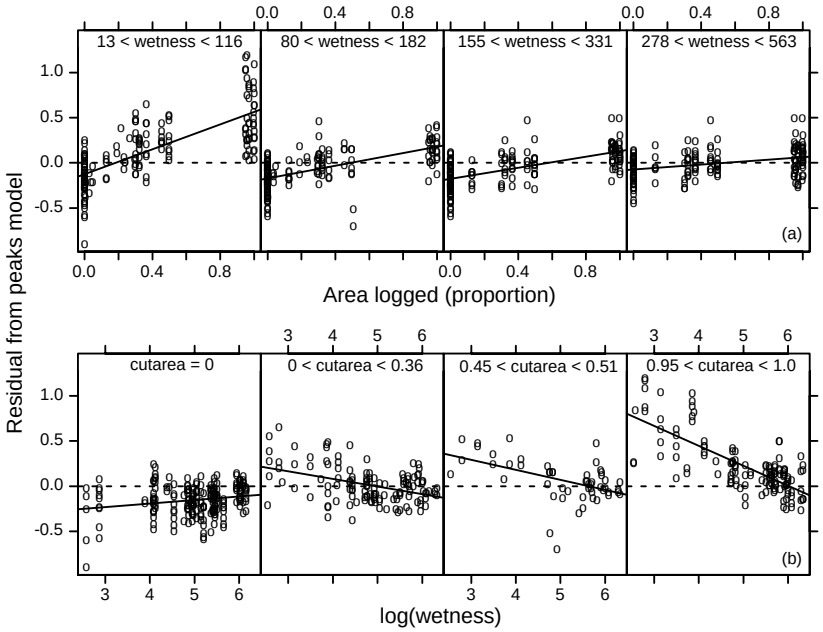

Figure 6. Conditioning plots of residual from storm peaks model (3) and interaction between area logged and antecedent wetness index with (a) wetness index fixed in each frame, and (b) proportion of area logged fixed in each frame.

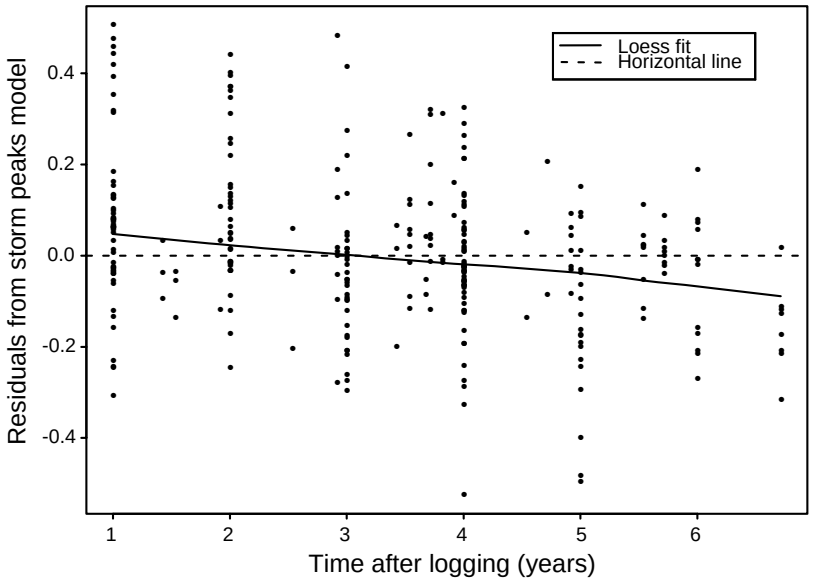

Figure 7. Relation between storm peak residuals and time after logging. Curve is fit by loess method [Cleveland, 1979]. Residuals are from least squares fit to the model $\log(y_{ij}) = \beta_{0i} + \beta_{1i} \log(y_{Cj}) + \beta_4 D_{ij} + \beta_6 D_{ij} \log(w_j) + \varepsilon_{ij}$.

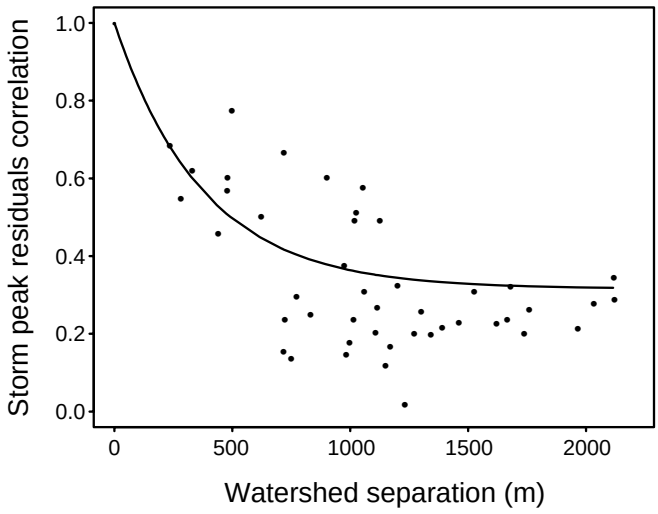

Figure 8. Relation between storm peak residuals correlation and distance between watershed centroids. Residuals are from maximum likelihood fit to storm peak model {(10),(16),(18)}. Curve depicts equation (16), with estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$.

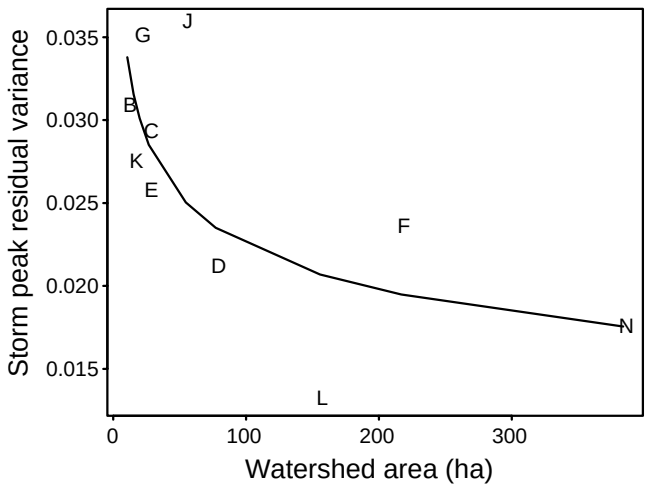

Figure 9. Relation between variance of storm peak residuals and watershed area. Residuals are from maximum likelihood fit to storm peak model {(10),(16),(18)}. Curve depicts equation (18) with estimated parameters $\hat{\theta}_3$ and $\hat{\theta}_4$. Letters designate watersheds (e.g. G is watershed GIB).

