Comments on Lee MacDonald's Presentation at the Elk River TMDL Information Workshop, May 7, 2014

May 29, 2014

Prepared by:
Jack Lewis
(USFS Redwood Sciences Laboratory – retired)

I was in attendance at the May 7 TMDL Workshop in Fortuna and was startled by the lack of objectivity in the presentation by Lee MacDonald (LM), who was hired by Humboldt Redwoods Company to speak at the meeting. I write this letter as a concerned scientist. I am not being paid to state a viewpoint by either the residents or industry. I will disclose however that I am motivated in part to defend and clarify the relevance of some of my work. For the past 30 years I have been engaged in research on the relationships between forest management, hydrology, and sediment production, mostly in the redwood region of California. I am very familiar with the Elk River data collected by both Salmon-Forever and Humboldt Redwoods Company. I am author or co-author or some of the studies cited by LM in his talk. As a statistical hydrologist I am keenly aware of the importance of uncertainty in data analysis. On May 7, LM understated the uncertainties in studies he cited to support his view and overstated the uncertainties in (or ignored) the evidence that didn't support his position, which I understand to be that careful timber harvesting at unlimited rates can do no harm nor delay the recovery of Elk River from historic damage. He repeated fallacies and misdirected criticisms of the regional turbidity study (Klein et al. 2012) that we have addressed in previous correspondence with this Board, and he offered a few new objections that demand a response. Unfortunately there was no opportunity for questions and answers after the presentation; however the problems were so numerous that the best way to effectively address them is through a point-by-point rebuttal, which follows. These comments are more in response to what LM said at the meeting than to his slides. His words seemed to betray a more slanted perspective than the slides, which were difficult to see for many in the audience. I have not consistently cited references or included a references section since this is just a letter, however I would be happy to supply literature references by email for any interested parties.

- *Peak flows vary a lot (so a little change doesn't matter).* A graph from a publication was presented showing peak flow changes after logging at Caspar Creek. LM pointed out that the variability was very great; in fact larger than the changes. His message was that "hydrology is messy" and the changes were therefore unimportant. Their statistical significance was not mentioned or challenged. However, a change that is small relative to overall variability can be very significant, both statistically and in relation to ecosystems and people. One only need look at global temperatures to realize that small changes in the mean, relative to overall spatial and temporal variability, can be extremely important to human endeavors. Small percentage changes in large peak flows may be very large in absolute terms and are usually accompanied by changes in frequency for a given magnitude that don't seem so inconsequential.

- *There has to be a proven mechanism for harvesting to affect turbidity before we need to pay attention to it.* LM implied, as he has done in the past, that there is inadequate evidence to link current logging practices with peak flows and turbidity or sediment production. It is true that canopy removal from a small portion of a watershed probably would not have a measurable affect on peak flows at the bottom. However, currently logging is not lightly dispersed across the Elk River landscape. HRC practices typically remove about 50% of a canopy in compact harvest units that have already been logged multiple times. Such harvesting certainly has the potential to affect slope stability. When headwaters harvest units occupy a large portion of small watersheds, the result can also have important local effects on peak flows. It is those local changes that cause increased scour and bank erosion. The number and length of affected channels have been augmented by historic logging practices. Skid trails and ongoing expansion of headcuts are not required for these existing sources to be important.

  What is the mechanism for increased peak flows? It is primarily reduced transpiration early in the fall and reduced interception during winter storms. Studies of interception under evergreen forest canopy repeatedly show that 22-45% of rainfall never reaches the forest floor. The numbers reported for second-growth redwood forest are, surprisingly, at the low end of that range (22-23%). The Caspar Creek interception rates may well be underestimated because of the measurement methodology, which ignored increases in raindrop velocity, wind, and splash that likely cause differential bias between platform collection under the canopy and in the open. Rainfall interception of 23% means that in a clearcut, effective rainfall reaching the ground is 100/(100-23) = 130% of what is normal under a forest canopy. HRC's own data from Doe Creek indicate that this number can reach 140% during the winter relative to a 60-year old canopy. This ignores the effects of transpiration, which in the redwood region is an important process year-round. The combined effects of lost interception and transpiration mean that up to 50% more water may percolate through the soil in a cleared forest than under a full canopy. That number is not surprising when one considers that runoff measured in second-growth redwood forest (Caspar Creek) averages about 50% of annual rainfall, and the primary losses are by evapotranspiration, which includes interception. Recent studies from around the world, including Caspar Creek, have much better time resolution than historic studies, and they universally indicate that evaporation continues in a wet canopy under most atmospheric conditions during storm events. Multiple studies indicate that interception is proportional to rainfall intensity; after the canopy wets up, an approximately constant proportion of the incoming rainfall is intercepted by the canopy and evaporates before reaching the ground (21% at Caspar Creek). When the canopy is removed, that water goes into the soils and the streams. In our climate, soils in canopy openings are thus primed all winter long (unless there are prolonged periods without rainfall), so the next big event is more likely to trigger landslides. If soils are saturated, winter streamflows are enhanced at least in proportion to effective rainfall (27%, based on the 21% evaporation rate). The percentages are higher than 27% after prolonged periods without rainfall when transpirational differences between harvested and unharvested areas become more important.

