Comments on the Draft Elk River TMDL
March 5, 2014

Prepared by:

Jack Lewis
(USFS Redwood Sciences Laboratory – retired)

and

Randy Klein
(Redwood National and State Parks – retired)

Introduction

We offer these comments on the North Coast Regional Water Quality Control Board's "*Peer Review Draft: Staff Report to Support the Technical Sediment Total Maximum Daily Load (TMDL) for the Upper Elk River*" (NCRWQCB, 2013) for your consideration. A journal article we recently published (Klein et al., 2012) played a significant role in constructing some of the TMDL's approaches to reducing turbidity and suspended sediment loads and is a central focus of some of the comments already provided.

Our comments are focused on offering some suggestions for the TMDL and responding to several documents: 1) a report produced by Humboldt Redwoods Company titled "Trends in Sediment-related Water Quality After a Decade of Forest Management Implementing an Aquatic Habitat Management Plan" by Sullivan et al. (2012), 2) a comment letter on the draft TMDL submitted to the Regional Board by Dr. Lee MacDonald dated Jan. 17, 2014, and 3) two comment letters submitted to the Regional Board from the two industrial timber companies subject to the TMDL: Green Diamond Resources Company (GDRC), letter dated Nov. 21, 2013, and Humboldt Redwoods Company (HRC), letter dated Dec. 2, 2013.

The Draft TMDL relies, in part, on a journal article we, and Dr. Matt Buffleben (currently with the State Water Quality Control Board), published in 2012 (Klein et al., 2012). Our publication was also a central focus of a portion of the HRC report, Dr. MacDonald's comments and the two industry comment letters, their main objections centering on the use of our findings in the journal article for setting limits on the rate of timber harvest. If there is anywhere that the lack of harvest rate limits has resulted in severely degraded water quality and harm to beneficial uses, it is the Elk River. Klein and others (2012) documented the poor water quality at several turbidity and sediment monitoring sites in the Elk River in water years (WY) 2004 and 2005. More recent work has shown that water quality remains poor in the Elk River and that downstream impacts remain severe. This despite road upgrading and decommissioning, application of modern BMPs, and the much-touted Aquatic Habitat Conservation Plan (AHCP) approved for Maxxam and later transferred to HRC. As is widely acknowledged, background erosion and sediment delivery rates are relatively high in the Elk River. This characteristic causes the Elk River watershed to be hyper-sensitive to land use disturbance, thus to avoid worsening CWEs, harvest rates must be lower there than harvest rates that might be applied in more inherently stable watersheds.

We acknowledge that very recent harvesting in Elk River by HRC has a much lighter touch on the landscape than harvesting methods (and rates) applied even a few years prior. We were kindly hosted by HRC on a brief field review in 2013. We observed recent harvest areas that appeared more benign than earlier harvests. Group selection is now widely applied rather than even aged management (clear cutting), and efforts were made to ensure no bare ground remained after operations were completed. But the Elk River remains heavily damaged from the earlier over-harvesting. When bought by HRC, those lands had (and still have) massive damage from the methods and rates of timber harvest imposed by the previous owner. With a goal of protecting or restoring water quality, past practices cannot be ignored in constructing an effective TMDL. Beneficial uses are blind to the sources of degraded water quality, whether they result from natural, legacy, or contemporary sources. A watershed as degraded as Elk River must be allowed a chance to recover, and the only way this will happen while continuing to allow extraction of timber resources is to limit the pace of that extraction through limits on harvest rates.

Clearly, imposing limits on the rate of timber harvest is controversial, and we are certainly not the first to broach this sensitive, politically and economically charged issue. The State's Forest Practice Rules could provide a safety net for watersheds that would otherwise be scheduled for overharvest by adopting harvest rate restrictions as BMPs. Had this been a best management practice when Pacific Lumber Company was taken by Maxxam in the 1980s, the devastation that took place in the Elk River and other areas could have prevented. Ongoing harvest in some areas is still occurring at too rapid a rate (e.g., GDRC's Maple Creek holdings), and will surely occur elsewhere.

A statewide harvest rate BMP, customizable for individual watersheds, presents a relatively simple means to avoid repeating past mistakes like those which occurred in the Elk River watershed. Silvicultural systems that retain more canopy can be accommodated by adjusting the silvicultural weighting factors used in computing clearcut equivalent area harvest rate, or by using newly available satellite-based datasets that quantify annual changes in forest canopy. Diversity among watersheds in terms of inherent erosional stability can be accommodated by setting numerical limits on the rate of timber harvest, as we suggest in Klein et al., (2012), and adjusting them through time as needed based on watershed performance (e.g., lowering of 10% turbidity to levels consistent with the North Coast Basin Plan).

Technical Comments

The comments that follow are arranged by subject matter and address criticisms from the documents mentioned above. For convenience, we refer to the comment letters with the following abbreviations: Green Diamond Resources Co. (GDRC), Humboldt Redwoods Co. (HRC), and Dr. Lee MacDonald (LM). Dr. MacDonald performed his work under contract by HRC and GRDC in response to the Peer Review Draft: Staff Report to Support the Technical Sediment Total Maximum Daily Load (TMDL) for the Upper Elk River (NCRWQCB, 2013). The letter from LM included a 12-page Appendix objecting to the Peer Review Draft's use of a paper we recently published in the journal Geomorphology (Klein et al., 2012: KLB) as justification for restricting timber harvest rates in the Elk River. The letter questions the validity of KLB for understanding the effects of timber harvest and guiding future management. Several of the points raised in the LM letter were mentioned in the HRC and GDRC letters. All three of the letters refer to a contradictory HRC Report (Sullivan et al., 2012) that analyzes a decade of HRC's monitoring data, concluding that our results were atypical of the 2003-2011 HRC monitoring period, and that sediment loads and turbidity are generally declining in Elk River under recent management. Having read the HRC Report in detail, and having already found one major error in the analysis, one of us (Lewis) recently undertook a re-analysis of the data from that Report. It is clear now that the major findings of the Report are based on incorrect analysis and interpretation of the data. This letter is primarily a response to the criticisms aimed at KLB, a discussion of its applicability and limitations, and an explanation of the problems in the HRC Report, which is presented in detail as Appendixes B, C, and D of this letter. In addition, because of its relevance to the linkage between forest harvest and delayed landslides, a previously unpublished report on redwood root biomass by Ziemer and Lewis is included as Appendix A.

The industry criticisms of KLB have focused on 4 issues:

1. The application of weighting factors fo r different silvicultural methods in characterizing harvest rates.
2. The uncertainty regarding causal linkages between harvesting and chronic turbidity.
3. Robustness of the results, including contradictions between KLB and the subsequent HRC Report.
4. The effectiveness of BMPs.

As explained in Appendix B of this letter, the relationship between harvesting and chronic turbidity in their data was misunderstood by HRC because of confounding effects in their model. In reviewing the HRC Report, fundamental errors in the trend analysis have also come to light (Appendix C). The report is cited by HRC and LM as showing significant declines in sediment yield but the declines seem to be limited to Freshwater Creek, as will be discussed below, with major details in Appendix D.

WEIGHTING FACTORS

LM: "*The weighting factors in Table 3 are not valid for all the key processes that cause forest management activities to increase in erosion and turbidity*". The silvicultural weighing factors were not intended to represent all the key processes that cause forest management activities to increase erosion and turbidity. They are intended to only represent changes in canopy cover, not myriad other factors that would only confuse the interpretation of the model. The harvest rate variable is not an attempt to account for all the unique influences of different harvest operations on erosion and sediment delivery. Harvest practices vary but the one thing they always have in common is canopy removal. While far from perfect, these weighting factors certainly improved the measure of canopy removal relative to simple acres harvested. If the model is to have explanatory value, different influences should be represented by different predictors that can be individually tested, not combined into an index that incorporates a complex model full of assumptions. Such a model would be useless for explanatory or regulatory purposes.

LM: "*As acknowledged by the authors (p. 143 in Klein et al., 2011), the weighting factors also implicitly combine both the relative effect on erosion and the relative likelihood that this sediment will reach the stream*". We made no such acknowledgment. We explained that the factors "may not properly reflect relative changes in hillslope hydrology and root strength" due to varying water utilization and root dieback. These are changes that are directly related to canopy removal, but which should vary depending on its spatial pattern. If better information were available relating the spatial pattern of canopy removal to transpiration, interception and root dieback, it might be reasonable to refine the weighting factors.

LM: "*there does not appear to be any consideration of how these weighting factors should change over time*". If time were to be represented in the model it would best be represented explicitly as a variable or, as we did, by aggregating different periods and letting the algorithm identify the most important harvest window. With only 28 data points, the potential for multivariate modeling is extremely limited, so the model had to be kept simple. Coding assumptions about recovery into equivalent clearcut area ('ECA', Sullivan et al., 2012) factors would have made it impossible to compare the predictive value of different 5-year harvest periods.

LM: "*the conclusion that 10-15 year ECA is the key control on 10% exceedence turbidities*". This was not a conclusion; in fact the data are inconclusive in that regard. The most predictive model incorporated the 10-15 year ECA and basin area, but the 0-15 year harvest rate had a higher simple correlation with turbidity. We only know that several of the harvest rate variables outperformed all other predictors. The importance of the harvest rate variables may be influenced by other disturbances associated with harvest rates, but none of those disturbances were statistically separable.

GDRC: KLB "*assigned the same weighting coefficients to clearcut areas and roads...This is a significant flaw that grossly overstates the impact of clearcut harvesting.*" As we have stated, the coefficients were intended to index canopy removal only. In characterizing harvest rate, it is appropriate that newly constructed roads and clearcuts should carry the same weight. Nothing is overstated by doing so. We also tested road variables in the model and they were correlated with harvest rates but not as well-related to turbidity. Assuming we had accurate road information, roads would have been included in the models if they were better predictors of turbidity, or if they had explained a significant amount of variability after accounting for canopy removal (Figure 1). If we had a larger sample size (i.e. more watersheds) they might have been significant. It is not our intent to deny their importance; nevertheless, in this data set, even basin area was a better predictor than any of the road variables.

CAUSAL PROCESSES

LM: "*The presumed dominant role of harvest-related shallow landslides due to canopy removal can only be inferred. The proposed imposition of a limit on harvest rate necessarily converts this statistical inference to a known and primary cause*." In fact, a limit on harvest rates does not assume a specific mechanism; it only assumes that a mechanism exists. Although we discussed landslides in relation to a delayed effect of harvesting, we do not presume that shallow landslides are necessarily the dominant process triggered by harvesting. Gullying, headcut formation and in-channel erosion likely play important roles as well. The Peer Review Draft emphasizes the importance of peak flows on in-channel erosion and near-stream landslides.

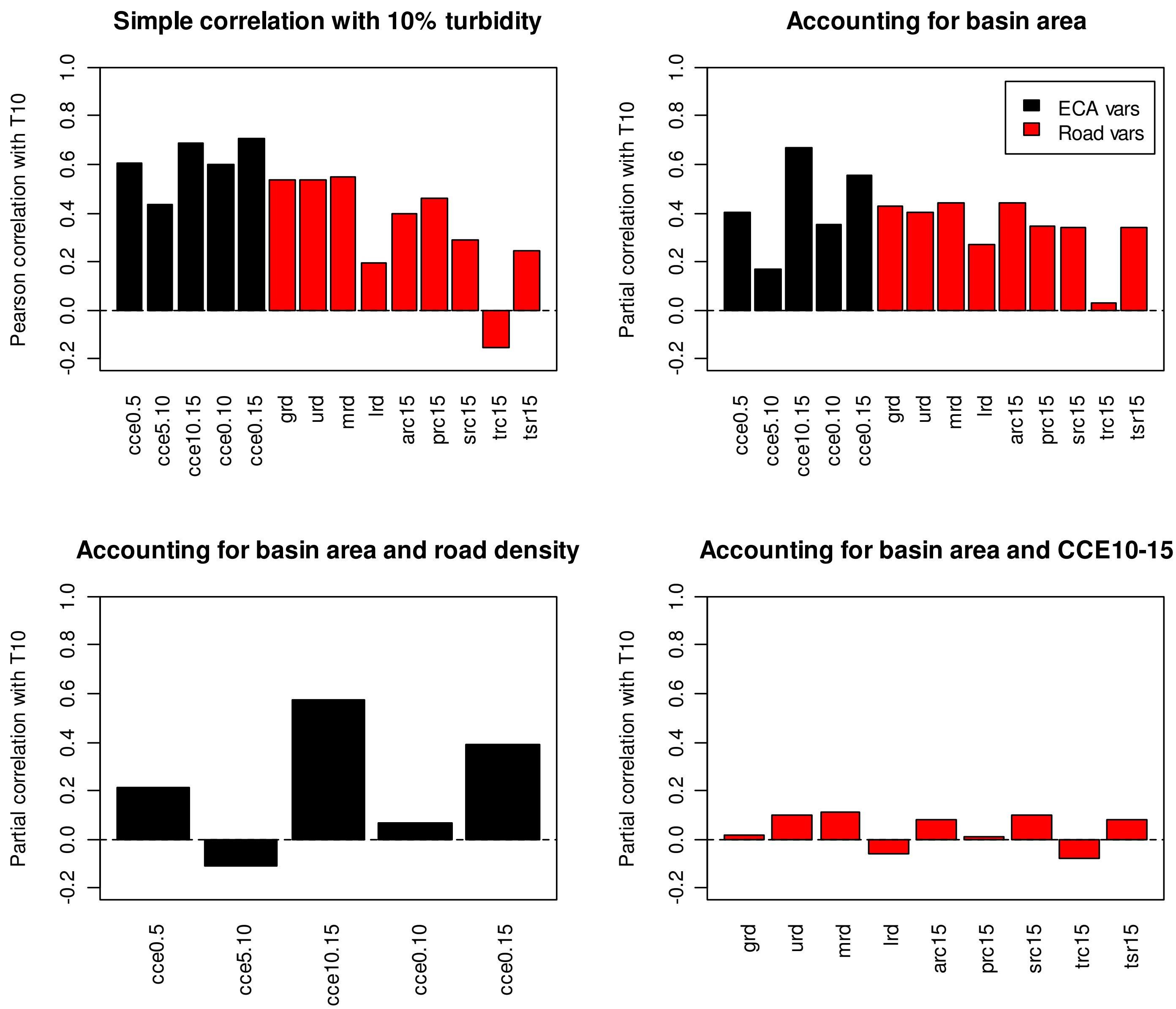


Figure 1. In the KLB data set, road variables generally had lower correlations than harvest rate (ECA) variables with 10% exceedence turbidity (T10). The upper panels show that the highest ECA correlates were the 0-15 and 10-15 year harvest windows (cce0.15 and cce10.15), while the highest road correlates were basinwide road density (grd), upper slope road density (urd), and middle slope road density (mrd). After accounting for basin area and road density, the harvest variables cce0.15 and cce10.15 still explained a great deal of variability in turbidity. On the other hand, none of the road variables explained much variability after accounting for basin area and cce10.15.

LM: "*The watershed-scale data and statistical analyses in Klein et al. (2011) do not provide any definitive links to the underlying causal processes*". The paper is an empirical statistical study, not an investigation of processes. We were fortunate to have assembled such an extensive turbidity data set from 28 watersheds under different ownerships, allowing us to show statistically for the first time what casual observers have known for decades: that the most heavily logged watersheds are generally the most turbid. Detailed process investigations have not been done in all these watersheds and were far beyond the scope of this paper. We agree with LM that, by itself, a statistical correlation is not an adequate basis for regulation. However, the processes that link forest harvesting to increased erosion and sediment transport are for the most part well understood. There is a wealth of literature establishing the linkages, and the TMDL refers to some of the more relevant papers, many of which originated from research at Caspar Creek. It is appropriate that those papers should be heavily relied upon because Caspar Creek is the only long-term experimental watershed in the redwood region. The recent summary of Caspar Creek findings by Cafferata and Reid (2013) states many of the relevant results, a sampling of which are quoted here with our occasional comments bracketed.