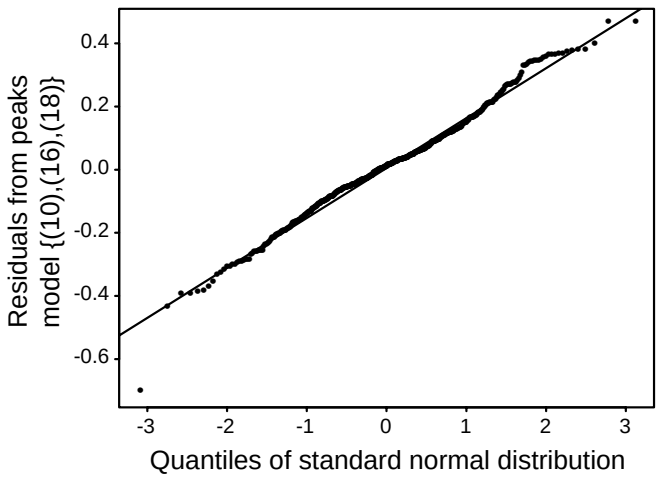

Figure 10. Normal quantile plot of residuals from storm peak model {(10),(16),(18)}. Line is least squares fit.

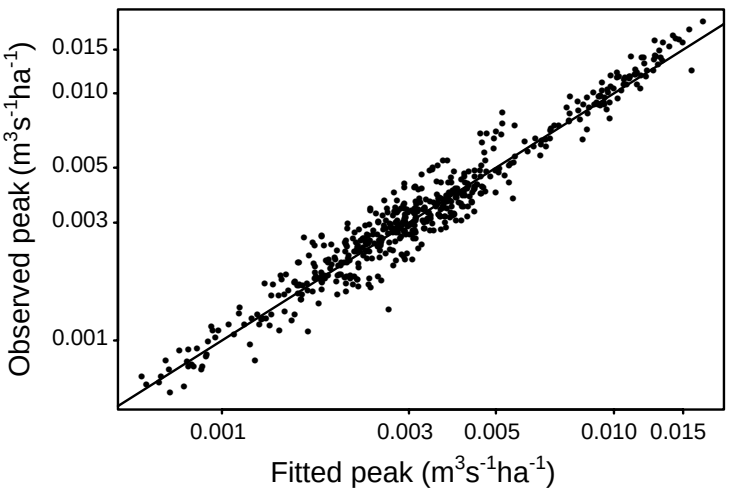

Figure 11. Observed storm peaks versus fitted values from model {(10),(16),(18)}. Line is $y = x$.

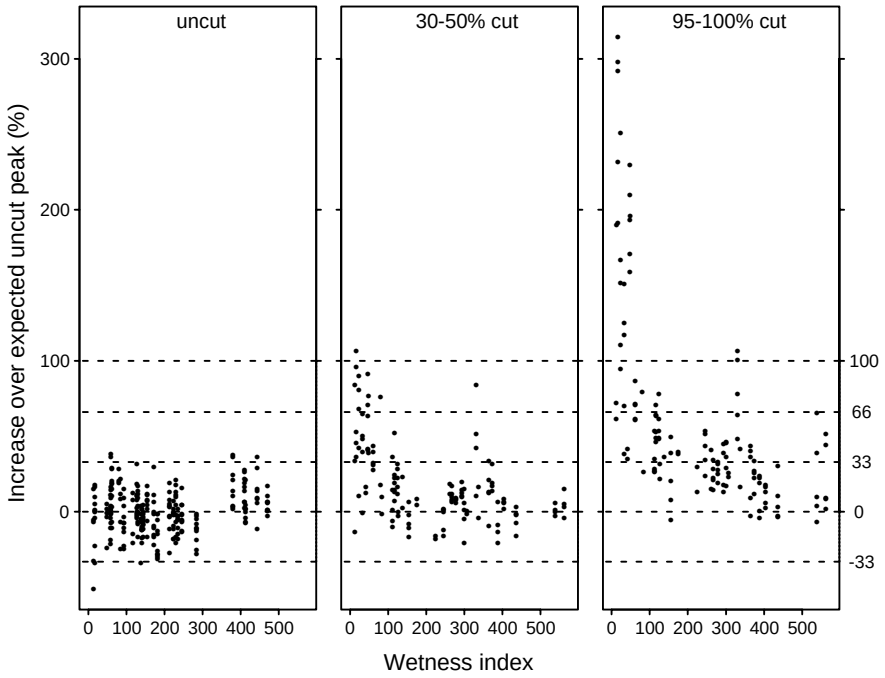

Figure 12. Percentage increase over expected uncut storm peak as related to antecedent wetness index for uncut (before treatment), partly (30-50%) clear-cut, and (95-100%) clear-cut watersheds. Bias-corrected predictions are from model {(10),(16),(18)} with disturbance set to zero.

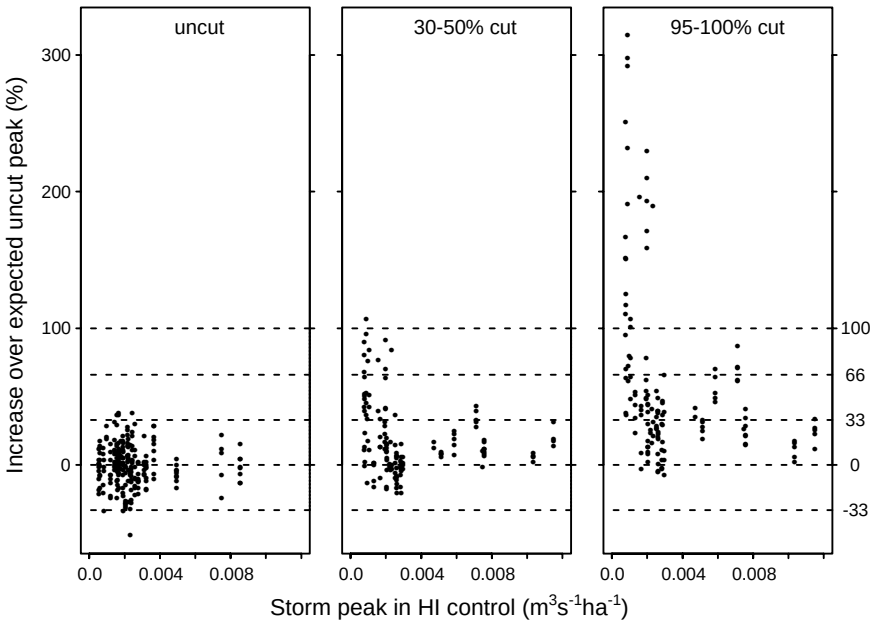

Figure 13. Percentage increase over expected uncut storm peak as related to peak size in the HI control for uncut (before treatment), partly (30-50%) clear-cut, and (95-100%) clear-cut watersheds. Bias-corrected predictions are from model {(10),(16),(18)} with disturbance set to zero.