- *Logging only affects small peak flows*. LM stated that harvesting affects peak flow responses to logging mainly in small events, primarily at the beginning of the rainy season, and that changes to large peak flows are unimportant. Nearly all studies have shown that percent changes are indeed greatest for small events. However, based on what is currently known about canopy interception, large flows must increase as well, and not trivially. Studies after logging the North Fork of Caspar Creek showed increases up to about a 2-year or 5-year recurrence interval. However peak flow studies have not been consistent in their results for the largest events, for which they typically do not show significant changes. Reasons for poor detection are (1) the largest flows are rare and unlikely to provide statistically useful sample sizes before substantial recovery has occurred, (2) except when gaged using weirs or flumes, measurements are usually very poor at flood flows, and (3) many analyses have assumed linear relationships before and after logging, which forces a convergence of effects when larger events have smaller percentage changes. While percentage changes may be smaller for larger events, there is no doubt that the absolute changes can be very large because of the proportionality of interception with rainfall.
- *Your model doesn't explain every data point.* LM showed a figure from the Klein et al. (2012) publication in Geomophology, pointing out that there was an overlap in chronic turbidity between high and low harvest sites, implying that either there was no real difference or those non-overlapping points could probably be explained by factors other than harvest rates. He failed to point out that there was NO overlap at all between chronic turbidity for high harvest rates and pristine sites. However the fact that there is some overlap between turbidities in different classifications of harvest rates in no way implies anything about the statistical significance or cause of those differences. Harvest rates are continuous variables appropriately analyzed by regression; they were shown in the paper to be the most statistically significant predictors of chronic turbidity in 27 northern California watersheds. If the continuous variables are related, the significance of the turbidity difference between any two classes of harvest rates is simply a function of sample size (i.e. if enough sites are measured you will be guaranteed to detect the difference). As for the causes of those differences, we cannot be certain. There are certainly more than one, but the cause for which we have the best empirical evidence of a relationship to chronic turbidity is rate of harvest.
- *You should have invented a variable that describes everything about how management can affect turbidity.* LM repeated his complaint from previous presentations and letters to the Board that the factors used by Klein et al (2012) to calculate clearcut equivalent area don't explain infiltration, surface erosion from road surfaces, and landslide rates. Randy Klein and I responded to this objection in our previous letter to the Board. We had variables that described canopy removal, others that described roads, and others that described landslide potential. To fulfill their roles as explanatory variables in regression, different types of information must be coded in different variables. It just happens that none of the other variables could compete with the rates of canopy removal (for the 10-15 or 0-15 year periods prior to measurement) in explaining the variation in chronic turbidity. That doesn't mean none of them have anything to do with turbidity. This size of data set simply doesn't contain enough information to support their inclusion in the model.

