# Peer review of "FINAL REPORT – INSTREAM FLOW ASSESSMENT PILOT PROJECT"

**(draft of 9/9/05)**


**Derek B. Booth, PhD, PE, PG**
**Research Professor, University of Washington**

**Paul McElhany, PhD**
**Ecologist, NOAA Northwest Fisheries Science Center**

**William J. Trush, PhD**
**Senior Ecologist, McBain and Trush, Arcata, CA**


*Final Report—November 19, 2005*

## SUMMARY


The *Final Report – Instream Flow Assessment Pilot Project* (peer-review draft of September 9, 2005) had an overly ambitious scope executed within a limited timeframe, and consequently it attempts to accomplish too much with insufficient data and resources. The bulk of the report's presentation are dedicated to results of applying the Ecosystem Diagnosis and Treatment model (EDT), but the underpinnings of this model are obscure and possibly flawed, model limitations that overshadow other potentially useful contributions of the study.

The Draft Report made a noble effort to apply EDT to two pilot watersheds. In so doing, the Draft Report accomplished the important goal of revealing that this approach should *not* be used as the template for region-wide planning. The report demonstrates the pitfalls (and perhaps the unrealistic expectation) of using data-intensive, computationally complex models of any flavor for developing either regional or basin-specific management recommendations. The Draft Report also demonstrates, perhaps inadvertently, the critical necessity of relying on thoroughly peer reviewed models. EDT has many innovative attributes but also many questionable assumptions, of which almost none have been subjected to careful, unbiased scrutiny. Until that has occurred, it will be easy (and appropriate) to cast doubt on any management actions that have been guided by EDT results.

The greatest value of the Draft Report is to remind the reader that urbanization is damaging salmon populations, and that the integrated effects of both runoff alterations and water withdrawals produce recognizable hydrologic effects that have direct correlation with biological conditions. Translation of these effects into predicted numbers of fish is completely unnecessary for the region to decide whether, and how, to increase its efforts to reduce or reverse the consequences of urban development on the biological health of lowland rivers and streams. A number of alternative approaches already exist, or are well within reach, that could be used to improve the effectiveness of such efforts over the near term.

# REVIEW OF THE 9/9/05 DRAFT REPORT

## 1. ECOSYSTEM DIAGNOSIS AND TREATMENT MODEL (EDT)

### A. EDT and "Best Available Science"

A complete peer review of the EDT model would be essential to evaluate the quality of this study. We strongly recommend against further application of this model until such a peer review has been completed. EDT may or may not provide useful, credible outputs, but it does not pass even the lowest standards for a scientific peer-reviewed framework, and it does not meet the standards for "Best Available Science." This is an important criterion if its results are ever to be used to guide management actions in the State of Washington.

By the Washington Administrative Code, "Best Available Science" is expected to have the following characteristics in order to be considered a "valid scientific process" (WAC 365-195-905):

"1. Peer review. The information has been critically reviewed by other persons who are qualified scientific experts in that scientific discipline. The criticism of the peer reviewers has been addressed by the proponents of the information. Publication in a refereed scientific journal usually indicates that the information has been appropriately peer-reviewed.

"2. Methods. The methods that were used to obtain the information are clearly stated and able to be replicated. The methods are standardized in the pertinent scientific discipline or, if not, the methods have been appropriately peer-reviewed to assure their reliability and validity.

"3. Logical conclusions and reasonable inferences. The conclusions presented are based on reasonable assumptions supported by other studies and consistent with the general theory underlying the assumptions. The conclusions are logically and reasonably derived from the assumptions and supported by the data presented. Any gaps in information and inconsistencies with other pertinent scientific information are adequately explained.

"4. Quantitative analysis. The data have been analyzed using appropriate statistical or quantitative methods.

"5. Context. The information is placed in proper context. The assumptions, analytical techniques, data, and conclusions are appropriately framed with respect to the prevailing body of pertinent scientific knowledge.

"6. References. The assumptions, analytical techniques, and conclusions are well referenced with citations to relevant, credible literature and other pertinent existing information."

Although many of these characteristics are weak or altogether lacking in EDT, prominent among the model's shortcomings relative to this standard is the almost complete absence of peer review, a consequence of the proprietary nature of the model. Bisbal (2002) notes: "In academic circles, it is anticipated that 'good science' is formally published in professional periodicals and has satisfied the scrutiny and critical review by peers in that scientific discipline. Proponents of peer review argue that this criterion is an essential feature of science and that it provides for the objective validation and assessment of reliability of scientific information." EDT does rely on many published, peer-reviewed relationships between the physical environment and salmon. For example, the functional relationship between percent fines in spawning gravel and egg mortality has been peer reviewed. What has not been peer reviewed in EDT, however, is how these numerous functions are assembled in EDT and how they interact to predict salmon and steelhead population sizes. We strongly recommend against further application of EDT until such a multidisciplinary peer review has been completed for the entire model, the results shared with the scientific community, and identified problems have been addressed. There is a growing enthusiasm in the region to initiate such a review, but its completion (and its outcome) is too speculative to plan for.