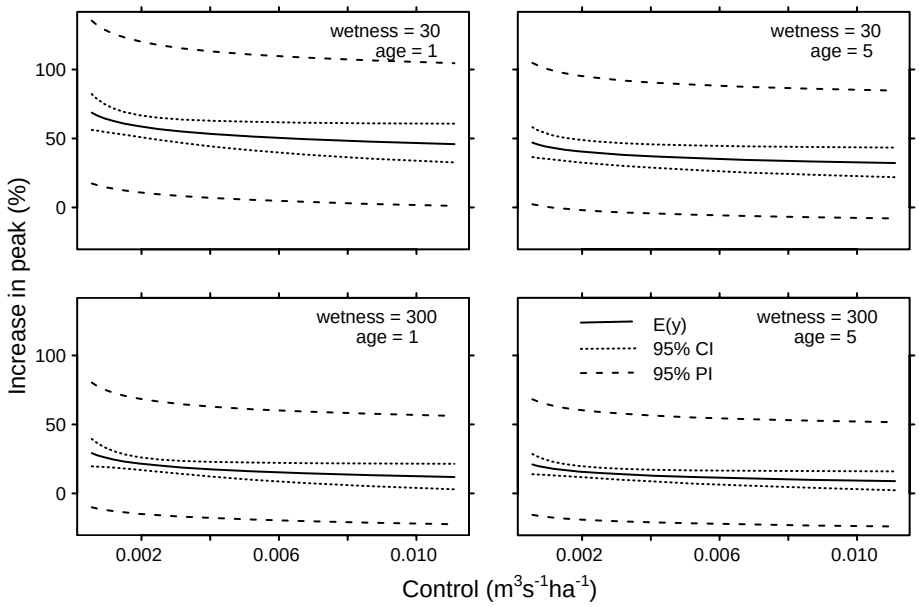

Figure 14. The effect of wetness and age after cutting on predictions from storm peak model {(10), (16),(18)} after clear-cutting 50% of a 20 ha watershed. Expected increases and 95% confidence (CI) and prediction (PI) intervals are shown for two levels of antecedent wetness 1 and 5 years after cutting.

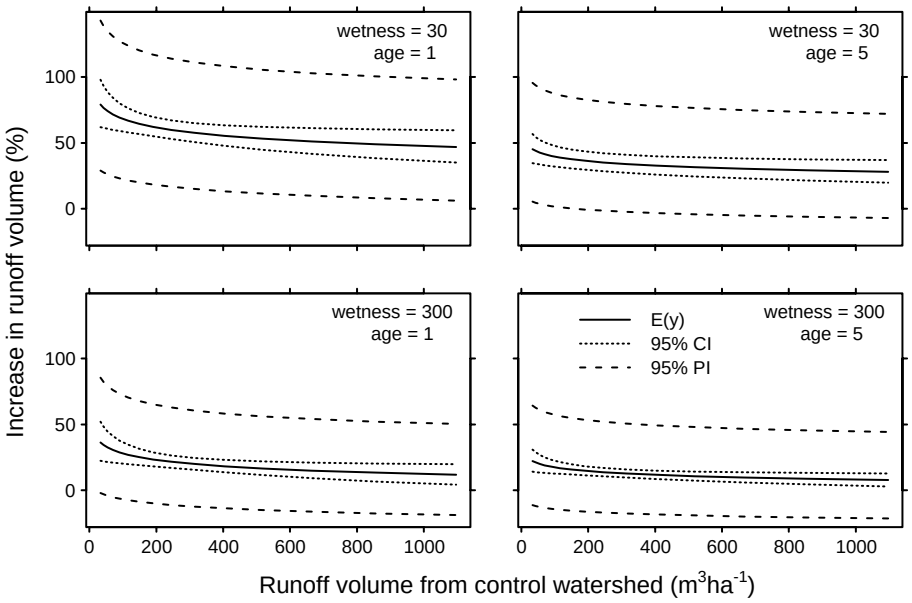

Figure 15. The effect of wetness and age after cutting on predictions from storm runoff volume model {(10),(17),(18)}, after clear-cutting 50% of a 20 ha watershed. Expected increases and 95% confidence (CI) and prediction (PI) intervals are shown for two levels of antecedent wetness 1 and 5 years after cutting.

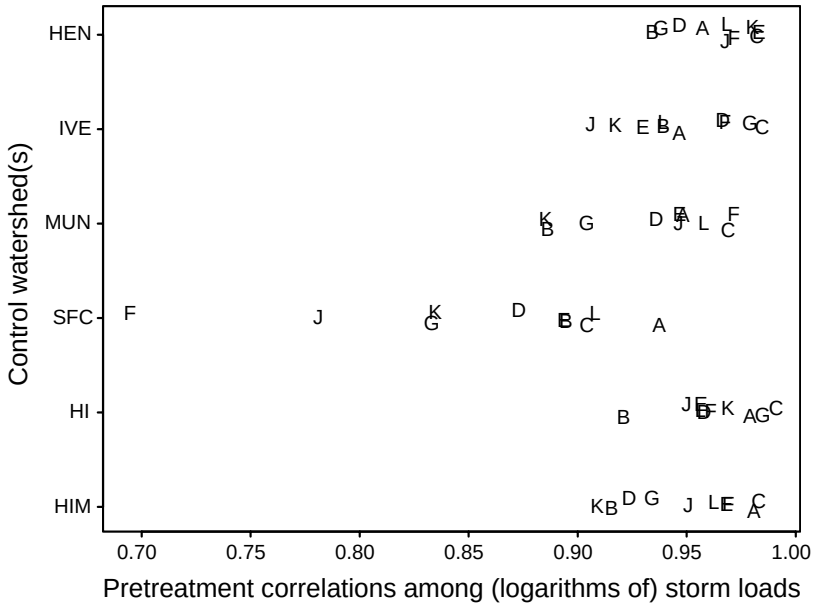

Figure 16. Pretreatment correlations between logarithms of storm sediment load at treated watersheds and alternative control watersheds. Letters designate watersheds (e.g. G is watershed GIB). Random noise has been added to the vertical plotting positions to improve readability.

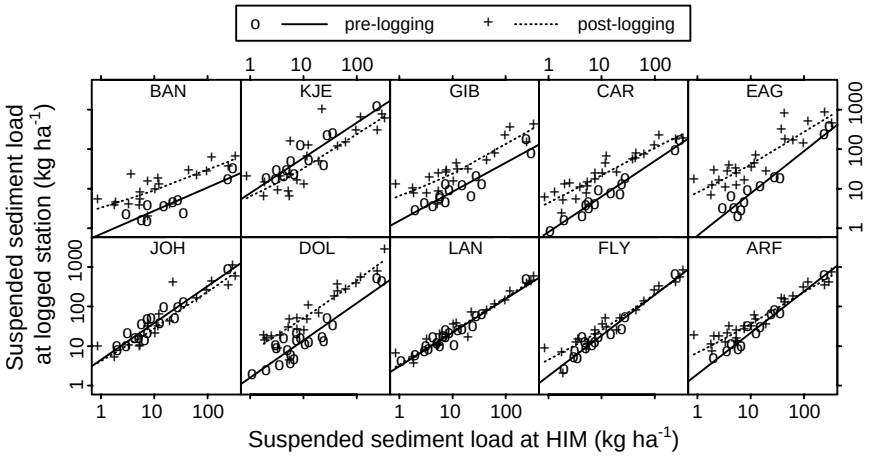

Figure 17. Relations between storm suspended sediment loads at logged subwatersheds in the North Fork and the the HIM control from 1986 to 1995. Post-logging relations were fitted by loess method [Cleveland, 1979]. The top row represents 95-100% clear-cut watersheds.