- In winter, when differences in soil moisture levels between logged and unlogged areas are minimal, peak flows increase after clearcutting due primarily to reduced interception loss after logging, and secondarily to reduced winter transpiration (Reid and Lewis 2007, Reid 2012).
- "After logging, the main sediment sources are large landslides and increased in channel erosion caused by higher flows." [refers to clearcutting]
- "Timber harvesting can modify transpiration and rainfall interception, increasing the amount of subsurface flow generated during storms; and road construction and heavy equipment use can compact soils and disrupt soil pipes. These kinds of changes can alter subsurface flow patterns and elevate pore water pressures during large storms, increasing landslide risk at some sites (LaHusen 1984, Montgomery et al. 2000)."
- "After clearcut logging in the North Fork, in-channel erosion (i.e., gullying, channel incision, and bank erosion) appears to be the major sediment source during periods without large landslides, and sediment inputs increased significantly along an undisturbed channel reach downstream of a logged subwatershed (Figure 23B, Reid et al. 2010)."
- "The largest landslides (>200 m3 or 260 yd3) did not occur until 9 to 14 years after North Fork logging and shortly after pre-commercial thinning, at a time when root strength is expected to be near its minimum value and hydrologic changes are once again important (Reid and Keppeler 2012)."
- "Caspar Creek studies showed that WLPZs and road repair work alone cannot prevent in-channel sediment increases because significant sediment inputs from in-channel sources can be generated by logging-related flow increases (Lisle et al. 2008, Reid et al. 2010)."
- "Data from Caspar Creek (Reid 2012) suggest that the [peak-flow] response for single-tree selection logging may be about 60% of that for the equivalent canopy removal by clearcutting". [More work needs to be done to quantify the relative impacts of different spatial patterns of harvesting]

The root-dieback linkage in redwoods is understood qualitatively but very difficult to study or quantify. However a major effort was undertaken by the Redwood Sciences Lab in the 1970s. Roots were extracted, sieved, and weighed from clusters of 1.33 m^3 excavations in 6 redwood stands and 7 mixed conifer stands of varying ages. Over 300 m^3 of soil was processed. The data from the redwood stands have not been published, but Bob Ziemer and Jack Lewis analyzed them, and the methods and results are included as Appendix A to this letter. The data indicate that, in the redwood stands, combined live and dead root biomass < 25mm reaches a minimum between 11 and 25 years after logging (Figure A5). Larger roots are even slower in returning to prelogging levels, requiring well over 60 years (Figure A4) after cutting an old growth forest.

Elevated pore water pressures are known to be a primary cause of landslides. In another recent HRC-funded study, Dhakal and Sullivan (2014) investigated throughfall and pore water pressures under a 60-year old forest canopy. Annual interception was 23%, nearly identical to that measured at Caspar Creek under a 120-year old canopy. They dismissed any influence of rainfall interception on pore water pressures because measured interception rates were small during the most intense rainbursts. Rainfall intensity is known to be important in triggering landslides, but, as they acknowledged, antecedent moisture conditions too are relevant in the rapid rise of pore water pressures at the onset of a rainfall event. The data in their Figure 4a show that, before the largest event of the year, 36% more water (20cm) had reached the ground surface in an opening than under the forest canopy. Excess water in canopy openings varied from 27 to 40% during the wet season (Figure 2). Such differences are magnified further by transpiration that occurs under the canopy. At Caspar Creek, transpiration accounts for approximately one third of water consumption under the canopy between October and April. Therefore excess water in canopy openings could exceed 50% at times during the wet season. Dhakal and Sullivan (2014) completely neglected the potential for all this excess water to affect pore water pressures. If they had placed some of their 83 piezometers in a recently harvested area they might have been able to detect elevated piezometric pressures as has been done at Caspar Creek. By analysis of pore water pressures in a clearcut and under canopy before and after harvesting Keppeler and Brown (1998) demonstrated clearly (their Fig .6) that harvesting elevates pore water pressures.

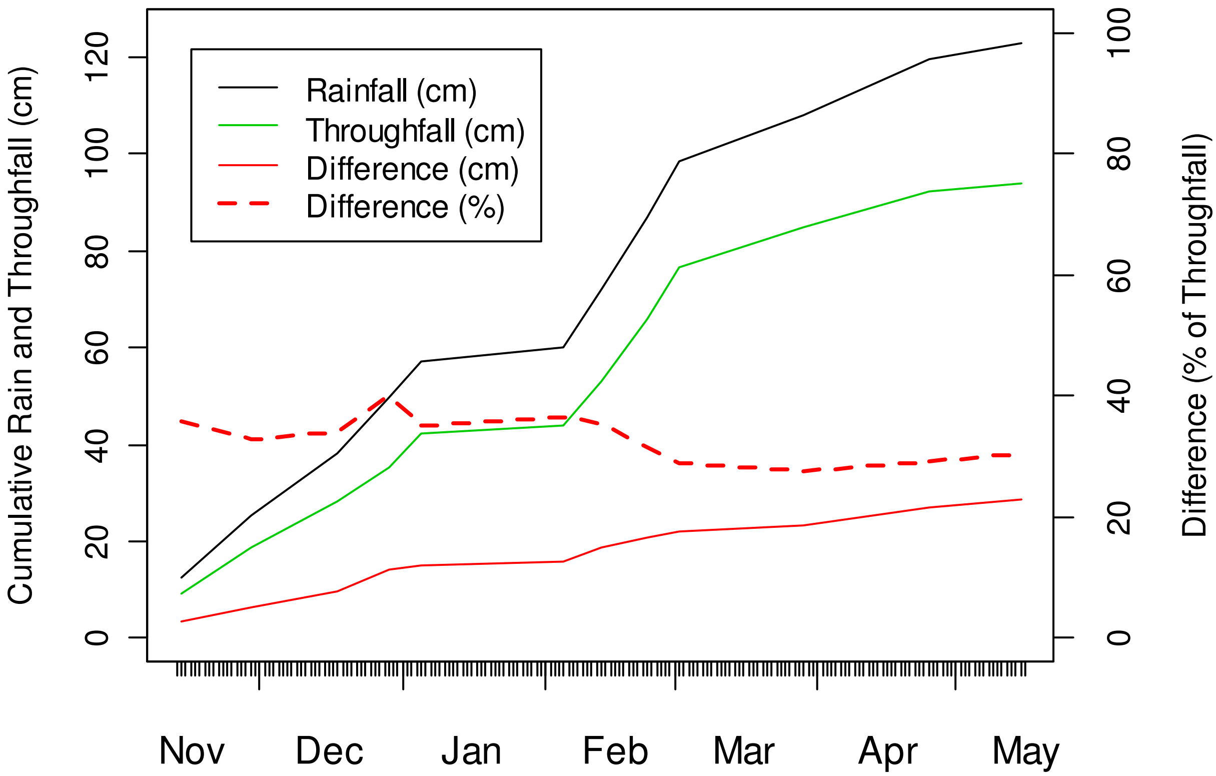


Figure 2. Throughfall measurements taken under a 60-year old canopy in Doe Creek, tributary of Elk River (data from Dhakal and Sullivan, 2014). The difference in effective rainfall, i.e. that which reaches the ground surface, is shown both in centimeters and as a percentage of throughfall. Expressing the difference as percentage of throughfall is appropriate because throughfall is the baseline amount that reaches the ground in an undisturbed forest.

ROBUSTNESS OF THE ANALYSIS

LM: "*Only the 2005 data were used for the multiple regression analysis*". Yes, because multiple regression generally requires large sample sizes to be informative, and that was the year for which we had the most complete data set (n=27). However, both 2004 and 2005 were used for the legacy analysis. In any case, the HRC Report, with even smaller sample sizes, showed that the 10-15 year harvest rate (CCE10-15) was significant in *both* years. Yet, the HRC Report claims that the 2005 results were anomalous and that the 10-15 year harvest rate (CCE10-15) was negatively associated with turbidity when all years were considered. The problem with that interpretation is that their mixed effects model (their Table 15) confounds the effects of harvest rate and location. If the site random effect is dropped from the model, the coefficient flips to positive and becomes highly significant (p=0.0001). The estimated coefficients for individual years are not shown in their Table 16 but they are *all* positive (Appendix B: Figure B1); most are not individually significant likely because of the small sample sizes. As in the full data set, coefficients for every set of 4 contiguous years are also positive and significant ($p < 0.05$). See Appendix B for a detailed analysis of this data set.

LM suggests that the results in Klein et al. (2012) are driven by a few data points and therefore are not reliable. Permutation tests and cross-validation results in Klein et al. (2012) show very clearly that the regression results are robust with regard to variable identification and sample selection.

LM: "*It appears that most of the differences between the low and high harvest rate categories are due to relatively high turbidity values in three or four watersheds (Figures 2 and 3; Table 5 [refers to KLB]).*" Only Figure 2 has any information that would be able to support such a statement but in fact the information does not support it. There are 6 data points in the High category that are above the highest value in the Low category, and there are 6 values in the Low category that are below the minimum value for the High category. These 12 points in non-overlapping regions represent the majority of the 15 watersheds in those two categories. The primary purpose of Table 5 and Figure 2 was to characterize the difference between turbidity in "Legacy" watersheds (which had the same value of CCE10-15 as pristine watersheds) from that in the other categories. The difference between turbidity exceedence for Low and High harvest rates was not tested in that context because harvest rates are a continuous variable more appropriately analyzed using regression. A glance at Figure 4 shows that statistical results are not being driven by 3 or 4 data points. Even if the four most turbid watersheds were discarded, a relationship remains. In a data set of this size statistical significance could depend on 3 or 4 values, but not in this case.

TRENDS IN ELK RIVER SEDIMENT AND TURBIDITY

The HRC Report (Sullivan et. al, 2012) attempted to show a reduction in sedimentation and turbidity from 2003 to 2011. The report used two methods of trend analysis: (1) analysis of exponential decay coefficients, and (2) mixed-effects regression. As detailed in Appendix C, the exponential decay analysis was based on a false assumption. Since erosivity and sediment loads are not proportional to one another, there is no reason to think their decay coefficients should be comparable; results were predictable from the assumptions irrespective of trends. The regression analysis was better, but Freshwater and Elk River watersheds were not distinguished, no scatterplots were presented, and the data were unnecessarily divided into event and non-event years. Several analyses of recovery rates (Lewis 2013) with greater sample sizes show clearly that there has been no reduction in sediment loads or concentration in either the North or South Forks of Elk River. Appendix D details the problems with the HRC regression analysis. Combining Elk River with Freshwater Creek gives the impression of overall declines, but only Freshwater Creek has experienced significant reductions in turbidity or sediment loading.

Figure 3 shows 10% turbidities for many of the stream gaging stations analyzed by KLB, brought forward in time with the latest available data. For a hydrologic perspective, annual maximum peak flow is plotted for one gage (JBW). Consistent with Lewis (2013), Freshwater Creek (FTR) turbidity has been declining since WY2006 despite some relatively high hydrologic stresses in recent years (WY2011 and 2013). However, Elk River sites continue to be the most turbid, with North and South Fork Elk River sites (KRW and SFM, respectively) the highest. Despite large differences in watershed size and geologic makeup, the three pristine watershed (LLM, PAB, and ESL) exhibit drastically lower turbidities than either legacy or actively managed watersheds. Notably, the chronic turbidity levels for this relatively large time span stack up very consistently with the harvest rate categories defined in KLB for WY2004 and 2005. (The only changes that might have occurred since then would be swaps between low and high harvest rates; the pristine and legacy watersheds will not have changed category.)

BMP EFFECTIVENESS AT PREVENTING CUMULATIVE WATERSHED EFFECTS

A primary question is ‘how well do best management practices (BMPs) reduce the harmful effects of timber harvesting and related activities?’ BMPs can, if implemented rigorously, greatly reduce hydrologic and sediment effects on any given plot of land. The problem with relying too heavily on current BMPs is that even with their rigorous implementation, removing trees from hillslopes causes hydrologic and physical changes, inevitably increasing sediment delivery to streams through multiple pathways and mechanisms. For every acre of harvested land, there will be an effect of some magnitude. BMPs leak sediment no matter how effective they are. The effects are at the least additive, causing cumulative watershed effects (CWEs) in the most basic sense. When harvest rates are low, CWEs in a watershed can remain within tolerable limits. When BMPs need to be applied over too much area, CWEs increase to levels that harm beneficial uses.

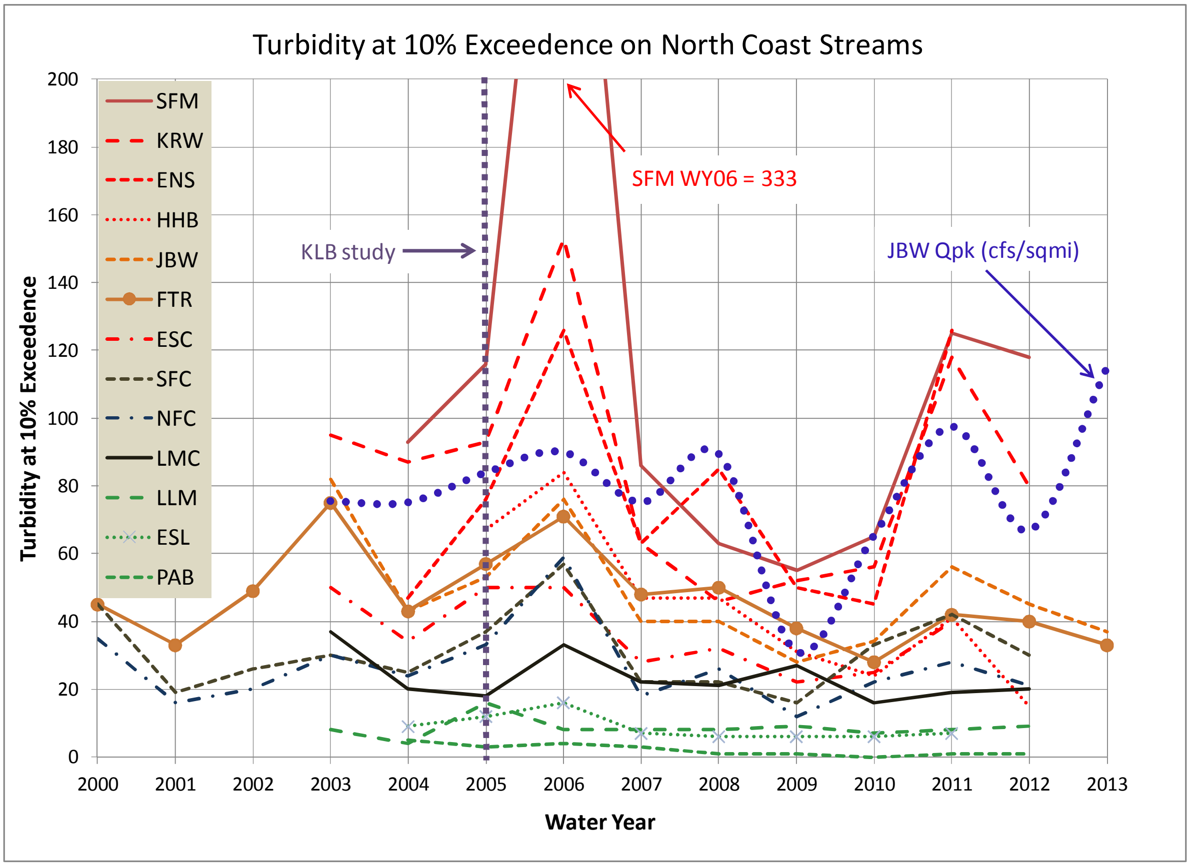


Figure 3. Turbidity at 10% exceedence for gaging stations analyzed by KLB, brought forward in time with the latest available data. Stream codes are the same as those used in KLB and lines are color-coded to indicate harvest rate categories from KLB (red = high; gold = low; black = legacy; and green = pristine).

LM suggests that studies be undertaken to assess the effectiveness of BMPs in Elk River. That sounds like a good idea that could produce useful results for refining BMPs. Problems with that approach are that proving effectiveness or ineffectiveness of specific BMPs is difficult if not impossible in an operational setting and, as discussed in the previous paragraph, does not address cumulative effects of harvesting with BMPs. The full effects of harvesting with BMPs cannot be known until all the effects have occurred, a period extending well beyond any relevant management time frame. For example, root decay minimums and resulting landsliding and soil pipe collapse may take ten or more years to occur. In addition, even quantifying short term BMP effectiveness will inevitably miss some of the negative impacts from harvesting with BMPs. Dispersed and short-lived evidence of surface, rill, and small gully erosion features are rapidly covered by leaf fall and rounding off of scarps and other indicators of past erosion. Inventorying and accurately quantifying erosional features occurring in areas where BMPs have been applied is a challenge even in experimental watersheds. In actively managed watersheds with a long history of disturbance such as Elk River the challenges are even greater. Judging the effectiveness of a BMP based on the occurrence of erosional features is highly subjective. BMPs are not applied one at a time on a landscape. If erosion is observed, which BMP is to blame? How do we know the erosion did not result from a legacy disturbance or natural processes? Attributing cause and effect without large numbers of before-and-after measurements on both control and treated sites relies entirely on field judgment. The dearth of reliable confirmatory evidence for BMP effectiveness is a concern in California, but there is good reason for its scarcity.