Although the Draft Report notes that EDT use is "common," this fact should not be considered an indication of whether the model has been adequately evaluated. The primary validation tests that have been applied to the model in the past seem to be whether or not the results are in agreement with previous expectations. While agreement with preconceived notions is comforting, it is not a scientific validation and begs the question of why use such a complex model if conclusions could have already been made by other means.

A complete peer review of EDT is outside the scope of our review, just as it was outside the scope of the Draft Report. This remains a critical need, however, before any use could be responsibly made of the "predictions" of EDT. For a future such review, peer reviewers of the EDT model would require access to the following information:

- Complete documentation of the EDT model;
- Any peer reviewed papers on the model;
- Several written example applications of the model;
- Any written studies validating the model (tests of accuracy);
- Any written studies on model parameter sensitivity; and
- Any previous reviews (or partial reviews) of the model

The EDT model and EDT code are both proprietary, which has contributed to the lack of access to the inner workings of the model. We encourage the consideration of an OpenSource framework for model development (http://www.opensource.org/docs/definition.php). In OpenSource computer programs (and, by extension, model development), all source code is publicly available. The entire community of researchers can access the model and contribute to its improvement or the development of specialty modifications. The result is a completely transparent process and a product(s) that represents the combined thinking of many researchers. The process has limitations, but it has proven very successful in developing complex computer programs (e.g., the Linux operating system). Openness and transparency are particularly critical when considering the management of a public resource.

## B. False precision

The EDT model output provides point estimate predictions of population performance down to the nearest fish. This leads to a false sense of model precision and implied accuracy (i.e. "pseudo-precision"). The model has, literally, tens of thousands of parameters, all of which must be estimated and often with great uncertainty. This uncertainty in the model inputs must be propagated up into the model outputs to be of use for management decisions. For example, the base-line abundance prediction for adult Chinook salmon in Pilchuck Creek is 6,559 fish; but is that ± 10 fish or ± 6,000 fish? In evaluating the consequences of input uncertainty, it is necessary to consider all the model parameters, not just those readily assessable to end-users. Users of EDT generally have access to the habitat parameters (and in the Draft Report, some "rule" formulas) but not to the many other parameters internal to the model.

The phrase "In the context of assumptions made for the purpose of this project," used in many places in the Draft Report to qualify the salmon population estimates, has immense implications. For many policymakers and resource managers/users this phrase would be dropped, resulting in the conclusion that "Pilchuck Creek Chinook salmon are, and will be, unaffected by human development in the watershed." While EDT authors note that "EDT is not designed to prove anything but rather to apply existing knowledge to important watershed problems" (Lestelle et al. 2004, Section 6, p. 26), model-derived estimates are presented much more assertively. If a policymaker asked how good the science is, Lestelle et al. (2004) assert "Almost all rules in EDT are currently rated a 1 or 2 proof level." Proof Level 1 requires "thoroughly established, generally accepted, good peer-reviewed empirical evidence in its favor," while Proof Level 2 requires "strong weight of evidence in support but not fully conclusive." A policymaker or planner couldn't ask for more. The review team, however, cannot find any basis to agree with either statement.

Although the Draft Report frequently qualifies and reminds its readers that the predicted population numbers have more significance relatively than absolutely, population estimates are presented to the nearest fish. This is a high level of precision that gives the appearance of an equally high level of accuracy (and thus correctness). It will likely encourage planners and/or agencies to consider these valid population numbers. Indeed, the reason for applying EDT is that real fish numbers are expected. Otherwise, why not take a much simpler approach, such as using percent impervious area and a few other independent variables to fashion a multivariate model as in May et al. (1997)? The answer seems simple: salmon are important in the Pacific Northwest. Land-use planners and state/local governmental agencies may make painful decisions to protect salmon, but they will do so only for a dependent variable that voters can get behind. So the talk in the introduction about relative numbers, or that the numbers aren't real, is disingenuous. EDT is in the spotlight because it predicts numbers of salmon lost or gained.

The study never addresses whether EDT was an appropriate model for the pilot project application. Answering this question seems like a primary reason for doing the pilot, but the issue was strikingly absent from the Draft Report discussion and conclusions.