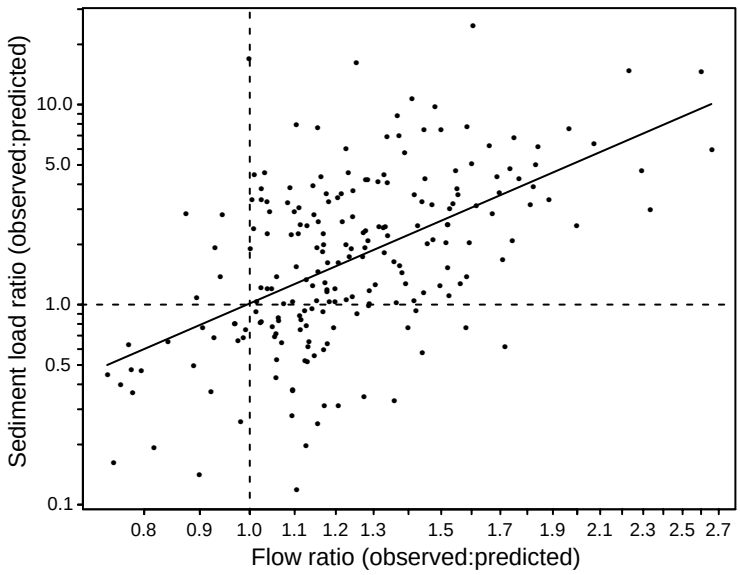

Figure 18. Relation between post-treatment sediment load departures from pretreatment relationship (3) and flow departures $\Delta q_{ij}^{(1)}$. Departures are expressed as the ratio of observed to predicted response.

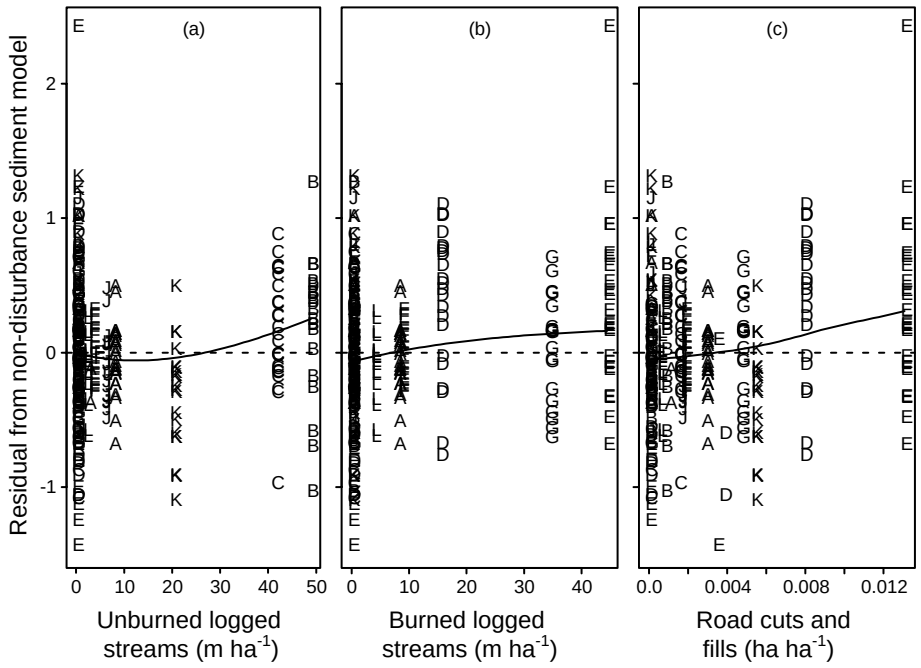

Figure 19. Relation between sediment load residuals and disturbance per unit watershed area. Curves are fit by loess method [Cleveland, 1979] to least squares residuals from the model: $\log(y_{ij}) = \beta_{0i} + \beta_{1i}\log(y_{(\mathrm{H})j}) + \beta_2 \Delta q_{ij}^{(1)} + \varepsilon_{ij}$. Disturbance variables shown are (a) length of stream in burned clear-cut areas, (b) length of stream in unburned clear-cut areas, and (c) road cut and fill area. Letters designate watersheds (e.g. G is watershed GIB).

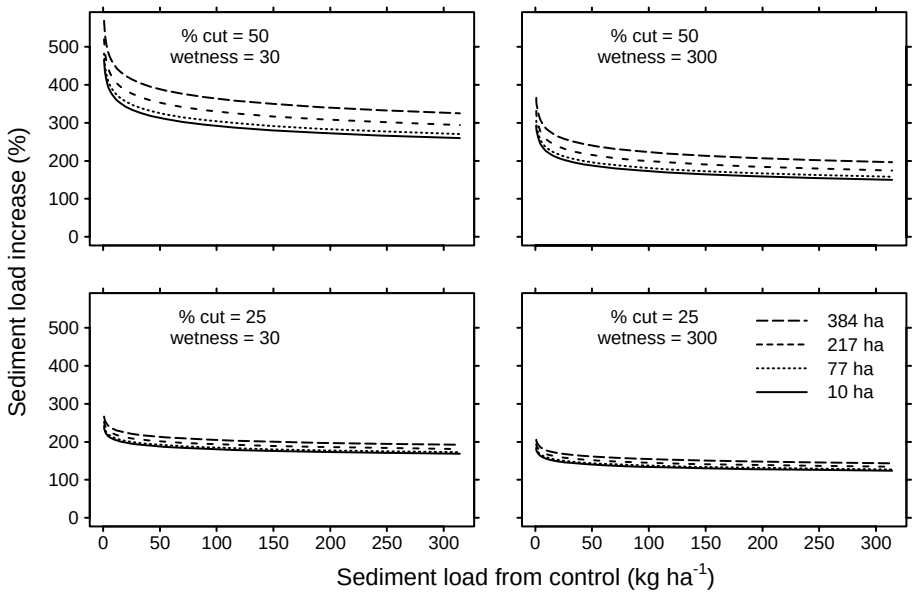

Figure 20. Effect of watershed area on predictions from sediment model {(24),(17),(18)} for two levels of cutting and two levels of antecedent wetness. Watershed areas are those of ARF, FLY, DOL, and BAN (Table 1). Predictions are for first year after cutting with $x_{ij}^{(1)} = x_{ij}^{(2)} = 12$ m ha^{-1}.