According to LM "*Recent data suggest that sediment from harvest-induced landslides is already being reduced by changing harvest practices, particularly in susceptible terrain (Sullivan et al., 2012; NCRWQCB, 2013), and the data in the Draft Peer Review indicate a reduction in sediment sources over time (Table 4.32, NCRWQCB, 2013. This means that BMPs, changes in management, natural recovery, and restoration efforts are reducing sediment production and delivery to stream*s." If indeed the propensity for new erosion and sediment delivery is lower today than several years ago, reduced timber harvest rates may be an important cause. However, sediment source inventories ignore the important role of storm history and the considerable lag between harvesting and some effects noted above (landsliding and soil pipe collapse). Analyses of suspended sediment data collected by HRC

(Appendix D) and Salmon Forever (Lewis, 2013) show that turbidity and suspended sediment concentrations and loads are not declining in the Elk River. These data are much more accurate than source inventories, and the analyses account for weather patterns through various characterizations of rainfall and discharge.

LM goes on to state *"Their conclusions with respect to the validity of **current** BMPs are also undermined by the fact that the strongest relationship between 10% exceedence turbidity was with ECA from 1990-1994, but forest practices and BMPs have changed substantially over the intervening twenty years (see Sullivan et al., 2012; NCRWQCB, 2013; California's Forest Practice Rules from 1990 and 2013). In the absence of any specific data on the implementation and effectiveness of **current** BMPs, Klein et al. (2011) cannot be used as proof that current BMPs are inadequate."* As illustrated in Figure 1 above, discussed in Appendix B, and acknowledged by LM (see next paragraph), harvest windows other than the 1990-1994 period were also well correlated to 10% turbidity. We did not conclude that current BMPs are invalid; however if harvesting by itself increases turbidity, then BMPs that ignore harvest rates cannot fully address the impact. The absence of specific data on BMP effectiveness is irrelevant to that argument, which is based on simple logic. Of course KLB did not prove that harvesting by itself causes turbidity to increase, but as discussed under "Causal Processes" above, we believe there is sufficient evidence from site-specific studies to support a causal relationship, and (thanks to recent improvements in stream monitoring technology) KLB provides the first peer-reviewed regional study corroborating the site-specific research with an empirical statistical relationship.

We agree fully with LM's statement "*Klein et al. (2011, p. 141) note that "Other models using just harvest rate (including annual mean harvest rate 0-15 years prior to the turbidity record) also performed well." This suggests that recent sources of sediment also are contributing to the observed turbidity, and that harvest-induced landslides due to a reduction in root strength almost certainly is just one part of the explanation for any relationship between forest management and turbidity.*" This supports our line of reasoning that using contemporary BMPs do not ensure healthy watersheds and avoid harm to beneficial uses. Limits on the rate of timber harvest offer a straightforward means of keeping CWEs in check and an opportunity for the Elk River watershed to recover.

As we suggest above, a statewide harvest rate BMP, customizable for individual watersheds, presents a relatively simple means to avoid repeating past mistakes like those which occurred in the Elk River watershed. Diversity among harvesting styles and inherent erosional stabilities can be accommodated and refined through time by adaptive management.

PROSPECTS FOR REDUCING SEDIMENT VIA HARVEST RATE CONTROLS

Manipulating harvest rates alone probably cannot reduce sediment and turbidity to the proposed TMDL allocations (i.e. 20% above natural conditions measured in Little South Fork) in a few decades (See Appendix B). Significant questions remain as to whether turbidity can be reduced dramatically (1) with ongoing timber harvesting or (2) without harvesting. Turbidity probably cannot be reduced more than 40% by reducing harvest rates (Figure B4), and reductions may take longer than predicted (Figures B5-B6). In KLB, legacy watersheds with no harvesting in the past 15-45 years have significantly higher chronic turbidity than pristine watersheds, although the range of turbidity overlapped considerably. In the low harvest rate category (< 1.4%), only 1 of 6 sites fell within the range of the pristine watersheds in 2004 and 2 of 7 in 2005.

Upper Salmon Creek was added to the Headwaters Forest Reserve in 1999 and provides another reference point for Elk River. The watershed, which borders Elk River to the south and is geologically very similar, was the subject of a recent master's thesis (Bailey, 2013). In 2012 (an average year, slightly drier than 2005), the 10% exceedence turbidity remained at 34.4 NTU, higher than any legacy watersheds studied in KLB. Technically Salmon Creek would have been classified in KLB's low harvest rate rather than the legacy group, since a small part of the watershed had been logged 15 years earlier. Contributing to the relatively high turbidity, a large landslide was triggered below a recontoured road in March 2012. Because its setting and history are so similar to Elk River's, and logging has been permanently eliminated, continued monitoring in Upper Salmon Creek is highly desirable and would be very relevant to learning about recovery rates possible in Elk River.

LOAD ALLOCATIONS

Because of the very large interannual variability caused by the sequence of rainfall and infrequent extreme events, it is difficult to accurately determine the long-term mean annual sediment load or turbidity 10% exceedence level in most watersheds. This problem has been pointed out in the other comment letters. If a target is defined as the mean annual value of some such parameter derived from a monitoring program, one extreme weather event could easily mean the difference between compliance and non-compliance, regardless of watershed management. A more practical approach to specifying load allocations would be to account for annual precipitation, runoff, or some combination of driving factors that explains most of the natural variability in loads. The models presented in Lewis (2013) provide a means for determining management-related trends in suspended sediment concentrations and storm event loads. These models use covariates such as antecedent rainfall, storm peak flow, and storm flow volume to account for weather-induced variability so that management-related variability can be evaluated. If the response in such a model is a log-transformed variable, trends can be expressed as a percent change between any two points in time, independent of the driving (weather) factors. Therefore we recommend that TMDL allocations (targets) be expressed as a percent change starting at a specific date, for specific watersheds where current monitoring systems can be used to verify compliance. The obvious watersheds are North and South Fork Elk River where both HRC and the residents maintain gaging stations.

WHAT CAN BE DONE TO MAKE A DIFFERENCE NOW?

While harvest rate limitations can reduce sediment and turbidity, recovery mechanisms are needed beyond management of streamflow and slope stability through harvest ceilings. Erosion sources must be stabilized or eliminated, and much progress is being made where practical in that arena. For the residents, perhaps the most serious concern is flooding and that can only be relieved in the short term through impoundments (anathema) or modification of the channel and riparian corridor. Flushing sediment out of the lower reaches will take years but can and should be accelerated by limiting channel encroachment by riparian vegetation. Dredging in the vicinity of the flood-prone properties could provide a quick fix but would create a great deal of disturbance, and the channels are likely to refill unless managed to be self-scouring. Removal and ongoing management of riparian vegetation would be less disruptive and would allow the channels to clean themselves over the long term. (Caveat: Our recommendations for the riparian corridor are based on a general knowledge of conditions in the nuisance reaches of Elk River, rather than specific fluvial/geomorphological investigations).

REFERENCES

Bailey, Kathleen N. 2013. Characterization of discharge, turbidity, and suspended sediment, Upper Salmon Creek Watershed, Humboldt County, California. M.S. Thesis, Humboldt State University, May 2013.

Cafferata, P., and L. Reid, 2013. Applications of long-term watershed research to forest management in California: 50 years of learning from the Caspar Creek Experimental Watersheds. California Forestry Report No. 5, The Natural Resources Agency, Sacramento, CA. 110 pp.

Dhakal, Amod S., and K. Sullivan, 2014. Shallow groundwater response to rainfall on a forested headwater catchment in northern coastal California: implications of topography, rainfall, and throughfall intensities on peak pressure head generation. Hydrological Processes 28: 446-463.

Keppeler, Elizabeth T.; and D. Brown. 1998. Subsurface drainage processes and management impacts. Proceedings of the conference on coastal watersheds: the Caspar Creek story, 1998 May 6; Ukiah, CA. General Tech. Rep. PSW GTR-168. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 25-34.

Klein, R.D., J.Lewis, and M.S. Buffleben. 2012. Logging and turbidity in the coastal watersheds of northern California. Geomorphology 139-140: 136-144.

LaHusen, R.G. 1984. Characteristics of management-related debris flows, northwestern California. Pp. 139-145 in: O'Loughlin, C.L. and A.J. Pearce (Eds.). Symposium on Effects of Forest Land Use on Erosion and Slope Stability, 7-11 May 1984, Honolulu, HI. Environment and Policy Institute, East-West Center, University of Hawaii, Honolulu, HI. 310 p.

Lewis, Jack and Elizabeth T. Keppeler. 2007. Trends in streamflow and suspended sediment after logging, North Fork Caspar Creek. In: Redwood Science Symposium: What does the future hold? March 15-17, 2004, Rohnert Park, CA. General Tech. Rep. PSW GTR-194. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 11 pp.

Lewis, Jack. 2013. Salmon Forever's 2013 Annual Report on Suspended Sediment, Peak Flows, and Trends in Elk River and Freshwater Creek, Humboldt County, California. SWRCB Agreement No. 07-508-551-0.

Lisle, T.E., L.M. Reid, and N.J. Dewey. 2008. Effects of clearcut logging on channel form in a coastal redwood forest. Abstract H13C-0932, American Geophysical Union Fall Meeting, Dec 2008; San Francisco, CA.

Montgomery, D.R., K.M. Schmidt, H.M. Greenberg, and W.E. Dietrich. 2000. Forest clearing and regional landsliding. Geology 28(4): 311-314.

NCRWQCB, 2013. Peer review draft: staff report to support the technical sediment total maximum daily load for the Upper Elk River. North Coast Regional Water Quality Control Board, Santa Rosa, CA.

Reid, L.M., N.J. Dewey, T.E. Lisle, and S. Hilton. 2010. The incidence and role of gullies after logging in a coastal redwood forest. Geomorphology 117: 155-169.

Reid, L.M. 2012. Comparing hydrologic responses to tractor-yarded selection and cable-yarded clearcut logging in a coast redwood forest. Pp. 141-151 in: Standiford, R.B.; T.J. Weller, D.D. Piirto, and J.D. Stuart. (Technical coordinators). Proceedings of the Coast Redwood Forests in a Changing California: a Symposium for Scientists and Managers. General Technical Report PSW-GTR-238. USDA Forest Service, Pacific Southwest Research Station. Albany, CA. 675 p.

Reid, L.M. and E.T. Keppeler. 2012. Landslides after clearcut logging in a coast redwood forest. Pp. 153-162 in: Standiford, R.B.; T.J. Weller, D.D. Piirto, and J.D. Stuart (Technical coordinators). Proceedings of the Coast Redwood Forests in a Changing California: a Symposium for Scientists and Managers. General Technical Report PSW-GTR-238. USDA Forest Service, Pacific Southwest Research Station. Albany, CA. 675 p.

Reid, Leslie M. and Jack Lewis. 2007. Rates and Implications of Rainfall Interception in a Coastal Redwood Forest Pp.107-117 in: Standiford, Richard B.; Giusti, Gregory A.; Valachovic, Yana; Zielinski, William J., Furniss, Michael J., technical editors. 2007. Proceedings of the Redwood Science Symposium: What does the future hold? March 15-17, 2004, Rohnert Park, CA. General Tech. Rep. PSW GTR-194. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture.

Sullivan, K., D. Manthorne, R. Rossen, and A. Griffith, 2012. Trend in sediment-related water quality after a decade of forest management implementing an Aquatic Habitat Conservation Plan. Humboldt Redwoods Company, Scotia, CA. 187 pp.

APPENDIX A. ROOT BIOMASS IN REDWOOD AND MIXED CONIFER FORESTS
(Robert R. Ziemer and Jack Lewis, when employed at PSW Redwood Sciences Lab)

SUMMARY

The death and decomposition of tree roots following logging might reasonably be expected to result in a loss of soil cohesion due to weakened root reinforcement. Ziemer (1981a) has shown how under certain circumstances this can lead to slope failure. To test his hypotheses, Ziemer investigated the abundance and strength of live and dead roots of various species in the laboratory (Ziemer, 1981b). Soil samples were collected from redwood and mixed conifer forest stands at various stages of succession following logging and the roots separated and weighed. A cluster sampling method was used within each forest stand. Previous publications reported only on the mixed conifer data. This paper summarizes Ziemer's root biomass data for both mixed conifer and redwood stands and investigates optimum cluster size and required samples sizes for future sampling designs.

METHODS

Studies of tree roots have been plagued by an inability to produce statistically significant results. The tremendous variances are seldom reported. Reynolds (1970) reported standard errors from ten random 425 cc augur samples ranging from 11 to 95 percent of the mean for less-than-6 mm root biomass. His average was about 30 percent and he estimated that 100 samples might be required to show significant differences between depths or concentric zones around a tree. Compared with our findings, that appears to be an optimistic estimate. Ziemer (1981) estimated that if 3200 cc samples were used, sample numbers on the order of 10^5 would be required to determine a trend in biomass related to time after logging.

It was on that basis that a sampling unit of one square meter by 1-1/3 m depth was chosen the root biomass study. The large sampling unit was designed to reduce the number of required samples to a manageable level. Using such a large sampling unit however made simple random sampling difficult, and a cluster sampling design was chosen instead.

Random samples of 2-4 clusters were selected from 6 redwood and 7 mixed conifer stands selected to represent successional stages following logging. Within each forest type, stands of similar soils, surrounding vegetation, climate, and management history were selected. The redwood stands included old growth and ages 5, 11, 24, 43, and 65 years. The mixed conifer stands included old growth and ages 3, 5, 7, 12, 20, and 24 years.

Each cluster consisted of several one-meter sampling units contiguous to an unsampled two-square-meter access pit. Sampling units blocked by tree trunks or boulders were not sampled. This resulted in varying cluster shapes and sizes. The number of sampling units measured in each cluster varied from 4 to 11. Most of the redwood clusters had 6 sampling units. The soil from each sampling unit was screened, and roots were extracted, separated into live and dead components, washed, sorted into 6 size classes, dried at 70° C., and weighed.

ESTIMATING BIOMASS AND ITS VARIANCE

When analyzing a cluster sample with unequal-size clusters, the best approach is often to assume a linear relationship of cluster total (y_i) to cluster size (m_i) and employ a weighted regression approach, wherein weights are inversely proportional to the variance about the regression line. One of three variance models is usually employed: (1) $\sigma_i^2 \propto m_i$, (2) $\sigma_i^2 \propto m_i^2$, or (3) σ_i^2 constant. These variance models result in different estimators. If the number of sampled clusters is n, then the best linear unbiased estimators (BLUE) for the cluster mean per sampling unit under the three models are, respectively:

$$\bar{y}_{CL,rm} = \sum y_i \Big/ \sum m_i \quad \text{(ratio of means)} \tag{1}$$

$$\bar{y}_{CL,mr} = \frac{1}{n} \sum y_i / m_i \quad \text{(mean of ratios)} \tag{2}$$

$$\bar{y}_{CL,reg} = \sum y_i m_i \Big/ \sum m_i^2 \quad \text{(regression)} \tag{3}$$

The three estimators are equivalent when m_i is constant. The ratio of means estimator is equivalent to the grand mean. Because clusters with more sampling units influence this estimator more than smaller clusters, the ratio of means estimator is biased, but it is often the most precise of the three. The mean of ratios estimator is equivalent to the mean of cluster means and is unbiased. The regression estimator is not considered here because, for root biomass, the constant variance assumption is unrealistic. The variance of cluster biomass is expected to increase with cluster size, and, if sampling units are independent, variance should be proportional to cluster size. Because there was significant variance among clusters and cluster sizes varied from 4 to 11, the risk of bias in the ratio of means estimator seemed considerable. Therefore we elected to employ the mean of ratios estimator, i.e. the mean of cluster means, for estimating root biomass.

An estimator for the variance of any of the $\bar{y}_{CL}$ estimators is:

$$\hat{V}\left[\bar{y}_{CL}\right] = \frac{N-n}{N} \frac{s_{BLUE}^2}{n\bar{M}^2} \tag{4}$$

where N is the number of clusters in the population, $\bar{M}$ is the mean cluster size in the population, and

$$s_{BLUE}^2 = s_y^2 - 2\bar{y}_{CL} s_{my} + \bar{y}_{CL}^2 s_m^2 \tag{5}$$

in which s_m^2, s_y^2, and s_{my} are the sample variances and covariance between m_i and y_i. In our stands we did not know N or $\bar{M}$. These are not an inherent property of the stands, since cluster sizes were based on practicalities, but the mean sampled cluster size ($\bar{m}$) can be substituted as an approximation for $\bar{M}$. Also, we know that n is small relative to N so the finite population correction factor, $(N-n)/N$, can be ignored. We therefore used the approximate variance expression:

$$\tilde{V}\left[\bar{y}_{CL}\right] = \frac{s_{BLUE}^2}{n\bar{m}^2} \tag{6}$$

The standard error of the estimated cluster mean per sampling unit is then $s_{BLUE} \Big/ \left(\bar{m} n^{0.5}\right)$.