### C. Model uncertainty and error

The Draft Report was inconsistent and ultimately misguided in applying the EDT model and presenting its results. It was generally careful to describe limitations and caveats associated analysis, and the examination of the rule parameters showed a willingness to move beyond simply accepting the "black box." However, the Report concludes by saying "Overall, the methodology employed in this study appears to be a sound way of comparing scenarios, getting a sense of their relative rank and at least a semi-quantitative appreciation of the degree to which they differ from one another" (page 5-33). We consider it premature at best and quite possibly mistaken to conclude that EDT soundly compares scenarios. There is no justification for this conclusion.

In addition to considering how EDT input uncertainty affects model output uncertainty, it is necessary to evaluate model selection uncertainty. Even if every parameter in the model were known perfectly, it is still unclear if EDT represents the "best" model structure of the possible models that could be created with the same information. Model selection procedures, such as AIC, exist for comparing different models, but these methods would be difficult or inappropriate to apply to EDT. Currently, very few models attempt the same predictions as EDT with which to make a comparison. However, some principles from formal model-selection analysis could be applied. One principle is parsimony: formal model-selection procedures apply a sort of penalty for an excessive numbers of parameters if those additional parameters do not provide substantially more information. The EDT model tends not to adhere to these principles and hence the EDT model is over-parameterized, and therefore overly complex, for most purposes. In attempting to include every possible functional relationship, the model greatly outstrips the available empirical information for habitat attributes and for functional relationships between habitat and fish performance. The large number of parameters and the complex structure of the model make it very difficult to tease out causal relationships; its data requirements are so onerous (particularly for elements identified as "critical" for the typical land-use scenarios investigated) that EDT's application cannot be defended pragmatically, either.

The Draft Report does a good job describing how the EDT habitat input parameters are generated from the streamflow modeling and other sources (Chapter 4). The Draft Report also provided a good qualitative description of the level of uncertainty associated with the input values, but a quantitative description of uncertainty would have improved the analysis. For example, if bed scour were assigned a value of 2, what is the possible range? Such information could provide a basis for necessary sensitivity analysis.

The ability of the Draft Report to evaluate even the relative effect of the different flow scenarios is questionable, because the relative effect depends in part on the values of the parameters held constant in the analysis. A sensitivity analysis could help understand this effect. Some preliminary sensitivity analyses on EDT suggest that uncertainty in these internal parameters (e.g., "benchmarks") may have a much greater impact on the results than the user-accessible parameters (e.g., habitat attributes).

The Draft Report presented some of the rule curves as "sensitivity" measures. Although looking at the curves can provide some information, the relationship between the rule curves and the EDT output is not simply proportional. In particular, the amount of time a trajectory spends in a particular life stage also affects the output, and the parameters controlling movement in time and space affect any particular outcome.

The Draft Report discusses this issue in the context of the "rules" analysis, and although this is a good start it is still very incomplete. The Draft Report notes that many of the habitat parameters (e.g., direct flow parameters) are likely to have little effect on model output. Work by McElhany and others show that about 20% of the rule functions could be removed from the model with little impact on results. A thorough sensitivity analysis that compared the scenarios under a variety of other constant conditions would be necessary to determine the robustness of the scenario conclusions. As a fairly simple (but still very incomplete) first step, the scenario analysis could be done assuming current conditions instead of pristine conditions (see also below). A systematic assessment would include:

- Conducting a general global sensitivity analysis of the model to understand basic model behavior and sensitivity of model output to input uncertainty.

- Based on sensitivity analysis, greatly reduce the number of model parameters and model complexity so that model behavior and relationships to empirical observations are more transparent.

Although these steps may be outside the ostensible scope of the Draft Report, it is the review team's unequivocal judgment that they would be necessary before *any* use of EDT model outputs in real management actions occurred.

**D. Assumptions in the modeling**

For purposes of evaluating the effects of flow alterations on fish populations, the Draft Report recognized the need to minimize the modeled variability in other non-flow attributes of the study basins. In the judgment of the review team, however, its approach was critically flawed by choosing the assumption that habitat composition, woody debris, riparian, and estuary conditions mimic historic or pristine conditions (e.g., Draft Report p. 5-19, and repeated frequently throughout the document).

The review team struggled with this basic assumption in the Draft Report for several reasons. First (and the Draft Report recognizes this), physical habitat responses can't really be separated from flow responses. This assumption immediately puts the Draft Report at a disadvantage and on the defensive to justified criticism. The attempt to separate the effect of flow changes within Church and Pilchuck creeks from other environmental changes, within and outside watershed boundaries, seems a strange way to "test drive" the EDT model as a land-use planning tool. Would EDT predict 100 adults if the assumption was removed? This would have been a fairer test of EDT and far better opportunity to evaluate the wisdom of any potential future use of the model.