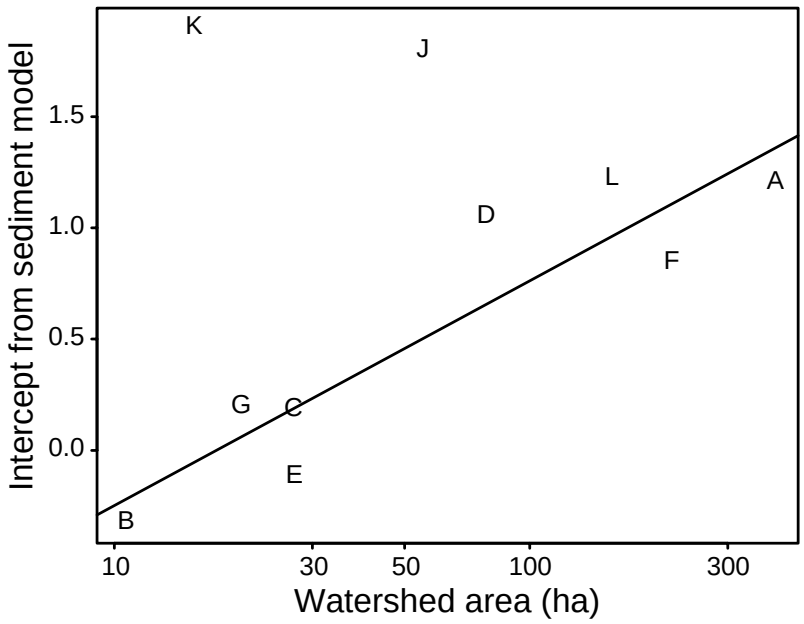

Figure 21. Relation between watershed area and fitted intercepts b_{0i} from model {(24),(17),(18)}, with β_6 fixed at zero. Watersheds JOH (J) and KJE (K) are omitted from regression. Letters designate watersheds (e.g. G is watershed GIB).

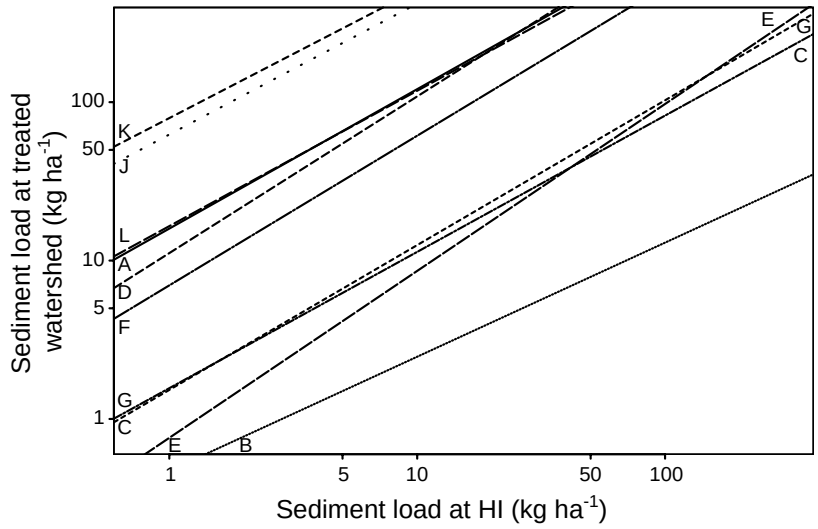

Figure 22. Regression lines for each watershed based on intercepts b_{0i} and slopes b_{1i} of sediment model {(24),(17),(18)}, with β_6 fixed at zero. Letters designate watersheds (e.g. G is watershed GIB).

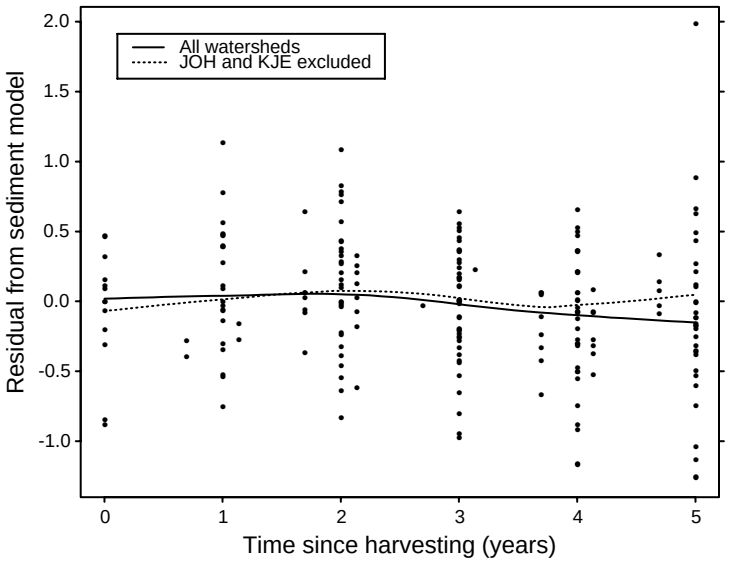

Figure 23. Relation between residuals from sediment model {(24),(17),(18)} and time after logging. Curves are fit by loess method [Cleveland, 1979], with and without the anomalous watersheds JOH and KJE.

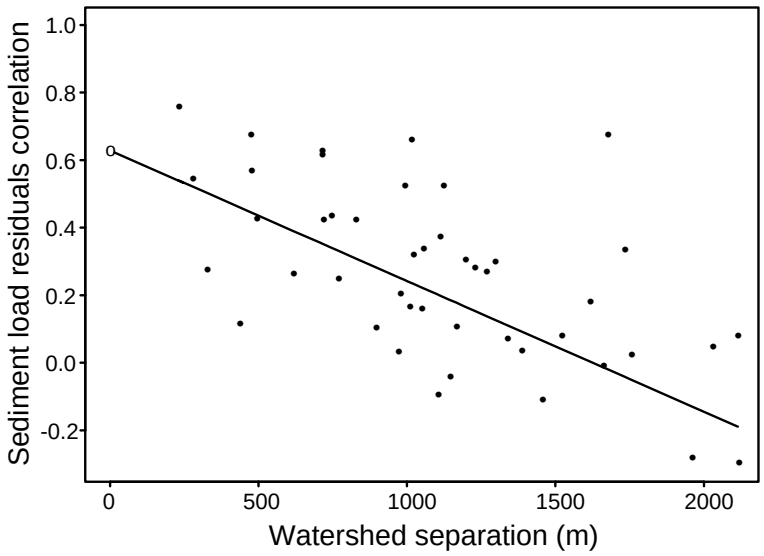

Figure 24. Relation between sediment residuals correlation and distance between watershed centroids. Residuals are from maximum likelihood fit to sediment model {(24),(17),(18)}. Curve depicts equation (17), with estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$.

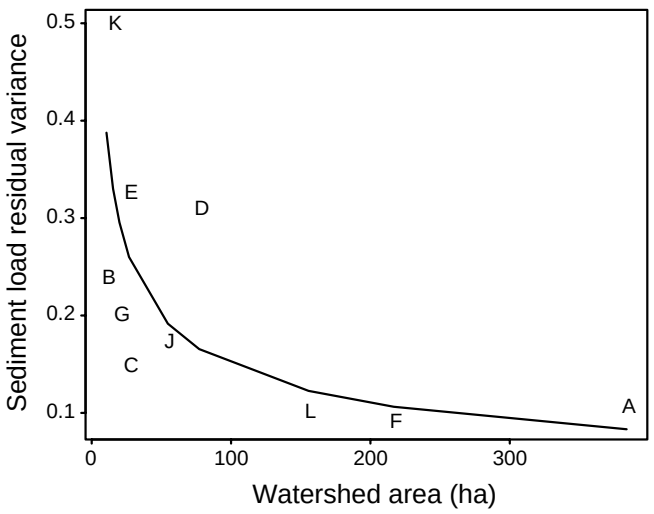

Figure 25. Relation between variance of sediment residuals and watershed area. Residuals are from maximum likelihood fit to model {(24),(17),(18)}. Curve depicts equation (18) with estimated parameters $\hat{\theta}_3$ and $\hat{\theta}_4$. Letters designate watersheds (e.g. G is watershed GIB).