RESULTS AND DISCUSSION

In the following sections, all biomass values refer to live root biomass unless explicitly stated otherwise. Dead roots comprised a very small portion of the total biomass in old growth stands (particularly in redwood) for all sizes and depths (Figs. A1 and A2).

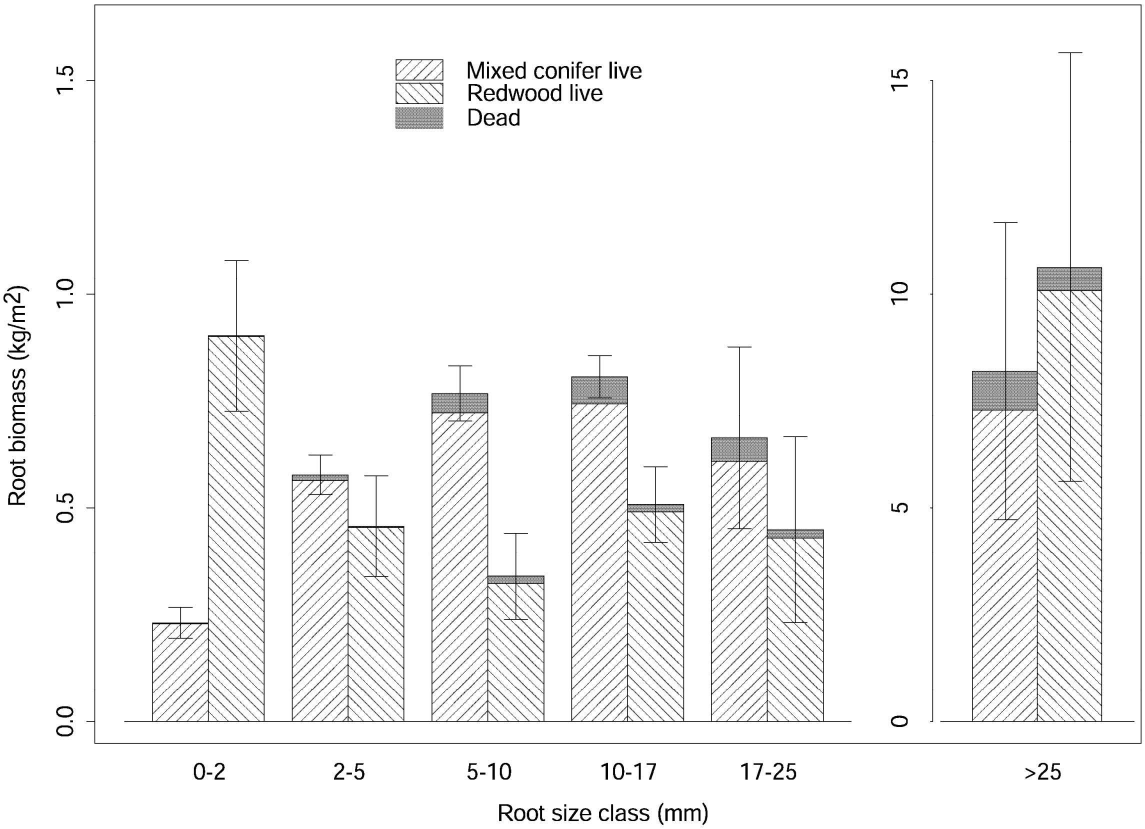


Figure A1. Live and dead root biomass by size class, all depths combined. Error bars designate one standard deviation.

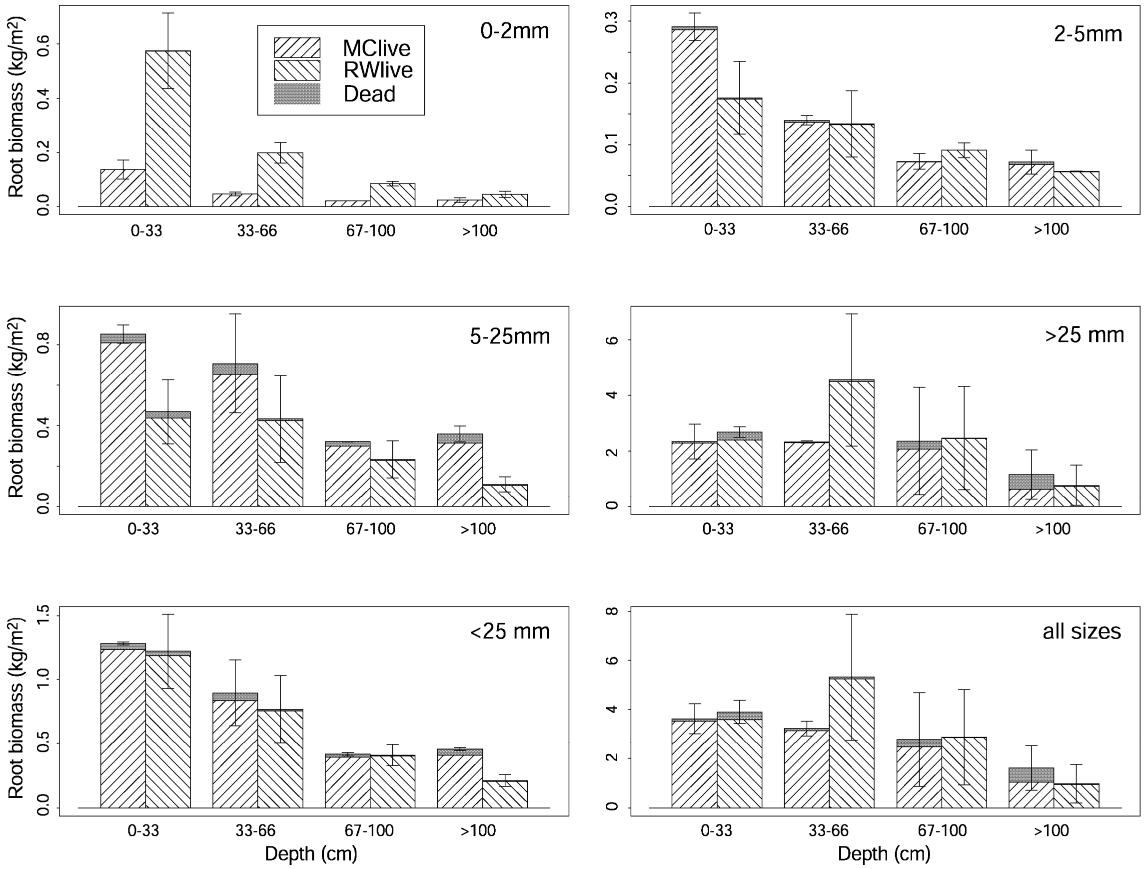


Figure A2. Live and dead root biomass by depth and size class. Error bars designate one standard deviation.

Total and Fine Root Biomass in Old Growth Stands.

Total live root biomass was 10.2 kg/m^2 in the uncut mixed conifer stand and 12.7 kg/m^2 in the virgin redwood. By comparison, in old-growth coniferous forests dominated by Douglas-fir (Grier and Logan, 1977; Santantonio et al., 1977) total root biomass ranged from 10.5 to 20.9 kg/m^2. Outside of these studies, the largest reported value for root biomass in coniferous forests has been 8.5 kg/m^2 in a 200-year-old stand of spruce (Picea abies) in the USSR (Santantonio et.al., 1977). On the other hand, in some tropical and subtropical forests, higher values up to 32.8 kg/m^2 have been reported.

Biomass of fine (less-than-5mm) roots was 0.79 kg/m^2 in the uncut mixed conifer stand and 1.35 kg/m^2 in the virgin redwood. This latter value exceeds nearly every value for fine roots yet reported in the literature, including tropical and subtropical forests. Fine root biomass in old growth forests dominated by Douglas-fir varied from 0.79 to 1.30 kg/m^2 (Grier and Logan, 1977). Studies from a wide variety of forests indicate surprising uniformity in fine root biomass, with values generally varying from 0.5 to 1.0 kg/m^2 in stands over 10 years old. Santantonio et al. (1977) indicates that fine root biomass often appears to reach a peak early in stand development, subsequently levelling off. However, neither of our forest types had reached their peak levels by the age of 24 years (Fig. A3a,c). It appears that coniferous forests of the Pacific Northwest, particularly redwood forests, are somewhat above average in fine root biomass. This could reflect extensive and persistent development of absorbing roots to exploit the plentiful soil moisture which is available throughout most of the year.

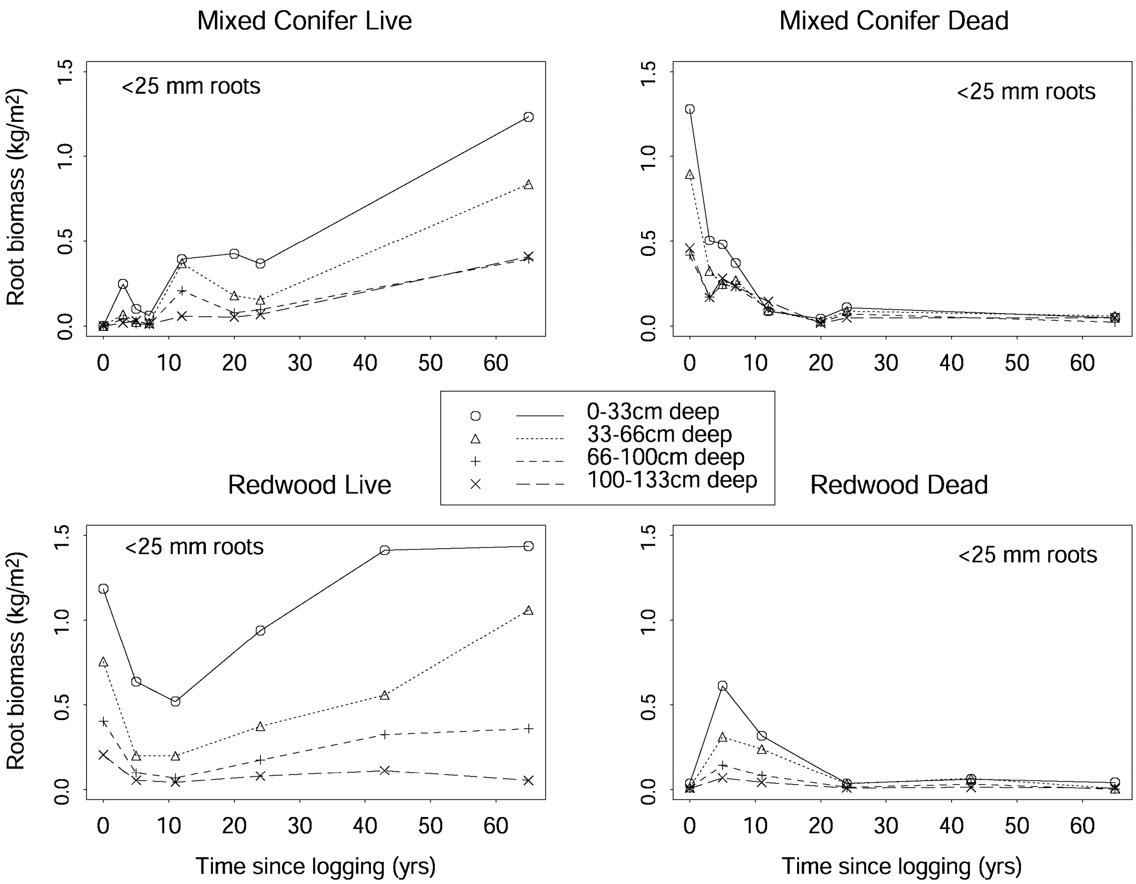


Figure A3. Live and dead fine root biomass (< 25 mm) by stand age.

Root Size Distribution in Old Growth Stands.

There is a marked difference in the size distribution of roots between the old growth redwood and mixed conifer stands (Fig. A1) Although total (live and dead) less-than-25 mm root biomass was very similar, biomass of less-than-2 mm roots was over 4 times as great in redwood. In less-than-2 mm roots redwood had an average 0.90 kg/m^2 contrasted with 0.23 kg/m^2 for mixed conifer forest. On the other hand, in 2-to-25 mm roots, redwood had only 1.75 kg/m^2 as opposed to 2.82 kg/m^2 in mixed conifer forest. Roots bigger than 25 mm in diameter were by far the largest component in terms of biomass, being about 10.6 kg/m^2 in redwood and 8.2 kg/m^2 in mixed conifer forest.

Depth Distribution of Roots in Old Growth Stands.

According to most studies to date, the majority of forest tree roots lie in the upper 50 cm of soil and most of the absorbing roots are in the upper 20 cm. This study is in general accordance with those, although our depths of measurement were in one-third meter increments. The distribution of roots does, however, depend on root size (Fig. A2). Less-than-2 mm roots are concentrated most heavily near the surface. Greater-than-25 mm roots are distributed evenly throughout the upper meter in mixed conifer and, in redwood, concentrated most heavily in the middle of the top meter, tapering off below a meter in both types. Intermediate size classes have intermediate depth distributions. Biomass is skewed towards the surface, but not as extremely as in the case of very fine roots. Considering all size classes, the depth distributions of redwood and mixed conifer forest are quite similar, with one possible difference. Mixed conifer had proportionally greater biomass below a meter in depth in nearly every size class. The individual differences are not all statistically significant, but biomass of all roots less than 25 mm in diameter was about twice as great in mixed conifer.

Changes in Live and Dead Root Biomass After Logging.

Root biomass in several different ages of cutblocks and second growth stands were charted along with the old growth forests (Figs. A3 and A4). Since the cutblocks were similar in soil type, depth, slope, aspect, elevation, original forest density, and silvicultural history, these graphs may be considered as representations of chronological development in the two forest types.

In mixed-conifer, the old growth is placed at 65 years after logging, on the assumption that most of the net change to old growth root biomass levels will have occurred by then. All roots are assumed dead immediately after logging. Thus live root biomass at age zero is plotted as zero and dead root biomass is plotted as the sum of live and dead roots from the old growth stand. The changes in live and dead roots reflect a successional sequence of extensive colonization with bracken fern (Pteridium aquilinum) by age 3, followed with brushfields by age 12. The overall trend of increasing live root biomass with time since logging was interrupted after each of these stages reached its peak. It seems unlikely that in any given cutblock an actual decrease would occur as a result of plant competition at such times. The declines shown in Figure A3 therefore probably reflect differences in succession between the cutblocks studied. Of course, if vegetation were killed in an attempt to establish conifers more quickly, a decline would be expected, but these cutblocks were not treated as such. The dead mixed conifer biomass curve reflects the decaying of roots that were killed by logging, with insignificant bumps corresponding to the decline of the fern and brushfields.

In redwood, the 65-year value is actually from a 65-year-old stand. Since redwood is a sprouting species, roots do not die immediately upon cutting. Thus the old growth live and dead root biomass are plotted at age zero. The live root biomass does, however, decline after logging as the roots come into equilibrium with the drastically reduced above ground biomass. Live less-than-25 mm biomass reached a minimum 11 years after logging. Thereafter it gradually increased to pre-logging levels by age 65, except in the layer below a meter in depth. As with the mixed conifer areas, in this layer live root biomass less than 25 mm remained near or below 0.1 kg/m^2 in all but the virgin stands (Fig. A3).