A more realistic, and arguably more useful, question would be to use EDT to evaluate the consequences of future flow changes given the *present* state of both watersheds and the basin. For example, an X % further reduction in present baseflows (presumably due to urbanization in Church Creek) will cause an X % further reduction in adult salmon. Because further reduction in baseflow is unlikely to initiate additional changes in physical habitat structure (though reductions should alter habitat availability), the present condition of both streams can be considered the likely physical condition in the near future.

Detailed fish counts display some of the problems caused by the Draft Report's "pristine" assumption. The highest annual count of adult spawning Chinook salmon in the 9.4-mile-long Pilchuck mainstem was 53 salmon in 1969 (Draft Report, p. 2-10). The quoted paragraph above refers the reader to Table 5-9 (p. 5-19), where EDT estimates an historic potential capacity of 6,911 adults and historic potential abundance of 6,520 adult salmon. Future Scenario 1 has EDT forecasting a capacity of 6,801 adults and an abundance of 6,414 adults. Are we to conclude that the huge discrepancy between observed population size (roughly 100 adults) and predicted current population capacity (roughly 6500 adults) is essentially solely attributable to (1) factors outside Pilchuck's watershed boundary and/or (2) changes in habitat composition, wood, and riparian integrity within the boundary? If EDT did not operate under its driving assumption of habitat composition, woody debris, riparian, and estuary conditions mimic historic or pristine conditions, would EDT predict 100 adults?

The Draft Report results lead to the conclusion that the Pilchuck salmon population is grossly limited by the number of eggs deposited. Smolt production should be relatively huge (relative to the number of eggs deposited) if only (1) above is happening with almost no smolts or pre-smolts survive swimming to the Pacific Ocean, never reach maturity in the ocean, or cannot return as adults. If only (2) above is happening then the habitat condition, wood, and riparian integrity must be hugely compromised relative to pristine conditions, resulting in very few smolts/pre-smolts leaving Pilchuck Creek regardless of the eggs deposited.

The Draft Report recognizes that the magnitude of the fish response under the pristine-habitat assumption would be incorrect (i.e. too high), but it asserts that the relative effect when comparing across management scenarios should be informative. This is a simplistic assumption, however, and may not be correct. There are potentially strong interactions between the parameters in EDT such that changing the parameters held constant in the model to different constant values could alter the relative effect predicted for the different scenarios. For example, the assumption of historical estuary conditions will lead to a larger number of trajectories surviving, relative to settings with current estuary conditions. This change in the number of surviving trajectories can change the relative usage of different reaches, which can affect relative predicted flow impacts under different scenarios. We have no way of knowing if this would affect the results because comparisons were not made, but changing any of the thousands of constant parameters in model has the potential to alter the relative flow effect.

Holding non-flow conditions constant is itself, of course, an artificial assumption to try to isolate flow effects. In reality, the changes associated with flow dynamics as a result of urbanization, for example, will often be associated with changes in other habitat features such as riparian condition or permanence of instream large woody debris. The report clearly stated the limited scope of habitat features evaluated, but it is worth further emphasis to managers—this is in no way a prediction of the relative consequences of urbanization or of other large-scale landscape change. As such, the ultimate utility of the entire modeling exercise would be limited even without our fundamental concerns about the model as a whole.

Curiously, this limitation is recognized elsewhere in the Draft Report, but that understanding is not mirrored in the EDT section itself. For example, the discussion of the NFFS Pathways and Indicators matrix properly notes (p. 5-39) that "habitat effects that might result from a deficit in late seral-old growth as reflected in the Matrix evaluation would not show up in EDT because of the initial and constant conditions for riparian and instream habitat." This recognition is never echoed, however, in the discussion of the EDT results themselves.

## E. Poorly parameterized variables

When data are lacking, EDT modeling left out that variable or substituted questionable surrogates. Furthermore, the fact that data triage was necessary was left out of the conclusions. For example, three fine sediment parameters are modeled as "Level 2" attributes in EDT: (1) fine sediment in spawning gravel ('FnSedi'), (2) embeddedness of the channel-bed surface by fine sediment ('Emb'), and (3) turbidity or suspended sediment ('Turb"). All three are strongly responsible for cumulative watershed impacts in response to gradually increasing land use (in both extent and intensity). While comments would be equally appropriate on any of these attributes, we focus only on Turb as an example of concerns because the suspended sediment responsible for turbidity is especially responsive to even small changes in land use.