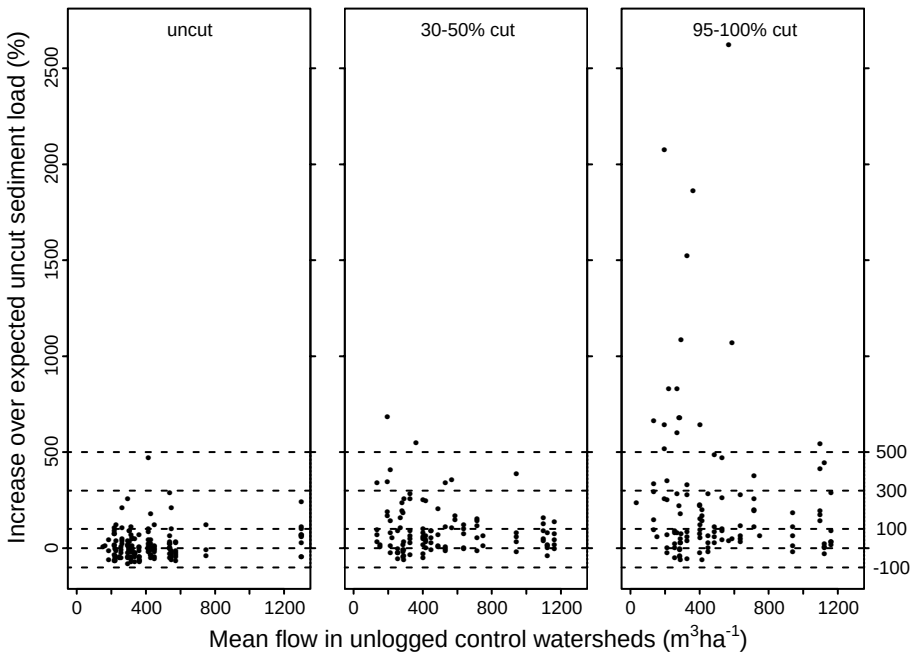

Figure 26. Percentage increase over expected uncut storm sediment load as related to mean of storm runoff volume in HIM control watersheds for uncut (before treatment), partly (30-50%) clear-cut, and (95-100%) clear-cut watersheds. Bias-corrected predictions are from model {(25),(17),(18)} with disturbance set to zero.

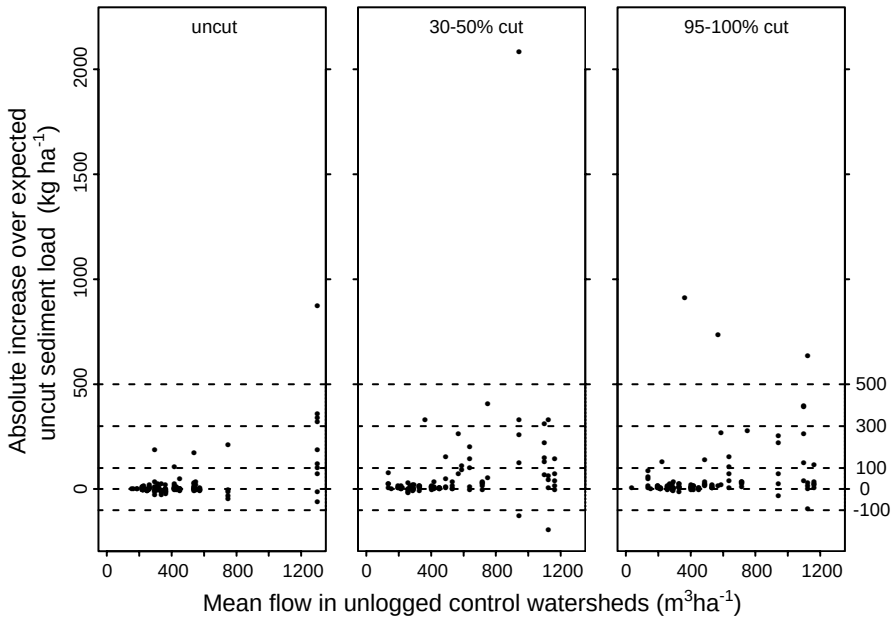

Figure 27. Absolute increase over expected uncut storm sediment load as related to mean of storm runoff volume in HIM control watersheds for uncut (before treatment), partly (30-50%) clear-cut, and (95-100%) clear-cut watersheds. Bias-corrected predictions are from model {(25),(17),(18)} with disturbance set to zero.

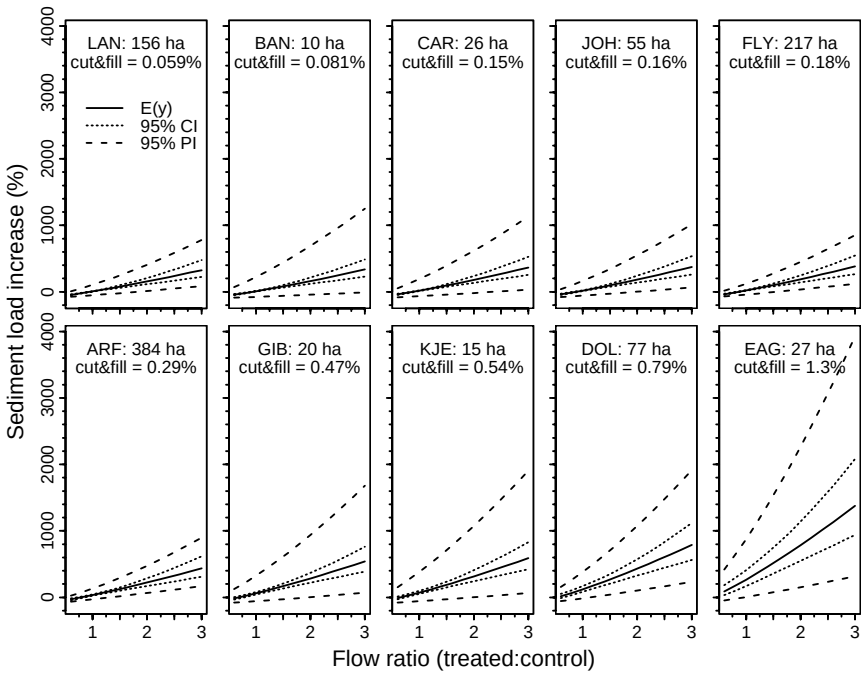

Figure 28. Predictions of sediment load as a function of flow ratio ($\Delta q_{ij}^{(2)}$) based on sediment load model {(25),(17),(18)}, with area interaction term for cumulative impacts (β_4) fixed at zero. Expected increases and 95% confidence (CI) and prediction (PI) intervals are shown for each treated watershed, ordered by proportion of the watershed occupied by road cuts and fills.

