COOPERATIVE MONITORING EVALUATION AND RESEARCH 03-312

Water Typing Model Field Performance Assessment Pilot Study

MARCH 2005

SUBMITTED BY TERRAPIN ENVIRONMENTAL
#PSC 03-286

Water Typing Model
Field Performance Assessment
Pilot Study

29 March 2005

Submitted to:

Washington State
Department of Natural Resources
Forest Practice Division
PSC 03-286

Submitted by:

C. Edward Cupp
Terrapin Environmental
988 Twisp River Road
Twisp, Washington 98856

Water Typing Model
Field Performance Assessment
Pilot Study

Table of Contents

Abstract: The Instream Scientific Advisory Group (ISAG) Statistical Team has developed a multiple parameter logistic regression model combined with heuristic stopping rules that predicts boundaries between fish habitat and non-fish habitat. This pilot study was conducted to assess the field methods and data processing protocols of a study plan designed to provide an unbiased characterization of model performance. This study also provides additional analysis to supplement to the model evaluation conducted during model development (Conrad et al. 2003). The field model performance assessment was conducted on 369 predicted end of fish points (EOFPs) within 15 randomly selected basins. Mean net and absolute prediction field error distances were estimated as -114 $\pm$ 78 feet and 176 $\pm$ 60 feet, respectively. When only predicted EOFPs associated with defined channels were included in the computation of model performance, a total of 158 EOFP predictions were excluded from analysis and the mean performance measures decline to -185 $\pm$ 129 feet and 287 $\pm$ 94 feet for net and absolute error, respectively. The model correctly classified fish absence or presence in over 90% of the DEM-portrayed stream network. Seventy five percent of the predicted EOFPs associated with defined channels were correct (zero error). Approximately 92% the of total error distance was associated with the 41 terminal EOFP predictions. Six terminal predictions situated upstream of steep channel features that are readily apparent on the DEM generated channel profiles accounted for 46% of the total error measured. The pilot study demonstrated the practicality of the study plan protocol and identified refinements to the procedures. Sample size requirements for further study application are addressed.

Water Typing Model
Field Performance Assessment
Pilot Study

March 29, 2005

Introduction

Revised Washington state forest practice regulations require the development of a GIS-based water typing system that delineates boundaries between fish habitat and non-fish habitat. The Instream Scientific Advisory Group (ISAG) Statistical Team has developed a multiple parameter logistic regression model combined with heuristic stopping rules that predicts the uppermost extent of fish habitat.

The ISAG Statistical Team made several assessments of the model using fish distribution data from thousands of previously surveyed streams. The methods used to estimate precision and balance at the stream network scale were not ideal due largely to the fact that the existing dataset was not specifically intended for model development and assessment purposes. There was no deliberate sampling design and very few complete basins examined in the field. Nor was there one standard protocol used for establishing the observed EOFP in the field. All error measurements were conducted in the GIS framework with no field measures of error.

The ISAG requested development of a field Study Plan to examine the predictive performance of the water type model across watersheds of western Washington. Assisting ISAG under Washington DNR Personal Service Contract 02-198, Terrapin Environmental developed a Study Plan to model performance across State and private forest lands in western Washington. The plan, entitled "Water Typing Model Field Performance Assessment Approach and Procedures" (Study Plan), was submitted for review in December 2003 and finalized in December 2004. Prior to final acceptance of the Study Plan, the ISAG requested the study protocol be tested as part of a pilot study.

The pilot study was conducted as a field test to assess the field methods and data processing protocols described in the Study Plan. The pilot study also provides additional analysis to supplement the model evaluation conducted during model development (Conrad et al 2003). In order to provide comparisons between mapped and field prediction error measures as well as unbiased characterization of model performance under a standard protocol, the pilot study was conducted on a random selection of basins throughout the affected forest lands.

Methods

The sampling design, field protocols, and data processing and analysis techniques examined in this pilot study are described in detail in the Study Plan. The Study Plan outlines the maps, photos, electronic shape files, databases, and other equipment required to efficiently conduct field surveys. The field survey techniques have been reduced to a very manageable set of data collection requirements and are described in detail. Rules for computation of performance metrics and data processing procedures required to characterize model performance are provided in the Study Plan.

We used the stratified random cluster sampling strategy to provide an unbiased characterization of model performance across Forest and Fish lands of western Washington. For the pilot study, we limited our sampling frame to include only fourth order sub-basins in which the entire length of stream was within State or private forestland ownership. Each fourth order sub-basin (FOSB) that was wholly contained within lands under the jurisdiction of the Forest and Fish Agreement was grouped by geographic region. Three FOSB were randomly selected from each of the five geographic regions under investigation for the pilot study (Figure 1).