Newcombe and Jensen (2000; referenced in the Draft Report) are incorrectly cited in Lestelle et al. (2000) as calculating the “severity of ill effects” (SEV) to fish based on the number of hours in a month that suspended sediment concentration exceeds some threshold value. Rather, SEV is calculated from the number of *continuous* hours that the threshold suspended sediment concentration is exceeded. This makes a very big difference in the SEV outcome: the EDT analysis will overestimate SEV and therefore overestimate turbidity effects on salmon populations using Figure 5-9 (Draft Report, p. 5-32). Any thorough peer review of EDT likely would have recognized this error already. Unfortunately, Figure 5-9 also does not account for cumulative non-lethal effects on juvenile salmonids. Instead, the rules curve does not begin to seriously affect juvenile survival until SEV values directly kill juveniles, scaled as approximately 2.5 in EDT. We know this is not correct, based on numerous studies of deleterious, but non-lethal, effects of fine sediment on fish. From this one example, EDT seems poorly equipped for handling cumulative watershed effects from fine sediment. As a consequence, the “assessment” of industrial forestry in Pilchuck watershed lacks a comprehensive review of likely impacts, of which the cumulative effects of fine sediment is an obvious one. The gross simplification of likely effects severely narrowed the scope of the analysis, and the lack of sediment data forced speculation for several variables to such an extent that those variables are not really characterized or evaluated at all.

As another example, redd scour is considered a prime culprit for lower Chinook capacity (due to higher peak flows), and yet the pathway leading to this assertion is extremely loose. EDT is apparently very sensitive to the Bed Scour parameter, and yet Bed Scour is rarely measured and is difficult to measure in such a way as to arrive at the answer needed for EDT attribute rating (p. 5-39). Section 3.5.6 reduces the complex phenomenon of sediment scour to an unreviewed and dubious hydrologic ratio. Continued application of EDT would need some process for dealing with this, insofar as many (most) watersheds will have even poorer data than are available here.

Given the opaqueness of EDT and its apparently rudimentary hydrology, the approach and justification articulated in Section 2.2.2 (“Flow-related attributes in EDT model”) should be critical. Yet this section has no reference to published literature, only to a draft internal county report. This is not scientifically defensible. Of the 6 EDT hydrology variables, two are properly eliminated for this study as being irrelevant to the scenarios being modeled. Of the four remaining, one (FlowDielVar) is replaced by a land-cover variable. What do prior studies say about the relevance and comprehensiveness of the remaining three variables? Are they the “right” three to capture the effects of human activity on flow regime? Might any significant factors be missing in this set?

Overall, a critical introductory section to the presentation of results (Section 3.5, “Scenario simulations and results”) is missing. It should cover the following: (1) what hydrologic parameters does the current scientific literature say are important? (2) What parameters will EDT need in order to run? (3) How closely do the lists in #1 and #2 align with each other, and what does that judgment tell us about EDT’s utility and defensibility? This third point (integration and evaluation) can be saved for later in the document, but astute readers will do the comparison even if it is not provided here. This section is too focused on getting to an easy-to-calculate set of parameters, and in so doing it only convinces the reader that the scientific underpinnings are weak and the results consequently unreliable.

Other sections of the text display this theme. Section 4.3.2.1, attempting to define a B-IBI and EDT "linkage," has no substantive basis. Pages 4-16 to 4-17, in particular, are very mechanistic and simply wrong, with the simplistic relationship presented there being contradicted by a number of published studies from both the Puget Lowland (e.g., Morley and Karr, 2002) and elsewhere. Similarly, Sections 4.3.2.2 and 4.3.2.3 (see also page 4-34 and 5-29) present simple single-valued TIA surrogate relationships for critical EDT parameters. They raise the question of why use a complex model at all.

**F. Accuracy and validation of model results**

The Draft Report concludes (Section 6.2.1 EDT Results, p. 6-2) that "In the context of assumptions made for the purpose of this project, Pilchuck Creek Chinook salmon were unaffected by the current level (or essentially any of the scenarios modeled) of human development in the watershed. Forest practices were found to neither shift flows on the mainstem of Pilchuck Creek to any significant degree nor cause significant increases in fine sediment in salmon spawning areas. Losses shown in Table 5-6 of Section 5 are attributed to default marine survival assumptions (see also Table 5-9 in Section 5). Currently, flow diversions are not significant on the mainstem of Pilchuck Creek or its tributaries." Note that these results are not acknowledged as the synthetic output of a black-box model with limited data, unrealistic assumptions, and some demonstrably wrong relationships—they are presented as facts.