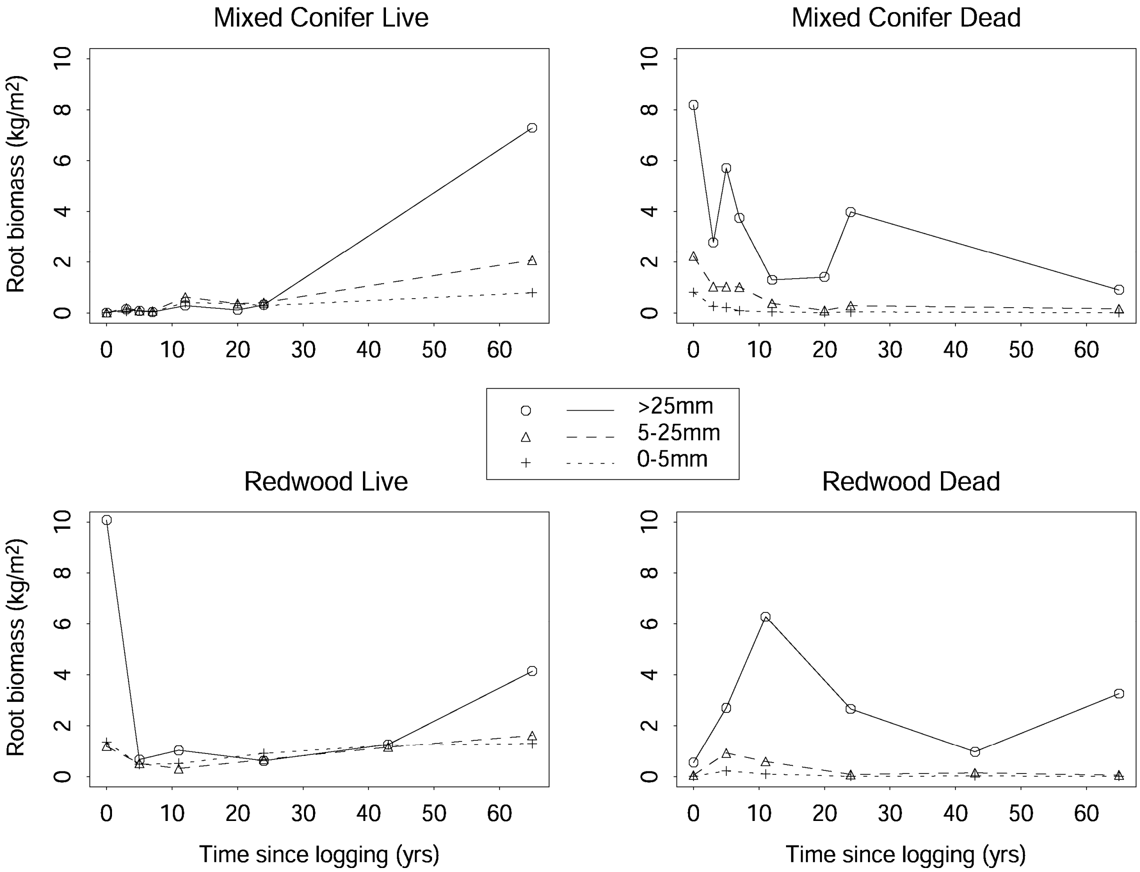


Figure A4. Live and dead fine root biomass (< 25 mm) by stand age.

Dead less-than-25 mm biomass in redwood areas peaked 5 years after logging, even though live root biomass did not appear to reach its minimum until age 11, particularly in the top soil layer. Apparently, decomposition of the large biomass component which had died in the first five years exceeded new senescence between years 5 and 11.

Live-plus-dead root biomass suffers a decline after clearfelling, and full recovery appears to take well over 25 years in the less-than-25 mm fraction (Fig. A5). Larger roots are even slower in returning to prelogging levels. In the mixed conifers the decline in roots appears to be more rapid and of greater magnitude than in the redwood forest. Biomass dropped from about 3.0 kg/m^2 to 1.5 kg/m^2 in only 3 years, and to 0.84 kg/m^2 after 20 years. In redwood, by contrast, biomass dropped from 2.7 kg/m^2 to 1.5 kg/m^2 in 11 years and thereafter began to increase again. Because some redwood roots survive logging, this is not a surprising result.

The following sections give methods and recommendations for optimum cluster size and sample numbers required to construct confidence intervals in future studies, based on the variance estimates from our clusters.

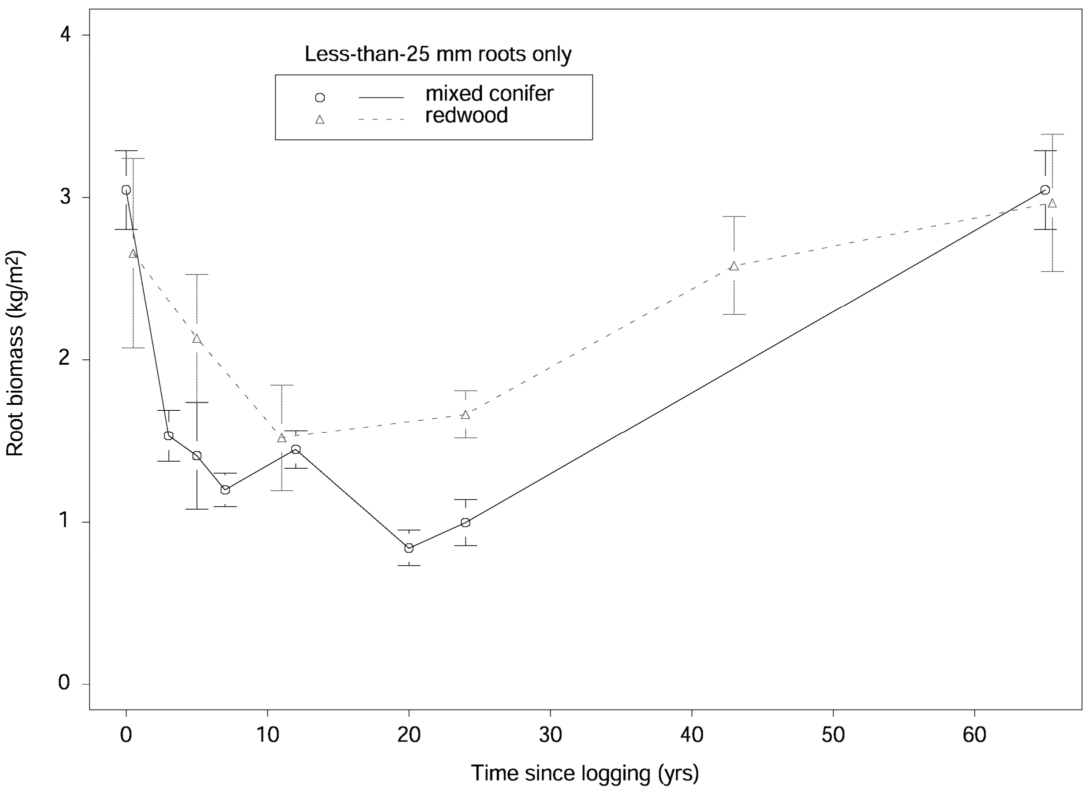


Figure A5. Combined live and dead fine root biomass (< 25 mm) by stand age.

THE REMAINDER OF THE ROOT BIOMASS REPORT IS A STATISTICAL ANALYSIS OF OPTIMUM CLUSTER SIZE AND REQUIRED SAMPLE SIZE TO ATTAIN VARIOUS LEVELS OF PRECISION. THESE SECTIONS HAVE BEEN OMITTED AS THEY ARE NOT RELEVANT TO THE ELK RIVER TMDL

REFERENCES: APPENDIX A

Cochran, William G. 1963. *Sampling Techniques* (2nd ed.). New York: John Wiley and Sons, Inc.

Grier, C. C. and R. S. Logan (1977). Old-growth Pseudotsuga menziesii communities of a western Oregon watershed: biomass distribution and production budgets. *Ecological Monographs* **47**: 373-400.

Guenther, William C. (1964). *Analysis of Variance*. Englewood Cliffs, New Jersey: Prentice-Hall. 199 pp.

Reynolds, E. R. C. (1970). Root distribution and the cause of its spatial variability in Pseudotsuga taxifolia (Poir.) Britt. *Plant and Soil* **32**: 501-517.

Roberts, J. (1976). A study of root distribution and growth in a Pinus Sylvestris L. (Scots Pine) plantation in East Anglia. *Plant and Soil* **44**: 607-621.

Santantonio, D., R. K. Hermann, and W. S. Overton (1977). Root biomass studies in forest ecosystems. *Pedobiologia* **17**: 1-31.

Ziemer, Robert R. (1981a). The role of vegetation in the stability of forested slopes. International Union of Forest Research Organizations, XVII World Congress. Kyoto, Japan.

Ziemer, Robert R. (1981b). Roots and the stability of forested slopes. Proceedings of the Christchurch Symposium, *International Assoc. of Hydrological Sciences Publ.* **132**: 343-361.

APPENDIX B. Re-analysis of HRC's data set relating harvest rate and turbidity. (Jack Lewis)

I computed the regressions in Table 16 of the HRC Report from the data in Appendix B of the Report and statistics matched fairly closely. The regression coefficients were not presented in the HRC Report, but here they are shown with error bars in Figure B1. For consistent presentation and in contrast with Table 16 of the Report, all the regressions for Figure B1 use the log-transformed response. Only 2004 and 2005 were significantly different from zero at the convention level of p=0.05, but all estimates were positive. The sample sizes in each year of this study are smaller than the sample size of 27 that we had for the regional study (KLB), and that might partly account for the lack of significance in the one-year models. The figure also hints that the impact of harvest rates on chronic turbidity may be declining, which is a reasonable hypothesis as selective harvesting and limitations on disturbance are implemented.

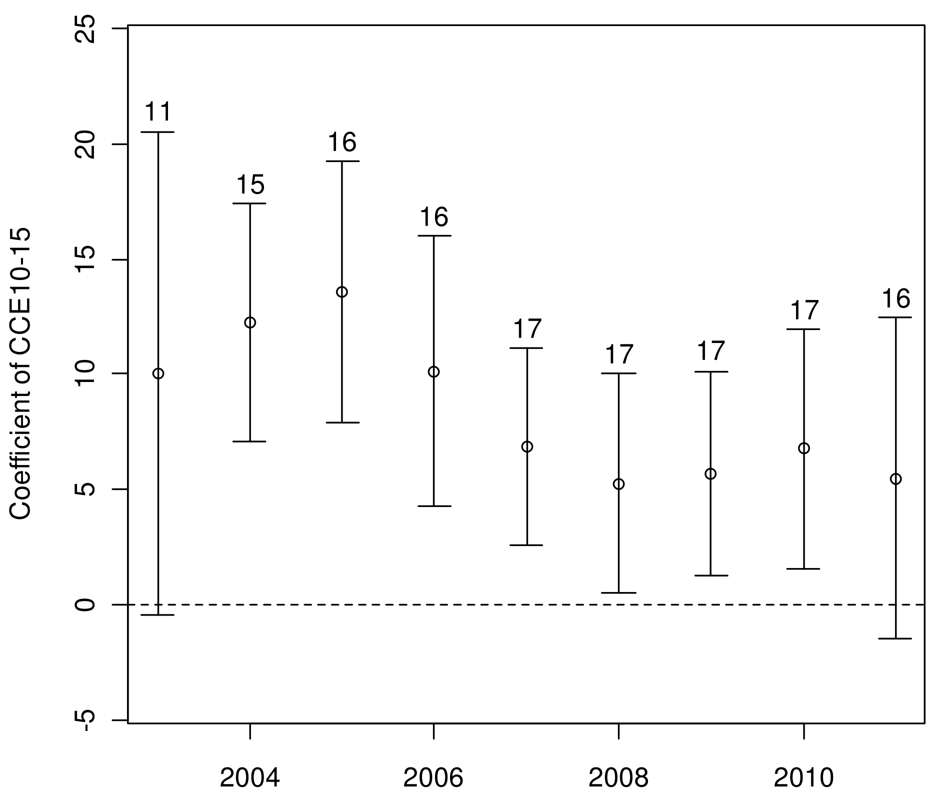


Figure B1. Regression coefficients from regression of log(T10) on basin area and 10-15 year historical clearcut equivalent area (CCE10-15). Error bars indicate plus and minus one standard error. The numbers on top of each bar indicate the sample size.

I was able to approximately reproduce the statistical model that combines all years (Table B1), shown in Sullivan's Table 15. There are apparently some small differences in the data sets used in computation for the report and that shown in its Appendix B[1], but the models appear to be in very close agreement (Table

[1] The HRC Report says one observation at site 509 caused some concern but my interpretation of the text is that all observations from that site were removed. Nonetheless, I obtained best agreement the HRC model when site 509 was included.

B1). Both Site and Year are represented as random effects in this model. A random effect is realized as a series of constants, conceptually representing the mean response for each level of the random effect (Site or Year). The constants are assumed to belong to a normal distribution describing the variability of the response among Sites or Years. Only one degree of freedom is used for each random effect, i.e. to estimate the variance of its distribution.

Table B1.

Model	Term	Estimate	Std Error	DF	t	Pr > |t|
Sullivan	Intercept	3.632	0.1947	8	18.65	<0.0001
Sullivan	Area	0.0103	0.0046	116	2.24	0.0272
Sullivan	CCE10-15	-3.270	1.3527	116	-2.42	0.0172
Lewis repro	Intercept	3.627	0.1951	139	18.60	<0.0000
Lewis repro	Area	0.0104	0.0046	139	2.27	0.0249
Lewis repro	CCE10-15	-3.286	1.3563	139	-2.42	0.0167
Lewis alt	Intercept	3.398	0.1369	139	24.82	<0.0001
Lewis alt	Area	0.0096	0.0014	139	6.75	<0.0001
Lewis alt	CCE10-15	7.039	1.7514	139	4.02	**0.0001**

A puzzling fact about these models is that the coefficient of CCE10-15 is significantly less than zero, while we saw (Figure B1) that the estimates for each individual year were positive. An alternative model (Lewis alt) is presented in Table B1 that estimates a random effect for Year but not for Site. *Without the Site random effect, CCE10-15 has a positive coefficient that is highly significant.*

Why would the sign of the coefficient be dependent on the inclusion of the Site random effect? Some watersheds did not experience a very wide range in harvest rates during the study period (Figure B2). For example sites 534 and 550 had CCE10-15 equal to zero every year, while 500 and 526 never experienced a value above 1%. Sites 502, 509, 510, and 523 also saw a narrow range of harvest rates. Statistically speaking there is no way to tell whether the low turbidity at site 534 is due to the zero harvest rate or to inherent factors of the location. The very small random effect constant for that site accounts for the low turbidity, leaving nothing to be explained by harvest rate. Including site as a random effect assumes that these sites with low harvest rates and low turbidity are inherently stable to begin with; if their low turbidity is accounted for by the random effect, the harvesting will appear to be unimportant or even inversely related to the response. This is the classic problem of multicollinearity in regression modeling, except that one of the effects in this case was modelled as a random effect. When variables in a model are highly correlated, their coefficients cannot be reliably interpreted and their signs may even be opposite of their true effect. Figure B3 shows the random effects constants for each site in relation to the site means for CCE10-15 and watershed area. The effect of location is highly correlated to CCE10-15 ($r = 0.619$) but not to watershed area. If the fixed effects being investigated are correlated with a random effect, it is best to leave the random effect out of the model. Otherwise its effect may be hidden or inverted by the confounding random effect. That appears to be exactly what happened in this case. While one could argue the relative importance of inherent site factors and harvest rates based on physical arguments, it is incorrect to dismiss the role of harvest rates based on the negative coefficient in the confounded model. Yes, location typically plays a role in this type of model; but harvest rates were associated with turbidity in both the regional data set (Klein et al., 2012) and the HRC data set. Cross-validation and permutation tests on the regional data set left little doubt that the association was not fortuitous.

The degrees of freedom differ but the coefficients are in very good agreement. Degrees of freedom seem to be reported differently by the two statistical packages: SAS and R. My model included 142 observations.

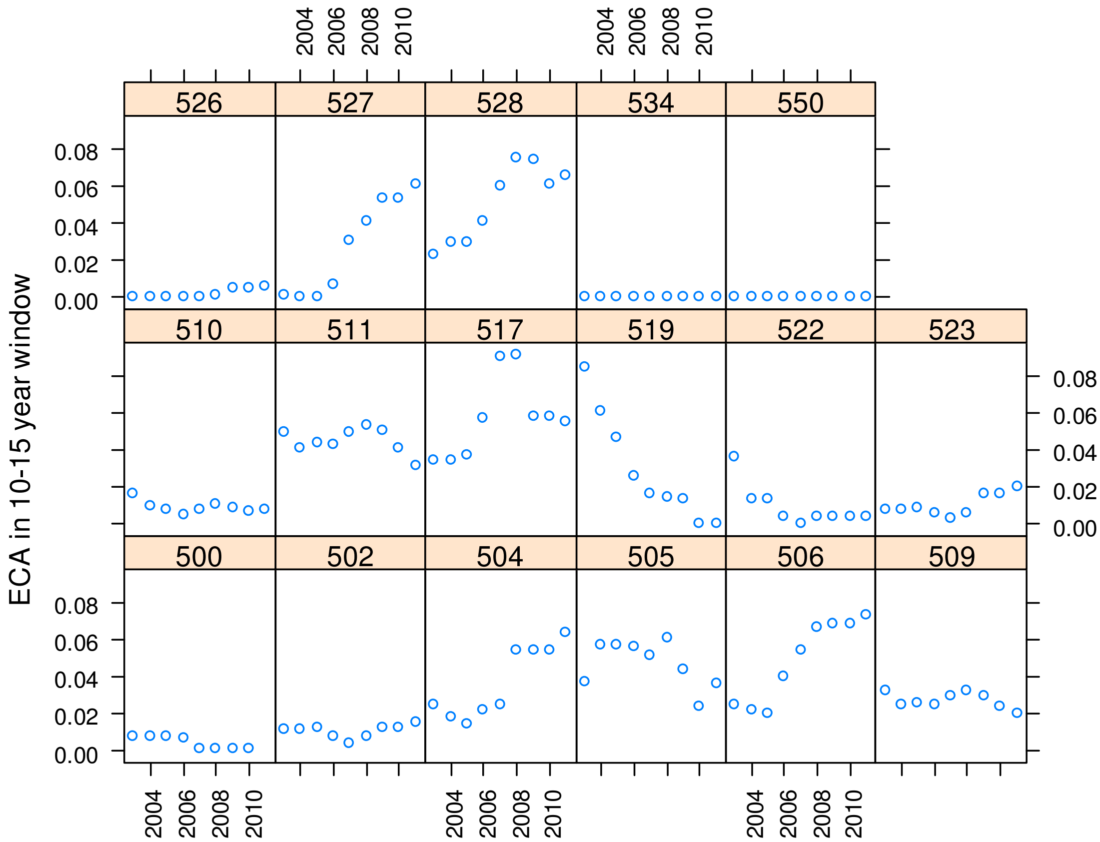


Figure B2. The distribution of CCE10-15 at each HRC gaging site.