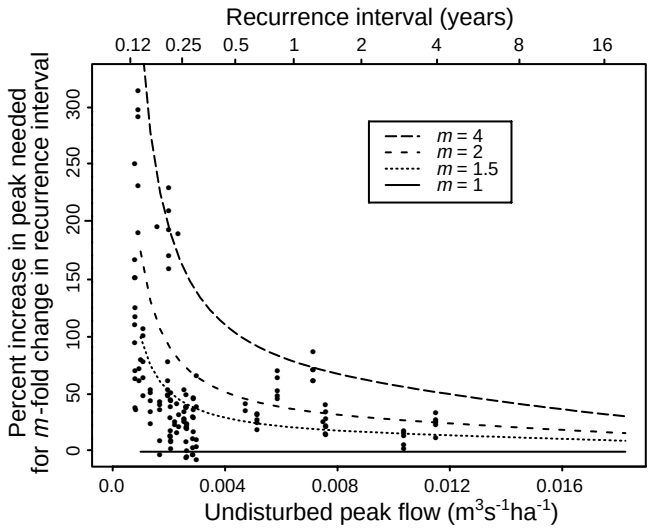

Figure 29. The curve shows the percentage increase in peak flow necessary to reach a size that formerly had 1 to 4 times the recurrence interval. The data points are from Figure 13 (third frame), representing the observed percentage increases in storm peak flow (based on the HI control, plotted on the abscissa) in 95-100% clear-cut watersheds.

Table 1. Basic watershed data and percentage in various conditions. Cut area includes portions of stream buffer zones corresponding to the proportion of timber volume removed.

Water-shed	Area (ha)	Cut area	Cable	Trac-tor	Road+ Lndg	Total Bare	Total Burnt	Dates logged
ARF	384	45.5	35.1	7.1	1.8	2.9	24.0	Spr89-Win92
BAN	10	95.0	77.3	13.4	2.6	3.2	0.0	Fall91
CAR	26	95.7	2.1	9.2	2.8	4.4	0.0	Fall91-Win91
DOL	77	36.4	27.4	5.9	2.5	3.7	33.9	Fall90-Fall91
EAG	27	99.9	79.0	15.4	4.9	8.5	97.8	Fall90-Fall91
FLY	217	45.4	34.6	7.6	1.6	3.0	30.4	Spr89-Sum91
GIB	20	99.6	54.9	39.4	4.2	7.9	98.2	Spr91-Sum91
HEN	39	0.0	0.0	0.0	0.0	0.0	0.0	
IVE	21	0.0	0.0	0.0	0.0	0.0	0.0	
JOH	55	30.2	26.4	1.3	2.0	2.1	0.1	Spr89-Fall89
KJE	15	97.1	85.2	3.9	6.5	6.9	0.0	Spr89-Fall89
LAN	156	32.2	27.8	1.9	1.0	1.9	20.3	Spr89-Spr90
MUN	16	0.0	0.0	0.0	0.0	0.0	0.0	
NFC	473	12.7	38.6[a]	7.6[a]	2.0[a]	3.2[a]	19.5[a]	Spr85-Spr86
		+36.9	38.6	7.6	2.0	3.2	19.5	Spr89-Win92

[a] not measured; assumed equal to Spr89-Win92 disturbance proportions

Table 2. Comparison of suspended sediment load estimation by time interpolation, stage interpolation, and SALT algorithms. The load was estimated for 5000 simulated SALT samples from each storm event.

				Percent RMSE			Percent Bias		
Station	Start of Storm	Load (kg/ha)	$\bar{n}$	Time Interp	Stage Interp	SALT	Time Interp	Stage Interp	SALT
ARF	950109	178.6	21.2	6.0	6.7	12.2	-2.3	-2.8	0.1
ARF	950113	123.6	22.9	2.8	3.4	8.2	-1.6	-2.0	0.1
ARF	950308	122.4	32.6	4.1	4.1	7.6	-0.3	-0.5	0.0
ARF	950108	99.2	8.6	14.2	14.6	19.8	-6.0	-7.2	-0.0
ARF	940216	33.6	16.5	7.0	6.7	10.0	-3.7	-3.5	-0.2
Mad	821214	846.3	41.8	2.1	1.8	10.0	0.0	-0.3	-0.1
Mad	830209	527.2	36.0	4.2	4.1	13.8	0.4	-1.3	0.1
Mad	830117	198.0	40.8	2.2	2.6	7.2	-0.4	-0.9	0.1
Mad	830225	134.4	22.9	7.8	7.6	19.3	-1.6	-2.6	0.3
Mad	831223	42.8	18.1	5.8	5.4	13.6	-2.7	-2.7	0.0
Mad	830221	33.2	15.7	7.5	8.1	16.1	-4.0	-4.9	-0.3
Mad	830212	27.2	14.0	8.1	7.4	16.2	-3.2	-3.9	0.0
Mad	830218	25.4	14.1	14.7	15.1	22.3	-3.4	-4.2	0.0

Table 3. Maximum likelihood parameter estimates for storm peaks model {(10),(16),(18)}, excluding β_{0i} and β_{1i}. p_N is normal probability value for H_0: $\beta = 0$.

Parameter	Effect	Estimate	Standard Error	p_N
β_2	Recovery	0.0771	0.0183	<0.0001
$\beta_3^{(1)}$	Fall logging	0.5939	0.0996	<0.0001
$\beta_3^{(2)}$	Winter logging	0.0000	0.2843	1.0000
β_4	Amount logged	1.1030	0.3409	0.0012
β_5	Storm size interaction	-0.0963	0.0484	0.0468
β_6	Wetness interaction	-0.2343	0.0251	<0.0001
β_7	Watershed area interaction	3.553E-4	2.861E-4	0.2142
θ_1	Correlation shape parameter	2.809E-3	6.188E-4	<0.0001
θ_2	Correlation limit parameter	0.4698	0.1564	0.0027
θ_3	Variance magnitude	0.2285	0.0242	<0.0001
θ_4	Variance shape	-0.0937	0.0238	0.0001

Table 4. Maximum likelihood parameter estimates for storm runoff model {(10),(17),(18)}, excluding β_{0i} and β_{1i}. p_N is normal probability value for H_0: $\beta = 0$.