There is no way to confirm or reject such EDT-generated "facts," produced as they were under the assumption of pristine habitat conditions—the only model results are for conditions that do not exist. Yet the point of the pilot assessment study was to build confidence among agencies and the public in applying EDT to real land-use planning. Some indirect validation of outputs, however, is still possible. The Stillaguamish Tribe's spawning survey in 2003-2004 (Table 2-3a, p. 2-8, 08/12/2005 Draft) only observed 35 live coho adults. EDT estimates (Table 5 in 08/16/2005 Draft) that the historical capacity of Church Creek is 1500 adult coho salmon and that only a small percentage of lost productivity is attributable to flow changes. Is almost all the difference (i.e. historical potential using 1500 adult capacity compared to 35 observed adults) attributable solely to physical habitat changes (and other changes, e.g., hatcheries, but not flow) in and out of the watershed boundaries?

Most importantly, EDT's capacity estimate of a sustainable 1500 adult population in the 11 $mi^2$ Church Creek watershed seems extremely high. A pristine watershed of similar size in Northern California may sustain (being conservatively very high) 300 to 500 adults; even this many would provide more than enough eggs to produce more than enough fry to saturate pristine rearing habitat. Using Figure 5-1 in the Draft Report to roughly estimate the length of channel in Church Creek suitable for coho rearing, the mainstem channel appears approximately 3 miles long. Appendix Table B-1 (Lestelle et al., 2004, p. B-3) uses a density of 0.5 fish/$m^2$ (1 $m^2$ = 10.76 $ft^2$) for 1-age resident rearing coho. Can 1500 adults be sustained by 3 miles of Church Creek, if all 1-age juveniles go on to become smolts and 3% of them survive to return as adults? The answer hinges on (a) what is a reasonable amount of 1-age juvenile coho habitat in a pristine stream channel and (b) how much habitat is necessary to sustain 1500 adults. Back-calculating (b), a 1500 adult capacity would require 1,076,000 $ft^2$ of 1-age habitat. A three-mile-long channel would need 70 $ft^2$ of 1-age habitat for every 1 ft of channel length, i.e., if every square foot of channel was favorable for rearing 1-age juvenile, mainstem Church Creek channel would have to average at least 70 ft wide at baseflows. This would be very wide for a pristine channel with an 11-$mi^2$ drainage area. If we used a width of 25 ft at baseflow and half the mainstem channel providing good rearing habitat (keeping the other assumptions, i.e. 3% smolt survival and a single 1-age juvenile per 21.5 $ft^2$), the sustainable population is 280 adults. The Draft Report (p. 2-9) implies important beaver pond, side-channel, and slough habitats have been lost to urbanization on Church Creek. This certainly would affect historic capacity estimates. Did EDT's data input include square footage estimates of these important lost habitats?

The problem of inappropriate background parameters seems particularly acute in this case because the assumption of historical habitat conditions leads to extraordinarily (i.e. implausibly) high fish performance predictions. Shifting non-flow parameters so far outside the likely range could have substantial impact on the relative predicted flow effects. This potential effect was noted in the study's examination of the rule functions, notably where the effect on flow was found to depend on the values of riparian condition. This is a significant concern.

## 2. INFLUENCE OF URBANIZATION ON SALMON POPULATIONS

Urbanization is damaging salmon populations, and the integrated effects of both runoff alterations and water withdrawals produce recognizable hydrologic effects that have direct correlation with biological conditions (albeit with much scatter). This is probably the most important message for the region to hear, and it needs to be articulated clearly. It is not a conclusion unique to this document—indeed, it is probably better developed elsewhere—but if this report did no more on this subject than summarize the literature reviews and independent studies of others, it would provide a useful service by letting the region move past the debate of whether or not impacts have occurred to a struggle over what to do about them. For example, Section 2.2.1 could be tailored to this study more closely. The generalities of IHA-RVA could be covered in a brief paragraph. Similarly, the history of the King County Normative Flow project is not very useful to most readers, but a summary of its literature review and findings ought to be useful for the reader of *this* document as well.

The conceptual model presented in Figure 2-2 lacks (at minimum) direct links from box 1 to box 3 and box 2 to box 4. Is such an oversimplified model needed? If what is being illustrated is EDT's framework, then it should be labeled as such and accompanied by an honest discussion of its (severe) limitations.