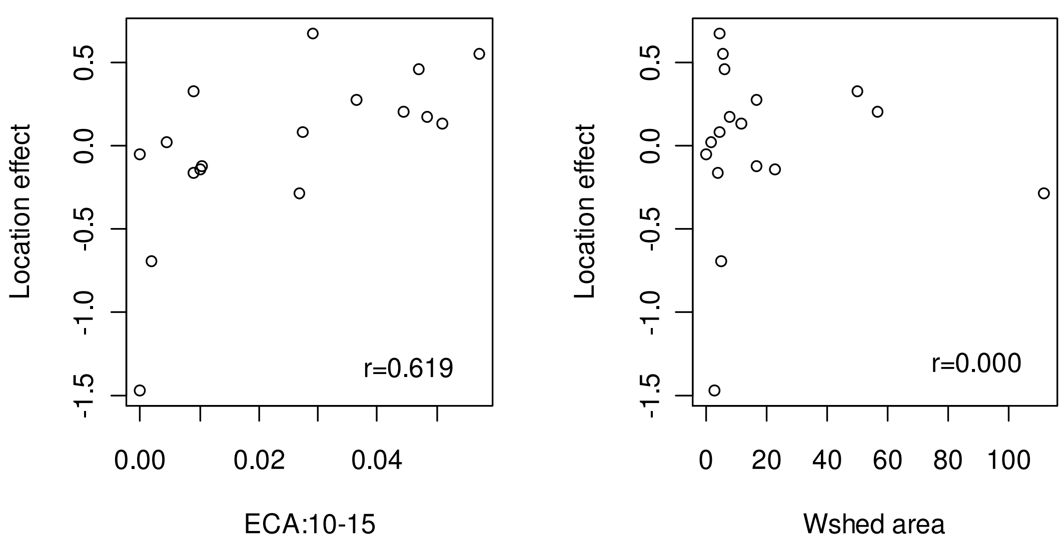


Figure B3. Scatterplots showing the correlations of harvest rate (10-15 year window) and watershed area with the estimated random effect of site location.

Did the effect of harvesting decline in the latter years?

The decreased sensitivity to harvest rate with time, hinted at by Figure B1, was not statistically significant (p=0.073) when the interaction was tested in the mixed-effects model with Year as a random effect, or in any of several standard regression models that I tried with the same response variable. This could be due to lack of statistical power. However, in models using square root transformations on all the variables (except time), the interaction tested significant. Models limited to Elk River showed no evidence of an interaction (p=0.868 for the square root model).

When I ran the mixed model (with basin area and no Site random effect) for log(T10) on groups of four contiguous years at a time, I found that the CCE10-15 variable was positive and significant for all groups of years (Table B2), but the estimate of the coefficient and its significance level declined each year. (These are not independent tests, so the Bonferroni adjustment does not apply). So there is some evidence that the sensitivity of turbidity to harvest rate may be declining under HRC and GDRC management. This may be related to turbidity and sediment trends in Freshwater Creek, because as mentioned above there was no interaction evident in the Elk River data.

Table B2. Coefficients of CCE10-15 in mixed models for log(T10), by 4-year period

Years	Coefficient	p-value
2003-2006	11.5	0.0005
2004-2007	9.6	0.0003
2005-2008	7.4	0.0030
2006-2009	6.3	0.0077
2007-2010	6.0	0.0085
2008-2011	5.6	0.0304

Other harvest windows

The main significance of the Klein et al (2012) paper was to establish a robust statistical association between harvest rates and chronic turbidity; it did not definitively establish the relative importance of each 5-year window. Consideration of the underlying processes and results from Caspar Creek suggest that all these windows should be relevant. Developing a model with proper weightings for each year in the harvest history would require a much larger data set. However, in the KLB data set, CCE10-15 was the best predictor among several highly correlated harvest rate variables, specifically when included in a model with basin area. I tested the other harvest variables using the HRC data, which included windows of 1, 2, and 10 years, as well as the 15-year window based on the sum of the 10-year window and CCE10.15. With or without basin area in the model, the 1-year window was borderline significant, the 2-year window was very significant, the 10-year window was not significant, and the 15-year window was highly significant. Just as in the regional study (KLB), none was as good a predictor as the 10-15 year window. The coefficients suggest that other time periods have a smaller effect on turbidity. However, I found inconsistencies in the coding of the 10-year window in Sullivan et al. (2012) Appendix B that lead me to suspect there are errors in the data.

Specifying an appropriate harvest rate: limitations

If one believes the association of harvest rate with turbidity is not fortuitous, then from the coefficient of CCE10.15 in the "Lewis Alt" model and its 95% confidence interval, we can display the sensitivity of T10 to changes in the harvest rate (Figure B4). The figure provides a more direct method than that proposed in the TMDL for specifying a harvest rate to achieve a given turbidity target. The graph depicts the curves $y=\mathbf{exp}(c_i x) - 1$, where $c_1 = 7.04$ is the coefficient of the 10-15 year harvest window, **exp** denotes exponentiation to the power e=2.718, the constants c_2 and c_3 are the upper and lower 95% confidence limits on c_1, and x and y are coded as proportions. The curves represent the range of expected change in turbidity 10-15 years after a change in harvest rate. The coefficient used to produce Figure B4 has a wide confidence interval but is reasonably stable. If the Year random effect is dropped from the model the interval changes from (3.6,10.5) to (2.1,10.3). Despite the uncertainty, Figure B4 makes it clear that changes in harvest rate provide limited leverage on turbidity. Given a harvest rate reduction of 4% per year, the maximum (>95% confidence) reduction in 10% turbidity that can be expected is 34%.

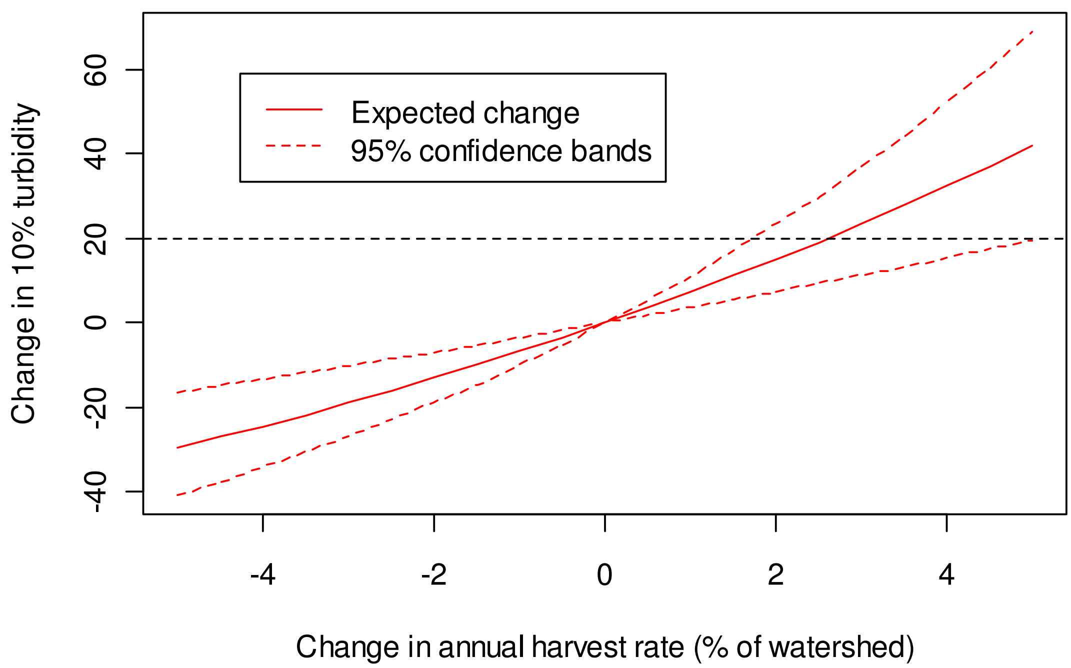


Figure B4. Expected effect of a change in harvest rate on 10% turbidity exceedence. The horizontal dashed line indicates the 20% regulatory limit for turbidity increases in California.

To make the situation worse, this model neglects increases in the first 10 years after harvest, and likely underestimates recovery time after reductions in harvest. The reduction lag time will depend on the processes that are contributing sediment. For in-channel processes, Cafferata and Reid (2013) argue that sediment recovery should lag behind flow recovery, as erosion processes may be operating in an expanded channel network, and streambanks may have been destabilized. In addition, full flow recovery lags harvesting by at least a decade (Lewis and Keppeler 2007). In relation to landslide risk, hillslope hydrology might largely recover in a similar time frame as flows, but root biomass recovery takes more than a decade for <25mm roots and many decades for larger roots. And, once triggered, stabilization of landslides can take many years. Figures B5 and B6 show the observed and predicted T10 for sites 519 and 522, sites that experienced the greatest declines in CCE10-15 during the study period. In both cases, the decline in predicted values is steeper than the decline (if any) in observed values. If this is generally the case then predicted decreases can be expected to lag behind reduced harvest rates by more than 10-15 years. The trend at site 519 is complicated by splash dam removal in 2005 (Sullivan et al., 2012) which may not be a typical disturbance but still could exemplify delayed recovery after destabilization.

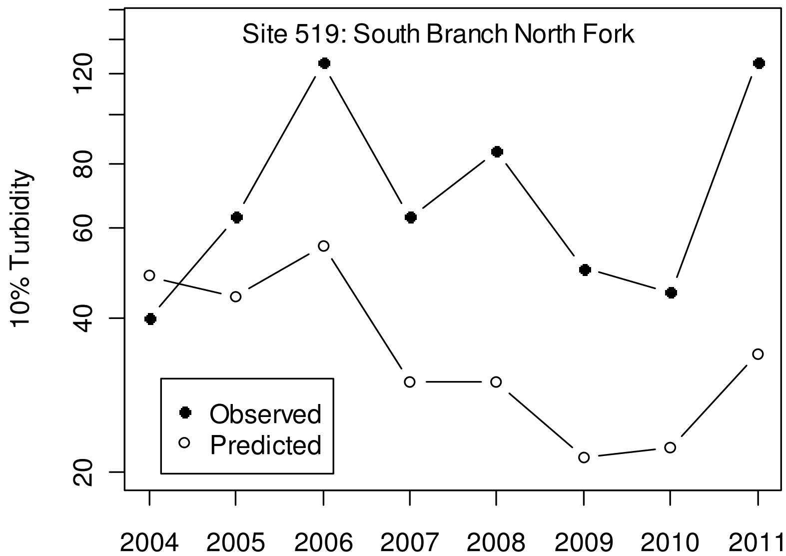


Figure B5. Predicted and Observed 10% exceedence turbidity at site 519, based on mixed-effects model (Table B1, Lewis Alt).

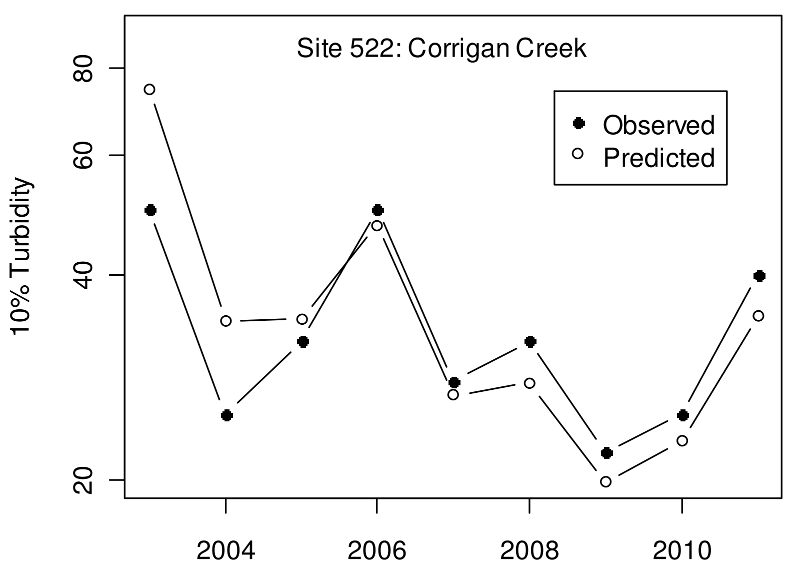


Figure B6. Predicted and Observed 10% exceedence turbidity at site 522, based on mixed-effects model (Table B1, Lewis Alt).

APPENDIX C. Trend Analysis Using Exponential Decay Coefficients
(Jack Lewis)

The HRC Report, in Appendix C, gives equations for Sediment Yield as a function of (1) Erosivity and (2) Sediment Yield as a function of Time. The text asserts that, if there are no management-related trends in yield, the exponential decay coefficient for Erosivity as a function of Time should be the same as that for Sediment Yield. Comparing the decay rates of E and S doesn't really make sense because E and S are two different quantities with different units. In addition, the steepness (first derivative) of the curve is $-bke^{-kt}$, so cannot be judged by k alone. The following analysis examines the properties of this method.

First, the relationships between Sediment Yield and Erosivity are shown as linear regressions:

$$S = a_0 + a_1 E \qquad (1)$$

and the exponential decay relation is

$$E = be^{-kt} \qquad (2)$$

Substituting E into the linear regression gives a different exponential form

$$S = a_0 + a_1 be^{-kt} \qquad (3)$$

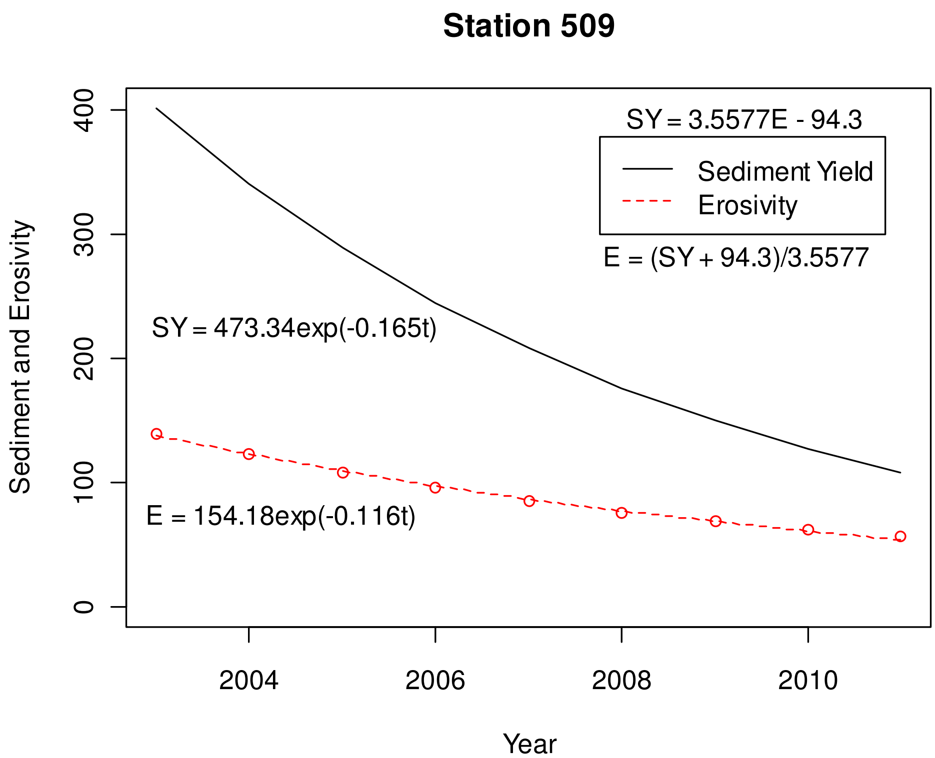
While equation (2) has a limiting value of 0 for large t, equation (3) has a limiting value of a_0. The original exponential decay form (2) is lost unless a_0=0, so it is not clear that the decay coefficient (k) would be unchanged if equation (2), with its zero asymptote, were forced to fit data conforming to the equation (3). But the assertion is easily tested by example.