Parameter	Effect	Estimate	Standard Error	p_N
β_2	Recovery	0.0912	0.0143	<0.0001
$\beta_2^{(1)}$	Fall logging	0.8117	0.0910	<0.0001
$\beta_2^{(2)}$	Winter logging	-0.196	0.225	0.3843
β_4	Amount logged	2.3054	0.2646	<0.0001
β_5	Storm size interaction	-0.1103	0.0467	0.0181
β_6	Wetness interaction	-0.2362	0.0236	<0.0001
β_7	Watershed area interaction	6.481E-4	2.578E-4	0.0119
θ_1	Correlation intercept	0.6697	0.0587	<0.0001
θ_2	Correlation slope	-1.898E-4	4.962E-5	0.0001
θ_3	Variance magnitude	0.1987	0.0190	<0.0001
θ_4	Variance shape	-0.0873	0.0209	<0.0001

Table 5. Percentage and absolute departures from predicted annual storm runoff volume (sum of storms).

	Uncut	30-50% Clearcut	95-100% Clearcut
Mean (%)	2	23	58
Median (%)	2	19	51
Mean (m^3 ha^{-1} yr^{-1})	54	415	1119
Median (m^3 ha^{-1} yr^{-1})	29	387	1050

Table 6. Chow test [Chow, 1960; Wilson, 1978] significance levels for hypothesis of no change in suspended sediment load after logging.

Watershed	HI control	HIM control
ARF	0.1649	0.0215
BAN	0.0128	0.0292
CAR	0.0000*	0.0001*
DOL	0.0198	0.0093
EAG	0.0056	0.0013*
FLY	0.3528	0.0955
GIB	0.0002*	0.0096
JOH	0.0983	0.0476
KJE	0.0026*	0.0384
LAN	0.8018	0.2453

* significant at nominal $\alpha = 0.005$ (experimentwise error rate = 0.05)

Table 7. Summary of changes in suspended sediment load (summed over storms) after logging in North Fork subwatersheds. Predicted loads are computed from pre-treatment linear regressions between the logarithms of the storm sediment load in the treated watershed and control HIM, the mean of the storm sediment loads at watersheds HEN, IVE, and MUN. Predictions were corrected for bias when back-transforming from logarithmic units. The number of years in the post-logging period varies from 4 to 6, depending upon when the watershed was logged and whether or not monitoring was discontinued in water year 1996.

Treated watershed	Number of years	Observed (kg $ha^{-1}yr^{-1}$)	Predicted (kg $ha^{-1}yr^{-1}$)	Change (kg $ha^{-1}yr^{-1}$)	Change (%)
ARF	4	505	591	-86	-15
BAN	4	85	28	57	203
CAR	5	240	108	132	123
DOL	5	1130	306	824	269
EAG	5	710	210	500	238
FLY	5	536	555	-19	-3
GIB	4	358	119	239	200
JOH	5	667	865	-198	-23
KJE	5	821	1371	-551	-40
LAN	5	420	400	20	5
NFC	6	465	246	219	89

Table 8. Increase in residual sum of squares after dropping variables from least squares fit to model (24).

Coefficient	Variable	SS Reduction
β_2	Change in flow	25.33
β_3	Burned stream channel	10.21
β_4	Unburned stream channel	3.51
β_5	Storm size interaction	1.62
β_6	Watershed area interaction	0.62

Table 9. Maximum likelihood parameter estimates for suspended sediment load model {(24),(17), (18)}, excluding β_{0i} and β_{1i}. p_N is normal probability value for H_0: $\beta = 0$. Control is HI, the mean sediment load from watersheds HEN and IVE.

Parameter	Effect	Estimate	Standard Error	p_N
β_2	Change in flow	1.3276	0.1609	<0.0001
β_3	Stream length, burned	0.0376	0.0057	<0.0001
β_4	Stream length, unburned	0.0204	0.0053	0.0001
β_5	Storm size interaction	-0.0051	0.0017	0.0031
β_6	Watershed area interaction	-3.316E-5	1.649E-5	0.0443
θ_1	Correlation intercept	0.6222	0.0846	<0.0001
θ_2	Correlation slope	-3.802E-4	9.218E-5	<0.0001
θ_3	Variance magnitude	1.0841	0.1565	<0.0001
θ_4	Variance shape	-0.2286	0.0338	<0.0001

Table 10. Maximum likelihood parameter for suspended sediment load model {(25),(17),(18)}, excluding β_{0i} and β_{1i}. p_N is normal probability value for H_0: $\beta = 0$. Control is HIM, the mean sediment load from watersheds HEN, IVE, and MUN.

Parameter	Effect	Estimate	Standard Error	p_N
β_2	Flow increase (log ratio)	1.3564	0.1414	0.0000
β_3	Road cut and fill area	107.11	13.071	0.0000
β_4	Watershed area interaction	-0.1822	0.0872	0.0367
θ_1	Correlation intercept	0.6848	0.0643	0.0000
θ_2	Correlation slope	-3.949E-4	7.618E-5	0.0000
θ_3	Variance magnitude	1.1839	0.1473	0.0000
θ_4	Variance shape	-0.2330	0.0290	0.0000

Table 11. Percentage and absolute departures from annual (sum of storms) sediment load predicted from HIM control. Parenthesized values omit outlier in middle frame of Figure 27.

	Uncut	30-50% Clearcut	95-100% Clearcut
Mean (%)	35	73 (67)	212
Median (%)	15	52	109
Mean (kg ha^{-1} yr^{-1})	65	263 (180)	262
Median (kg ha^{-1} yr^{-1})	1	46	59

Table 12. Apparent and cross-validated RMSE for model predictions.

Data Omitted	Data Predicted	Model			
		Peaks[a]	Volume[b]	Sed (HI)[c]	Sed (HIM)[d]
None	All	0.1589	0.1426	0.4584	0.5046
10% at random	All	0.1633	0.1483	0.4900	0.5238
None	Post-treatment	0.1654	0.1560	0.4644	0.5094
10% at random	Post-treatment	0.1692	0.1623	0.4948	0.5291
Systematic by station	Post-treatment	0.1739	0.1676	0.6966	0.6724

[a] model {(10),(16),(18)}, HI control, $\beta_7 = 0$

[b] model {(10),(17),(18)}, HI control, $\beta_3^{(2)} = 0$

[c] model {(24),(17),(18)}, HI control

[d] model {(25),(17),(18)}, HIM control

Table 13. Regression slope of observed versus predicted response.

Data Omitted	Data Predicted	Model			
		Peaks[a]	Volume[b]	Sed (HI)[c]	Sed (HIM)[d]
None	All	1.0039	1.0103	1.0012	0.9986
10% at random	All	1.0028	1.0047	0.9920	0.9947
None	Post-treatment	1.0077	1.0103	0.9921	0.9651*
10% at random	Post-treatment	1.0085	1.0020	0.9825	0.9611*
Systematic by station	Post-treatment	1.0014	0.9998	0.8601**	0.8775**

[a] model {(10),(16),(18)}, HI control, $\beta_7 = 0$

[b] model {(10),(17),(18)}, HI control, $\beta_3^{(2)} = 0$

[c] model {(24),(17),(18)}, HI control

[d] model {(25),(17),(18)}, HIM control

* $0.01 < p < 0.05$ for one-sided test of H_0: slope=1 (with H_A: slope<1)

** $p < 10^{-6}$ for one-sided test of H_0: slope=1 (with H_A: slope<1)