## 3. USE OF HSPF

Given HSPF's long history of use in the region and its transparency, there is little need to defend its very worth (as is needed with EDT). However, some acknowledgment of its limitations is needed, along with the likely sources of error in the outputs. In particular, Section 3.4 of the Draft Report needs to be more forthright in the limitations of the input data, be they in the applicability of the precipitation input for the calibration or the consequences of those limitations on the accuracy and uncertainty of the final outputs. Land-cover determination via LANDSAT classification is critical to the performance of the model, because errors in that classification will propagate throughout the remainder of the investigation. The description of the methodology (and its potential errors), however, is virtually absent in the Draft Report.

Two other aspects of HSPF merit further acknowledgment and, ideally, exploration. The model's performance in simulating base flow needs greater discussion, and the consequences of that performance on both flow and EDT models should be assessed. Finally, HSPF simulation of temperature is much less well documented than its simulation of flow, at least here in the PNW. Therefore, there is no context or guidance for how the temperature-simulation results should be viewed. They are presently offered without caveat or uncertainty—is this warranted?

Overall, the purpose of HSPF appears solely to provide an input parameter for EDT. This belies a long history of using HSPF directly in trying to characterize the impacts of urbanization on streams and the efficacy of proposed mitigation. With recent studies of urban influences on flow regime and their ecological influence, this ought to be a larger, not a smaller, facet of the HSPF analysis.

## 4. "INDEX" METHODS

Index methods of assessing habitat conditions provide robust, but noisy, relationships between land use and instream biology. From a management perspective, this noise may be regrettable but it is also real. Efforts to translate index-based results into "numbers of fish," however, require an assumption of false precision. More honestly, only the broadest ecological trends and corresponding management responses are going to be successful by comparisons between different watersheds. The current draft errs on the side of false precision by implying that one-to-one relationships between land cover and biological conditions exist. Were the document to take a more balanced approach, it would acknowledge the challenge of fine-tuning management actions by use of either broad regional trends or a mechanistic model with inadequate data and only tenuous connections to scientific understanding.

Section 4.5.2 ("Quality indices") is a very brief discussion of what was once characterized as a fully developed alternative to EDT. It provides little substance, however—reduction of May's scattergram (Figure 4-12) into a single-valued function loses most of what's important about these data, and which have been emphasized by more recent (and peer-reviewed) studies. When B-IBI can range between 17 (poor) and 32 (v. good) for a single TIA level, what has been leaned? *Not* that "25" is a good average, but that there's a lot of other factors that need to be understood. Given that we can find B-IBI scores of 25 for TIA between 4 and nearly 40% (i.e., rural landscapes to the equivalent of a Bellevue neighborhoods), how can this application be useful?

As currently presented, the "index" approach is a TIA–B-IBI relationship and requires no special name. Its value is real but very broad, and it is already well covered by other published documents. The Draft Report needs to set its sights significantly higher for this overall approach, or it needs to acknowledge that others have already extracted whatever value is likely to derive from such an approach.

## 5. MANAGEMENT RESPONSE

This study, together with others recently or long-ago published, are sufficient to assert that flow alteration as a result of land-use activities has, can, and will continue to affect instream biota. If the region's interest is in reducing the effects of instream flow alteration on those biota, subsequent efforts should explore how to reduce those flow alterations directly. Well-documented, well-tested methods and models are available to evaluate the management actions that seek to achieve that goal—that would be a useful "next step." At the current state of knowledge and implementation, however, any further efforts to take a shortcut to "numbers of fish" should be recognized as costly, misleading, and indefensible.

The Draft Report's summary (Section 6) broadly overstates any plausible or defensible outcomes of the modeling. For example, page 6-6 uses actual numbers of fish predicted by EDT to make assertions about limitations of fish numbers, despite the earlier promise that EDT results were only going to be used relatively, not as absolute numbers. Section 6.3.3 asserts a physical basis (scour and fine sediment) for a presumed (but unsubstantiated) "loss" of fish, but without some sort of physical modeling that is *not* provided by EDT, there is no basis for making such an assertion. The subsequent bulleted "management implications" thus lack a defensible basis, and they themselves also imply a level of effectiveness for the various mitigation measures that have never been demonstrated.

Simpler, more empirically based models will be more informative than EDT for making flow management decisions. Such simpler models may not directly predict fish performance, but a robust prediction of a good surrogate (e.g., natural river physical processes or healthy invertebrate community) is more useful than a questionable prediction of fish performance. It is worth noting the management insensitivity to model results, as expressed (for example) by the discussion on page 5-36. If this is true, why model? Similarly, the conclusions and recommendations at the bottom of page 5-32 don't seem to have required EDT.