- *Why didn't you use the variable these other guys came up with*? LM referred to a paper by Andrews and Antweiler (2012), in which bedrock geology had been related to sediment loads, and suggested that if those authors could do that why didn't Klein et al (2012)? Andrews and Antweiler characterized erodibility using "the relative mechanical strength of a fresh unweathered sample" determined by hitting it with a hammer. Based on this observation they assigned a number to each mapped lithology in the basin and weighted each lithology by the area of exposed bedrock of that type. Bedrock exposures are not common in the coastal watersheds of California; where they do occur they tend to be the most resistant rock types, so they don't provide an indication of the overall proportions of lithologies in a basin. It is unclear how this information about the abundance of atypically hard rocks could be relevant to basinwide erodibility unless it were inversely related to erosion (which it was not). The regression model in which this variable was "significant" incorporated 4 variables to explain the variability of 16 watersheds. Even a two-variable model would likely be overfitted to such a small data set so their model is very likely just fitting noise in the data. I don't doubt that basin geology is related to erosion rates, but I'm not sure this constitutes the proof. A numeric characterization of basinwide bedrock erodibility is not something easily obtained from a geologic map and a few rock specimens. However, we did consider landslide potential, basin relief and mean basin slope in our analysis (Klein et al 2012).
- *What about using uplift rates to represent basin erodibility*? While LM didn't explicitly ask that question, one of his slides suggested that latitudinal distance to the Mendocino Triple Junction might be a good indicator for uplift rates and basin erodibility. So I decided to try including this variable in the Klein et al (2012) regional analysis of chronic turbidity (10% exceedence level) even though the model already had two explanatory variables. I computed the north-south distance from the Mendocino Triple Junction (MTJ) and added it to the published regression model. The new variable was weakly correlated with the 10-15 year harvest rate (-0.29). (A high correlation would have been worrisome because it might have meant that harvest rate was in the model spuriously because of its relationship to uplift rates.) Including the MTJ distance in the model increased the adjusted R^2 of the regression model from 0.632 to 0.765 (a very big improvement) and increased the significance of the harvest rate variable from $p=0.00017$ to $p=0.00009$, without changing the sign of its coefficient. So it seems that accounting for erodibility in this way only strengthens the case for harvest rate being an important influence on chronic turbidity.
- *You told us there was one place that was improving (therefore your model must be wrong).* LM said that the apparently declining sediment loads that I have identified in Freshwater Creek disproves the linkage between harvest rate and turbidity reported by Klein et al (2012). The argument has no substance. LM did not show what our model predicts for Freshwater Creek. In any case, a prediction error for a single observation would prove very little. And harvest rate effects are often delayed until large storms occur and/or root strength declines, so the current trend could easily reverse itself. There are always unknown variables and processes behind the statistically unexplained variance, and prediction is never perfect. We would have liked to include more variables in our model, but this would have required data from many more gaging sites.

- *Your reference watershed is not identical to the rest of Elk River*. LM displayed a figure showing topographic roughness from the DSLED-ROUGH model (McKean and Roering 2004). He stated that this indicated deep-seated landsliding was present in much of Elk River but not in the reference basin (Little South Fork) and implied that therefore the reference basin must have naturally lower erosion rates than the rest of the watershed. These maps need to be interpreted cautiously. Many of the features identified are thousands of years old and show no evidence of late-Holocene movement (Mackey and Roering 2006). There has been no ground-truthing of identified features in Elk River, and no evidence was presented that would link roughness to current sediment loads. This landslide mapping tends to over-predict the extent of landslide terrain (Roering et al 2006) and "does not directly address potential sediment delivery from landslide-prone areas to a watercourse and/or other important receptors" (Stillwater Sciences 2007).

  There are no two identical watersheds anywhere on earth, so the best one can hope for in selecting study watersheds is similarity of geology, soils, topography, climate, and historic management. These characteristics are as uniform in Elk River as in typical experimental watersheds such as Caspar Creek and HJ Andrews. However, "reference" watersheds, if used properly, need not be identical to the watersheds being calibrated to them (see discussion at the end of this letter).
- *There are very high percentage errors in the measured sediment yields from the reference watershed*. LM showed four scatterplots of suspended sediment concentration versus turbidity in Little South Fork Elk River (ESL) to support an argument that data quality is poor and the reported sediment loads are unreliable. The relationships for years other than 2004 were very weak, indicating that sediment loads are highly uncertain. However, this argument is not particularly relevant: the inter-site differences far overshadow the uncertainty at ESL. In the 4 years shown, only two samples out of well over 100 exceeded a concentration of 100 mg/L. In 2004 and 2005 the turbidity data are of high quality (I have not seen the data from 2007 or 2008) but the relation between turbidity and sediment in 2005 is poor simply because no concentrations higher than 10 mg/L were obtained. While the sediment loads in years other than 2004 are not known precisely it is absolutely clear that they must be very small. In 2004, sampling used the TTS method which preferentially samples storm events and high turbidities. Using the same TTS sampling algorithm, 118 of 207 samples at SFM (South Fork main stem) exceeded 100 mg/L, while *none* (of about 50 points) exceeded that value at ESL. The maximum turbidity recorded in 2004 at ESL was 62 NTU on Feb. 16. In contrast, Salmon Forever's SFM station reached over 1500 NTU that day and exceeded 200 NTU during 12 different storm events that year. In 2005, the maximum turbidity at ESL was 125, while SFM reached 1600 NTU and exceeded 200 NTU in 20 different events. Chronic turbidity is less variable than the maxima and yet HRC's data show that on all three of their measures of chronic turbidity, ESL is far and away the cleanest of all monitored streams every single year. That includes several subwatersheds that have similar or higher percentages of Yager formation and one watershed (Bridge Creek) that like ESL is apparently without deep-seated landslides. Differences of such a magnitude cannot be explained by the variation in topography, bedrock, or rainfall within the Elk River. The most plausible explanation for such stark differences is the management history.