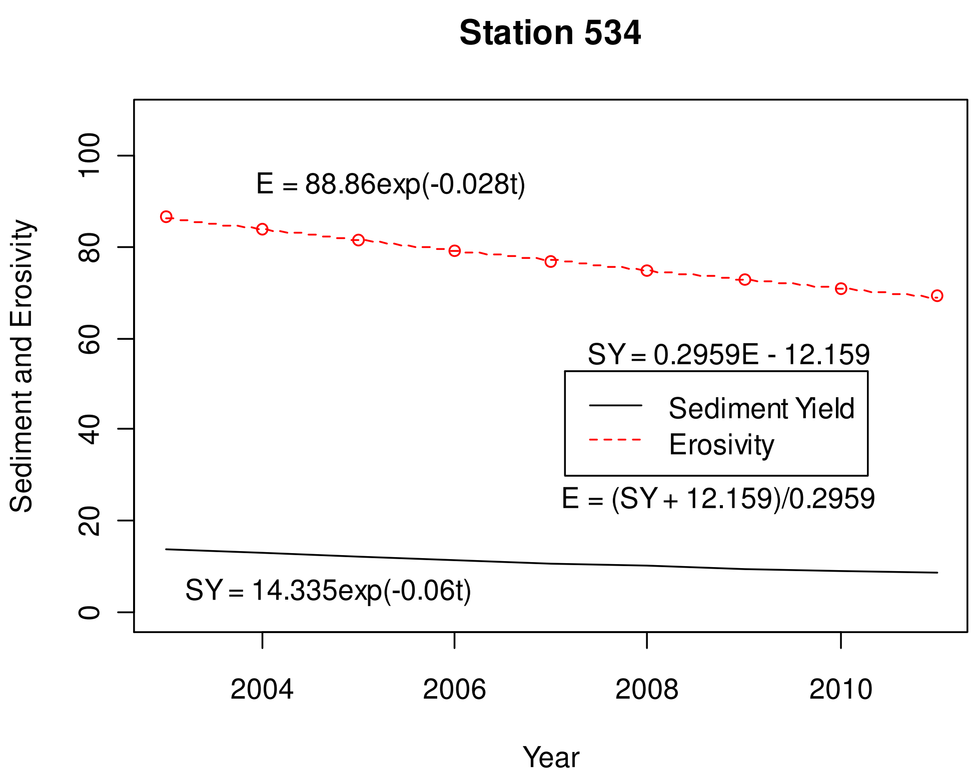
Since, in Appendix C, we are given the decay coefficients for Sediment Yield but not Erosivity, it is most convenient to work backwards, calculating "expected" Erosivity from Sediment Yield by inverting equation (1), and fitting equation (2) to expected Erosivity. I fitted equation (2) in R using nonlinear least squares. It is not a perfect fit unless the intercept of equation (1) is zero, in which case the decay parameter does not change. However the intercept a_0 is never zero (Table C1). Assuming an unchanging relationship (1) between Erosivity and Sediment Yield, the decay parameter (k_1) for Sediment Yield is smaller than that for expected Erosivity (k_2), at every gaging site (Table C1). *Reporting $k_1 < k_2$ as evidence of recovery is simply incorrect; it is to be expected as a consequence of the assumed relationships*. In the report, just one station (519) was reported as not showing a recovery in sediment. Station 519 is unique because it is the only station whose parameter k for Sediment Yield is positive (and whose intercept a_0 is negative). Its curve increases exponentially rather than decaying. However its curve for expected Erosivity actually *does* have a smaller k parameter than that for Sediment Yield; it just rises more steeply because coefficient b is much greater.

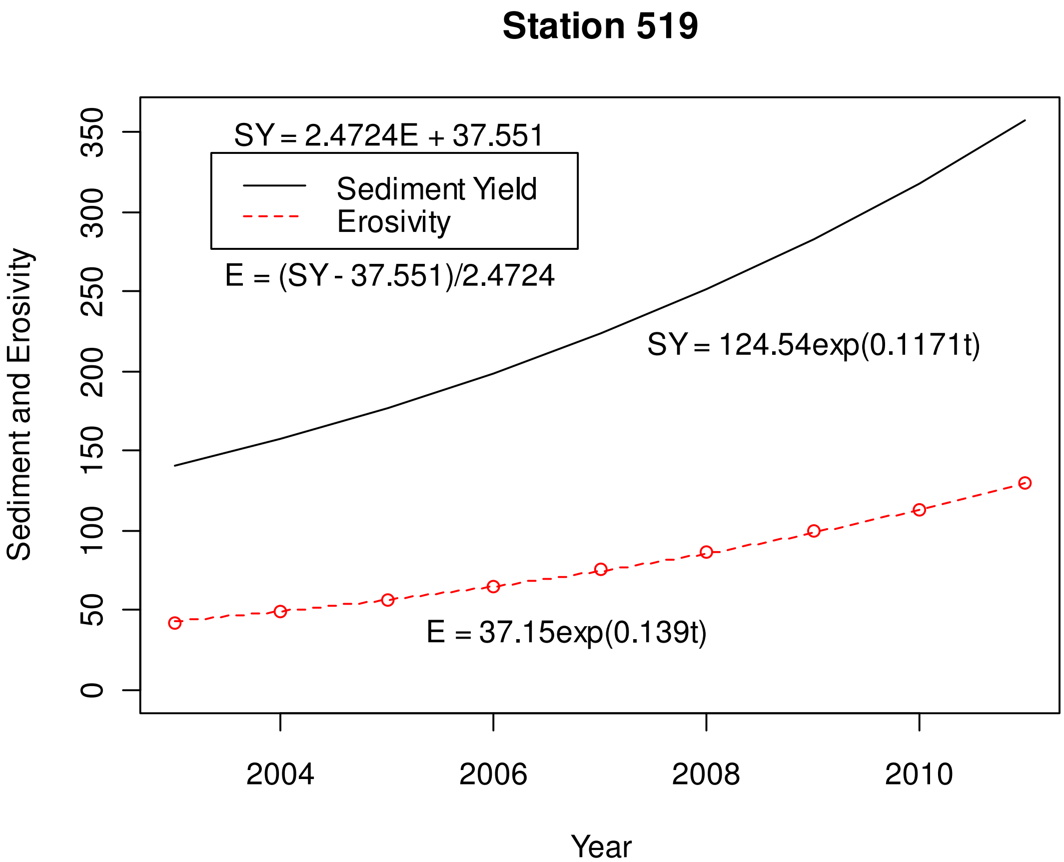
Gaging	*Eq(1)*		*Eq(2) Sediment Yield*		*Eq(2) Erosivity*		*Difference*
Station	a_0	a_1	b_1	k_1	b_2	k_2	$k_2 - k_1$
509	-94.30	3.558	473.3	-0.165	154.2	-0.116	0.049
511	-81.81	3.138	460.7	-0.191	165.7	-0.135	0.056
510	-100.40	3.991	710.2	-0.319	183.7	-0.210	0.109
183	-169.49	3.452	375.6	-0.237	144.1	-0.104	0.133
188	-93.12	2.490	710.2	-0.319	293.2	-0.215	0.104
533	-413.36	11.288	1144.8	-0.148	133.1	-0.086	0.062
519	37.55	2.472	124.5	0.117	37.1	0.139	0.022
522	-51.56	1.727	136.4	-0.076	107.8	-0.049	0.027
534	-12.16	0.296	14.3	-0.060	88.9	-0.028	0.032
517	-126.50	2.382	258.4	-0.246	146.3	-0.101	0.145
550	-245.21	3.863	710.2	-0.319	213.5	-0.137	0.182
502	-99.39	2.696	343.1	-0.216	153.2	-0.123	0.093
523	-104.32	2.544	189.9	-0.132	111.9	-0.065	0.067
504	-118.26	2.355	272.5	-0.236	151.6	-0.106	0.130
505	-82.26	2.542	233.8	-0.116	121.6	-0.072	0.044
506	-110.33	2.587	276.4	-0.196	140.3	-0.100	0.096
527	-125.64	2.378	262.7	-0.244	148.1	-0.102	0.142
528	-39.89	1.593	295.4	-0.276	195.5	-0.193	0.083
526	-59.60	1.651	151.2	-0.175	121.3	-0.093	0.082
500	-119.15	2.665	300.0	-0.219	145.5	-0.107	0.112

Table C1. Coefficients of equations (1) and (2). Last column is the expected difference in decay parameter k assuming a static relationship between Sediment Yield and Erosivity.

The following graphs show a few curves for Sediment Yield and equivalent Erosivity to make the discussion in the last paragraph more concrete. Station 509 illustrates the most common pattern. The erosivity curve is dropped but declines more slowly to reach the required asymptote of zero. At station 534, erosivity is numerically higher than sediment yield. Its decay coefficient is about half of that for sediment yield, but its curve is actually slightly steeper because its b is larger. At station 519 both curves are rising, and sediment yield is steeper than erosivity because b is larger. These differences in the curves exist despite an assumed constant relation between sediment yield and erosivity.


Station 509
SY = 3.5577E - 94.3
Sediment Yield
Erosivity
E = (SY + 94.3)/3.5577
SY = 473.34exp(-0.165t)
E = 154.18exp(-0.116t)
Sediment and Erosivity
0
100
200
300
400
2004
2006
2008
2010
Year


Station 534
E = 88.86exp(-0.028t)
SY = 0.2959E - 12.159
Sediment Yield
Erosivity
E = (SY + 12.159)/0.2959
SY = 14.335exp(-0.06t)
Sediment and Erosivity
0
20
40
60
80
100
2004
2006
2008
2010
Year



A more sensible approach, along the same lines that the Sullivan et al. (2012) report used, would be to fit model (3), with asymptote, to both erosivity and sediment yield. Then you would have

$$E = b_0 + b_1 e^{-k_1 t} \tag{4}$$

so, assuming there is no trend, the expected sediment yield is:

$$\begin{aligned} S_E &= a_0 + a_1(b_0 + b_1 e^{-k_1 t}) \\ &= a_0 + a_1 b_0 + a_1 b_1 e^{-k_1 t} \end{aligned} \tag{5}$$

Then fit another model of the form (3) to the observed sediment data

$$S_O = c_0 + c_1 e^{-k_2 t} \tag{6}$$

Now compare the derivatives of (5) and (6), which will define two rates of decline that vary with *t*.

$$\frac{dS_E}{dt} = -k_1 a_1 b_1 e^{-k_1 t} \tag{7}$$

$$\frac{dS_O}{dt} = -k_2 c_1 e^{-k_2 t} \tag{8}$$

These can be plotted to see if one is greater than the other for the observed range of t.

While that makes better sense, I still do not recommend it because (1) an exponential decay model is inappropriate for highly stochastic weather data that should be more or less stationary, (2) the model is a very crude fit to the data, and (3) there is no statistical significance test. Representing weather using covariates in a multiple regression or mixed-effects model is a far superior approach.

APPENDIX D. Trend Analysis Using Regression
(Jack Lewis)

I have been able to closely reproduce the regression analyses summarized in Table 10 of Sullivan et al. (2012) using the data printed in their Appendix B. They computed separate analyses for "event" years with large storms exceeding bankfull and "non-event" years. Analyses were also computed for two different response variables: Sediment Yield and Turbidity 10% Exceedence. The authors tested interactions between Year and Watershed (Elk vs Freshwater); a significant test would have indicated different linear trends for the two watersheds. Because multiple tests were done, an adjustment (a la Bonferroni) was made to the p-value to limit the risk of declaring any of the 4 tests significant in the absence of a real effect. Another set of tests was done for interactions between Site and Watershed, with similar results. One of the tests in each group indicated borderline significant interactions ($0.0125 < p < 0.05$), but the possibility of differences between watersheds was not considered further. It is important to keep in mind that failure to find a significant effect means just that; it does not prove the effect does not exist. Real effects are very often dismissed simply because sample sizes were too small to detect them. The authors should have presented scatterplots to back up their statistical analyses, as these can often provide more meaning to the statistics.

I don't feel it was necessary to split the data set into "event" and "non-event" years. Climatic stress is a continuum that produces a continuum of responses. Although such splitting can be rationalized with reference to theoretical thresholds necessary for landsliding etc, I have rarely found the loss of statistical power justifiable. If the data set had not been divided the authors might have found that the interaction was significant, and been motivated to split the data by watershed instead.

10% Exceedence Turbidity

I re-ran the analyses without splitting "event" and "non-event" years. In modeling the 10% exceedence turbidity (T10), I found that the (logarithm of) Erosivity (E) was a much better predictor than Annual Peak (Q), which was used by Sullivan et al. (2012), who defined E as the product of annual rainfall and maximum daily rainfall at the Eureka WSO. Therefore log(T10) was modeled as a function of log(E), Year and Site. I started with a fixed-effects model (i.e. multiple regression), because of the range of diagnostic plots available. The mixed model treats Site as a random effect, but the significance tests for the other effects (such as year) are not very sensitive to that choice. The fixed-effects model in the combined (Freshwater plus Elk) data set indicated a declining trend in T10. An intuitive way to visualize the trend is to drop year from the model, then plot the residuals from the reduced model against year. Statisticians use partial residual and partial regression plots, which have certain advantages, but I will use residual plots in this description because they are easier to understand. The residuals represent the remainder of each observed response that has not been explained by the predictor variables. Figure D1 shows the trend in residuals that has not been explained by log(E) or Site.

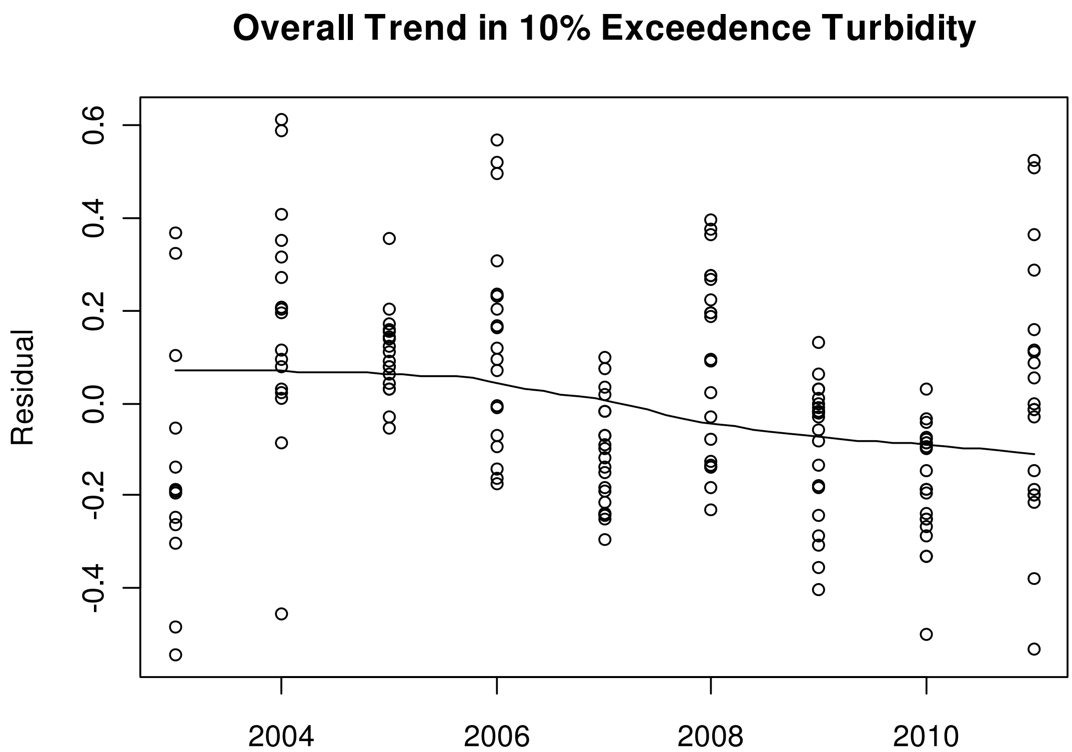


Figure D1. Trend in Turbidity not Explained by Erosivity or Site, complete data set

The random effects model was used to test year and its interaction with watershed. Year was very significant (p=0.0034) and the interaction was borderline significant (p=0.0221). The next step was to fit the same models to each watershed to see if Year was significant in both Freshwater and Elk River.