Given all the uncertainties, the review team is not convinced of the need to estimate populations, particularly if no one will feel (or should feel) that the numbers are credible. Rather, more attention to physical thresholds or direct correlation between flow parameters and biological health would be a much better way to assess land-use change and predict future change using many of the tools presented in the Draft. We suggest some "next steps" as alternatives to an ill-advised investment of scarce resources to further EDT-based analyses:

- Make use of existing published research, and any outcomes from King County's Normative Flow Project, to identify credible relationships between flow parameters and in-stream biological conditions (see Konrad and Booth, 2005). These relationships could then guide more targeted applications of HSPF to investigate the potential consequences of future flow alterations on stream health. This original goal of the Instream Flow Assessment Pilot Project was, and remains, an important and appropriate focus for salmon-recovery efforts in the region.
- Genuinely explore alternatives to EDT based on empirical studies. These models might not have the mechanistic realism attempted by EDT but they probably have more predictive power. Analyses suggest that using only a few habitat parameters in a simple regression can have predictive power similar to, and more justifiable than, the vastly more complex EDT model (e.g., Green et al., in press).
- Explore alternate ways to use the EDT habitat datasets, many of which contain quite a bit of information characterizing habitat condition. Simpler analyses could be conducted using these datasets that do not require dubious application of the EDT model.

- Consider other measures of stream health instead of (or in addition to) predicted fish numbers in making flow management decisions. There is no reason to expect that fish numbers can be predicted by the means that have been explored here.

Most broadly, land-use planning across the region will need to explicitly consider the health and recovery of salmon and steelhead populations if the goals of Shared Strategy are to be reached. Use of EDT was explored in the Draft Report because salmon and steelhead are important to the Puget Sound area, and the model offered the hope of concrete predictions of fish populations under alternative future scenarios. This hope was well-intentioned but simply cannot be accomplished at the present time.

"Salmon friendly" land-use planning can proceed, however, without estimating fish populations. Guidelines based on physical watershed variables, and that correlate with ecologically relevant responses, can be scientifically formulated and some have already been published for the region (e.g., Booth and others, 2004; McBride and Booth, 2005). Different ranges of watershed suburbanization can be identified that broadly correlate with alternative outcomes that preserve, chronically impair, or permanently degrade future salmon and steelhead habitat. Local communities and larger watershed authorities could more formally incorporate these suburbanization guidelines as rough, but unambiguous, land-use planning goals that explicitly acknowledge the consequences of future growth on salmon and steelhead. Fish population forecasts are not necessary for the region to decide whether or not to reduce the consequences of urban development on the biological health of lowland rivers and streams; nor are they necessary to implement fundamental salmon-protection strategies.

# REFERENCES

Bisbal, G.A. 2002. The best available science for the management of anadromous salmonids in the Columbia River Basin: Can. J. Fish. Aquat. Sci. v. 59, p. 1952-1959.

Booth, D.B., J.R. Karr, S. Schauman, C.P. Konrad, S.A. Morley, M.G. Larson, and S.J. Burges. 2004. Reviving urban streams: land use, hydrology, biology, and human behavior: Journal of the American Water Resources Association, v. 40(5), p.1351-1364.

Greene, C.M., D. Jensen, G.R. Pess, E.M. Beamer, E.A. Steel, in press. Effects of environmental conditions during stream, estuary, and ocean residency on Chinook salmon return rates in the Skagit River, WA: Transactions of the American Fisheries Society.

Konrad, C.P., and D.B. Booth. 2005. Ecological significance of hydrologic changes in urban streams: chapter in L.R. Brown, R.H. Gray, R.M. Hughes, and M.R. Meador, editors. Effects of urbanization on stream ecosystems. American Fisheries Society, Symposium 47, Bethesda, Maryland.

Lestelle, L.C., Mobrand, L.E., and W.E. McConnoha. 2004. Information Structure of Ecosystem Diagnosis and Treatment (EDT) and Habitat Rating Rules for Chinook Salmon, Coho Salmon, and Steelhead Trout. Mobrand Biometrics, Inc. 29 p. + Appendices.

May, C.W., Welch, E.B., Horner, R.R., Karr, J.R., and B.W. Mar. 1997. Quality Indicies for Urbanization Effects in Puget Sound Lowland Streams. Final Report for Washington Department of Ecology. Water Resources Series Technical Report No. 154. Department of Civil Engineering, University of Washington, Seattle, WA.

McBride, M., and D. B. Booth. 2005. Urban impacts on physical stream conditions: effects of spatial scale, connectivity, and longitudinal trends: Journal of the American Water Resources Association, Vol. 41, No. 3, pp. 565-580.

Morley, S.A., and J.R. Karr. 2002. Assessing and restoring the health of urban streams in the Puget Sound basin: Conservation Biology 16:1498-1509.