Little South Fork obviously produces a very tiny fraction of the sediment of the other gaged subwatersheds and it is the only one that has never been logged.

- *Things used to be a lot worse and the uncertainty in the evidence doesn't merit discussion.* As explained in the next bullet, the long-term erosion rates aren't really relevant to this discussion. The studies by Ferrier (2005) and Balco (2013) use measurements of a beryllium isotope (Be10) in quartz obtained from sediment samples to deduce erosion rates. LM stated that the uncertainty was only 10-20% due to laboratory procedures. He didn't mention that Balco's estimates for North Fork Caspar Creek were 82% higher than Ferrier's estimates *from the same laboratory data* (and 70% higher for the South Fork). These researchers just made different assumptions in computing the rates. The Be10 method of estimating millenial erosion rates requires at least a dozen critical assumptions that are difficult or impossible to verify. I won't list them all here but they include (1) that all surfaces in the watershed contributed sediment to the sample in perfect proportion to their *long-term* erosion rates, (2) that the only erosional process in the watershed is surface erosion, and (3) that quartz is evenly distributed throughout the basin. These and other critical assumptions are probably far from true in mountainous watersheds with high spatial variability and huge changes in management that have taken place since Europeans entered the scene. Current methods of monitoring using turbidity and pumping samplers cannot determine historic loads, but they give far more accurate estimates than indirect methods using cosmogenic radionuclides or rating curve estimates (Andrews and Antweiler) based on historically infrequent manual sampling.
- *Things used to be a lot worse (so we must be doing things right).* LM cited studies suggesting that current sediment yields are underestimates of the elusive "background" condition, supposedly represented by long-term yields over thousands of years. Because yields in the distant past appear to have been much greater than they are today he concluded that the relative impact of current management is therefore overestimated. That sounds a lot like "we don't need to worry about carbon dioxide levels in the atmosphere because they have been this high before (and the dinosaurs loved it)". The unstated corollary is that today's management failures, including chronically elevated turbidity are trivial distractions. Also assumed and unproven is that, when they arrive, the large events that control the long-term averages will produce the same responses regardless of management. However, arguments based on average long-term erosion rates overlook the problem that managed basins are in much poorer condition today than unmanaged basins and intensive harvesting will likely prolong those differences. Regardless of yields in a different era, current inputs do matter to the people affected by flooding and degraded water quality and aquatic/riparian ecosystems.

Elk River has been identified as an impaired watershed; the hardships caused by the impaired conditions were described clearly by the residents at the May 7 meeting. Remediation should be undertaken as soon as possible, and sediment production needs to be reduced. That probably will require controlling cumulative watershed effects though limitations on harvest rates. But it is difficult to specify quantitatively how much reduction should be required. Can we expect disturbed basins to perform the same as a selected reference basin? As we have seen, using a reference watershed as an absolute target condition for a managed watershed is difficult to justify because physiographic differences can always be found between any two basins. Historic management has in many cases caused profound channel changes that can lead to long-term differences. Even undisturbed basins that appear to be quite similar can differ markedly in hydrology and sediment loading. More importantly, in a dynamically variable system it is impractical to define progress in terms of attainment of a single number such as a long-term average erosion rate. The ‘reference condition’ consists not of a single value, but rather a range of turbidity and sediment yield values, owing to natural variability. Erosion rates vary in response to the weather and climate. These are not stationary time series' -- the means drift about and can seem to change abruptly after extreme years

However, progress can be measured in terms of performance for a given hydrologic input, indexed by responses in a control watershed, ideally one that is off limits to disturbance and is not rapidly changing. (Protection should eventually result in stability.) Through calibration and monitoring we can track changes in responses for an event of any given recurrence interval. Using a control watershed in this way, as a relative rather than absolute reference, does not require carbon-copy watersheds. It only requires that (1) the watershed pairs have very similar hydrologic inputs, (2) they tend to respond proportionally for a given hydrologic input and (3) the control watershed be protected and stable. (Since access to the Little South Fork is so difficult, Upper Salmon Creek in the Headwaters Forest should also be given serious consideration for this purpose, and two controls are better than one.) Responses in the reference watershed(s) can provide a rough guide for goal-setting, but a target such as 120% of background may be impossible to match numerically in a managed watershed. It may be more reasonable to set objectives and measure progress in terms of percent change from current conditions for a given size event.

Finally, selective harvesting should have smaller effects than the clearcut logging that was studied in Caspar Creek, but whether or not those effects are negligible and can permit recovery of the degraded ecosystem has yet to be proven. It should be industry's responsibility to demonstrate that recovery is taking place, not the burden of the already stricken residents to prove there is continuing harm being done.

Sincerely,

Jack Lewis

Jack Lewis