In Elk River Year was not a significant predictor in either the fixed or mixed model (p = 0.938 and p=0.879). The residuals from the regression without Year do not suggest a recovery trend (Figure D2). The uptick in 2011 parallels that found in my analysis of Salmon Forever's data (Lewis, 2013).

In Freshwater Creek, Year was a very significant predictor in both the fixed and mixed models (p < 0.0001). For these models Q was used in place of log(E) as the weather surrogate, but all models led to the same conclusion. The residuals from the regression of log(E) on Peak Q and Site are shown in Figure D3 and illustrate the declining trend in T10.

It is clear that combining the two data sets (and splitting on "Event" and "Non-Event") hid the fact that Elk River and Freshwater Creek are on two different trajectories.

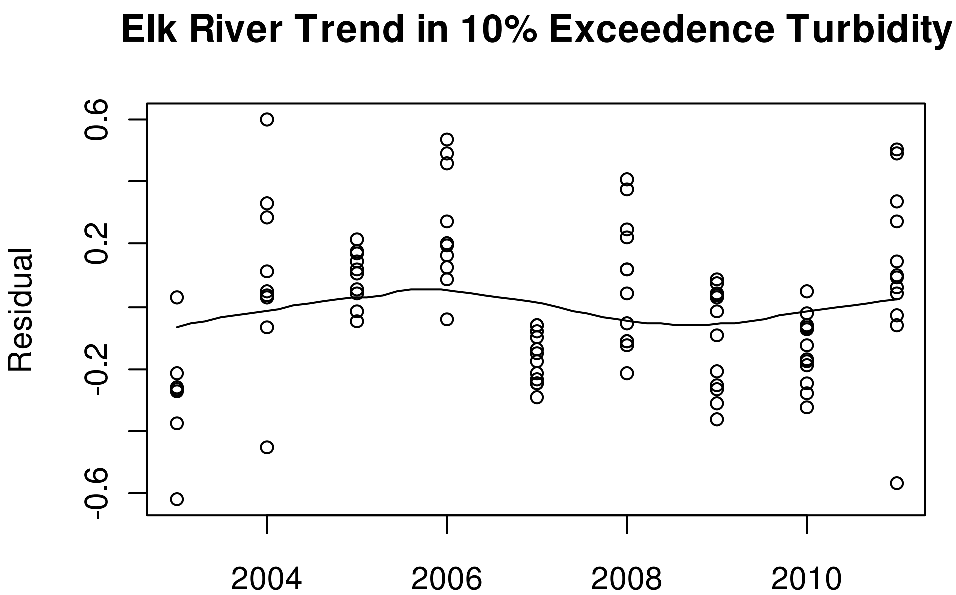


Figure D2. Trend in Turbidity not Explained by Erosivity or Site, Elk River only

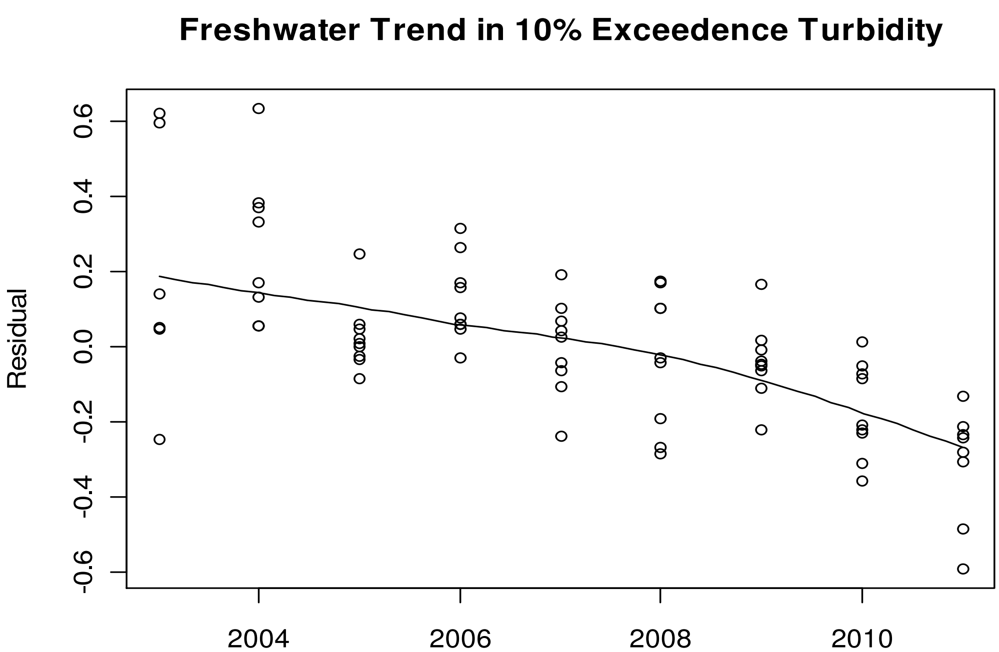


Figure D3. Trend in Turbidity not Explained by Erosivity or Site, Freshwater Creek only

Sediment Yield

I proceeded with the same analyses as for T10. The best weather covariate for the overall model was log(Q). The residuals from the regression model of sediment yield, log(SY) on log(Q) and site, show a declining trend through 2008, then a leveling (Figure D4). Tests for a linear trend do not respect the leveling, but are highly significant ($p < 0.0001$) in both fixed and mixed-effects models.

Tests of the interaction between watershed and year are borderline significant, and depend on the weather covariates included in the model. With just log(Q) in the mixed-effects model, the interaction was not significant ($p=0.094$). But adding annual rainfall to the model, as log(ppt), significantly lowered the unexplained variance and suggested that the interaction was significant ($p=0.023$). Separate investigations of the two watersheds was indicated.

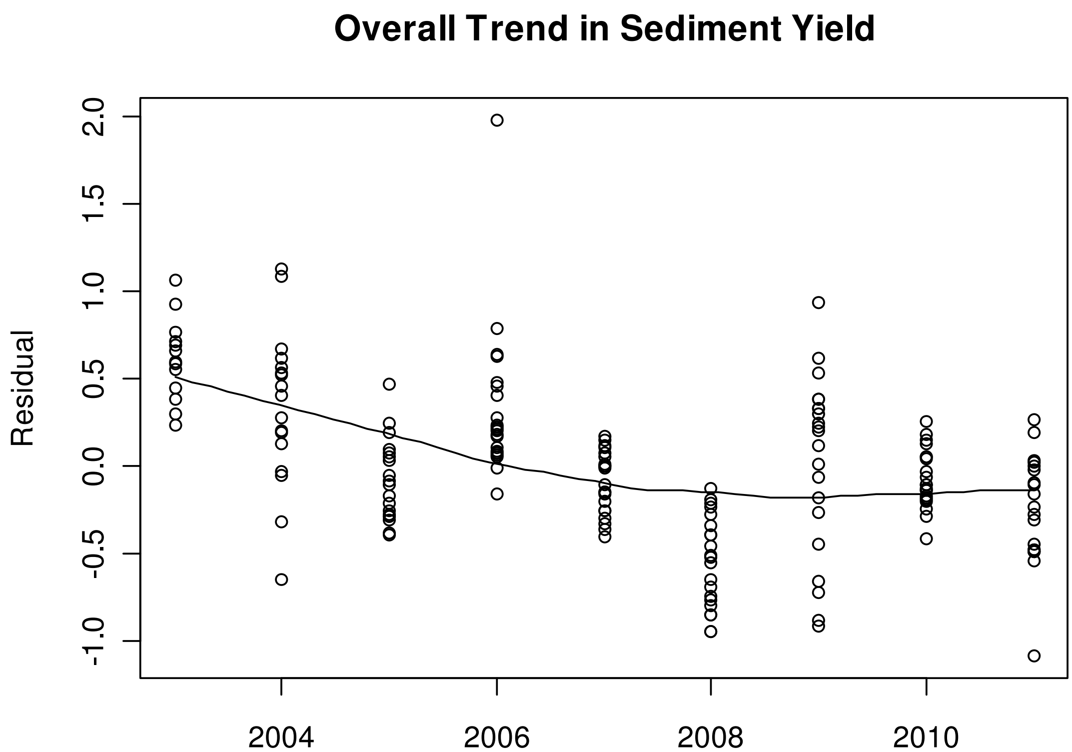


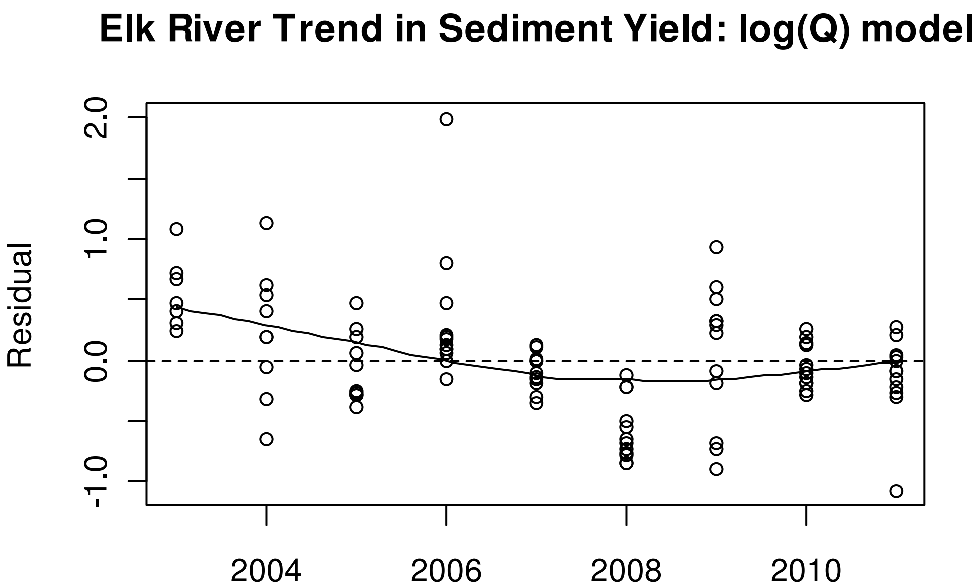
Figure D4. Trend in Sediment Yield not Explained by Erosivity or Site, complete data set

In Elk River, results are ambiguous, depending on which weather covariate is included in the model. The log(Q) model explains more variance but the residuals from the log(E) model are more normally distributed with equal variance throughout the range of the response. For significance testing, the latter should be preferred, but both residual plots are shown below (Figure D5). The trend is significant ($p < 0.0001$) in the mixed or fixed models using log(Q), but not in the models using log(E) ($p = 0.365$ or 0.394). This is a great illustration of why plotting the data is important. Even with a highly significant p-value, how convincing is the trend in the top plot? The trend seems highly influenced by 2003 and appears to have ceased after 2008. In fact if 2003 is omitted, the trend is no longer significant. This suggests to me that log(Q) may not adequately explain the variation due to weather. In modeling the Salmon-Forever data, I found that including both annual flow volume and annual peak flow significantly improved the models. Annual flow volume was not available to me for this analysis, but annual rainfall does improve the fit and renders the trend back to the realm of borderline significance. With regard to sediment yield reductions in Elk River these data are inconclusive; I do not find the case for a recovery trend to be compelling.

In Freshwater Creek, Year was a very significant predictor in both the fixed and mixed models ($p < 0.0001$). For these models, Q was used as the weather surrogate, but all models led to the same conclusion. The residuals from the regression of log(E) on Peak Q and Site are shown in Figure D6 and illustrate the declining trend in sediment yield.

As with turbidity, sediment yields at Freshwater Creek are declining but apparently they are not at Elk River. Combining the data sets hides the distinction.

I have taken a very detailed look at Salmon Forever's sediment concentrations and storm event loads (Lewis, 2013) and there is no evidence of a downtrend since 2008 in either the North or South Fork of Elk River. The Salmon-Forever analyses utilize far more data points than Sullivan's data set of yearly values. I found statistically significant increases between 2008 and 2013 in both the South Fork (SSC and loads) and the North Fork (loads only). Concentrations in the South Fork in 2013 were about 35% higher than the 2003-2013 average after accounting for discharge and rainfall.



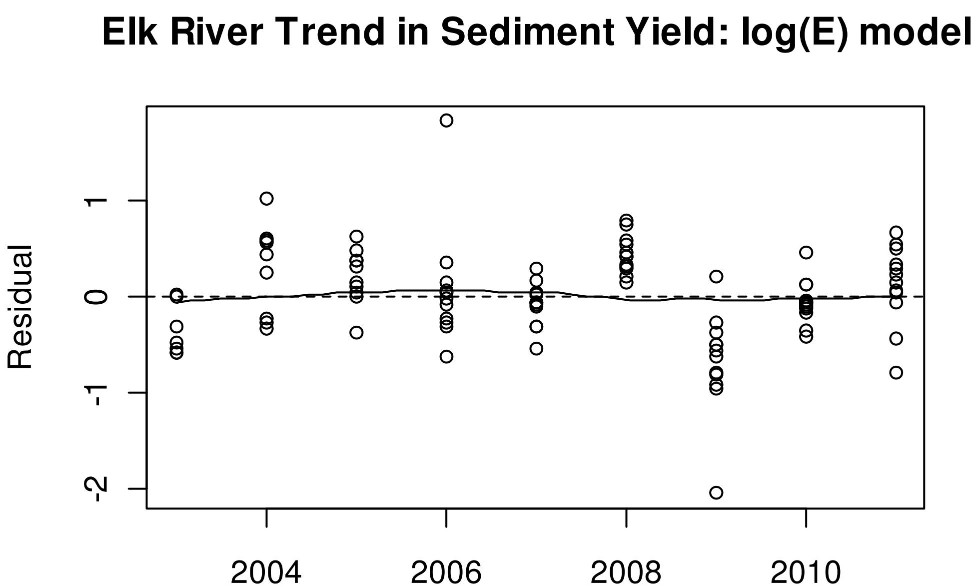


Figure D5. Trend in Sediment Yield not Explained by Erosivity or Site, Elk River only; apparent trend depends on covariate used to account for weather: (1) discharge, (2) erosivity

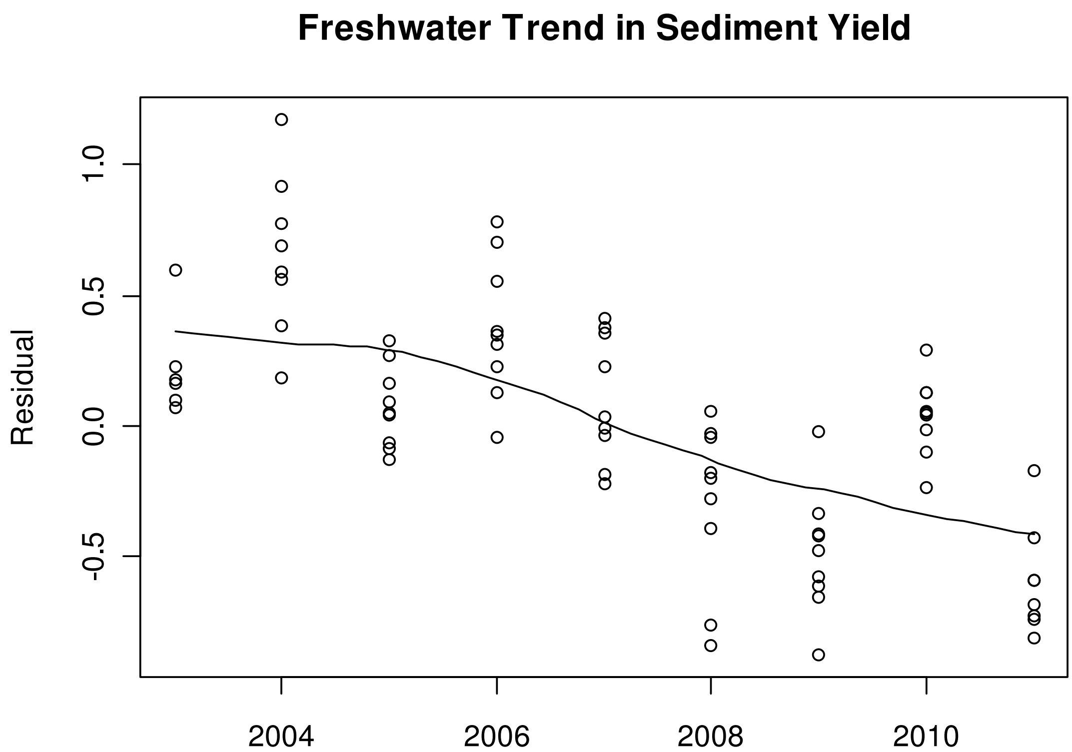


Figure D6. Trend in Sediment Yield not Explained by Erosivity or Site, Freshwater only

REFERENCES: APPENDIX B, C, and D

See REFERENCES for main body of letter on pages 10-11.