Upon ISAG's request, we surveyed two additional FOSB located within the Coastal region. Previous surveys in portions of these basins had indicated that model predictions had substantial error (large underestimation) in distinguishing fish habitat. Our efforts were intended to provide a complete survey of the two basins in order to reconnoiter these apparent trends from the partial surveys and to further explore factors affecting model performance. Although data from these basins are included in assessment of mapping discrepancies, they are excluded from the model performance estimates that follow because of the biased process in which they were selected for sample.

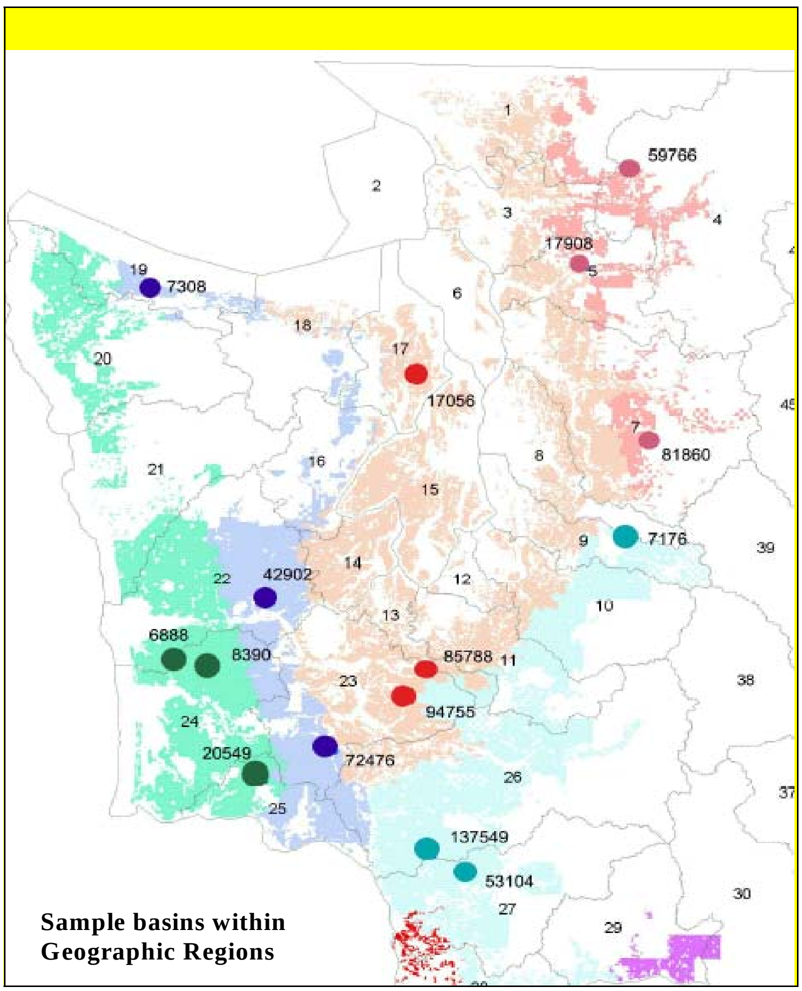

Figure 1. Fifteen Fourth order sub-basins randomly selected for sample under the pilot study.

Results and Discussion

During July through September 2003, field crews investigated 15 randomly selected FOSBs containing 376 predicted EOFPs. Two additional coastal basins were examined containing an additional 73 predicted EOFPs. Another 124 observed EOFPs were encountered in streams that were not portrayed by the DEM-generated stream network. The upstream extent of fish use was determined for streams in each sampled basin, thus verifying fish presence / absence for the entire stream network and the accuracy of 449 EOFPs predictions.

Discrepancies between Mapped and Field Located EOFPs

Inaccuracies and misrepresentations of stream locations derived through DEM analysis can increase the difficulty in locating predicted points on the ground or placing field observed points on a map. Because of this, ISAG expressed concern regarding potential bias in locating predicted EOFPs in the field or identifying field observed points on DEM maps, which in turn effect measurements of model prediction error. To address this potential for bias, we examined the correspondence between the map and field locations of predicted EOFPs, as well as how well the DEM-portrayed the distribution and abundance of tributary junctions.

The DEM-portrayed location of the stream channel and its tributary junctions are commonly not aligned precisely with field conditions. Using digital imagery of aerial photographs scaled with DEM maps (on which the model predictions are generated), and GPS coordinates when available, we evaluated the accuracy of 209 DEM predicted EOFP locations. Seventy three percent of the DEM-portrayed EOFPs were considered to be within 100 feet of the field location of the point with 91% located within 200 ft (Table 1). Twenty nine percent of the lateral confluence and the tributary junction predictions were out of alignment by more than 100 feet. In each of these situations, the predicted EOFP was field located at the actual tributary junction and not at the literal mapping coordinate as portrayed on the DEM network. Likewise the field measured prediction error (field ΔEOFP) was determined between the tributary junction field location and the field observed EOFP (Figure 2).

Table 1. Count of predicted EOFPs with grouped by accuracy classes of the locations portrayed by the DEM stream network. Numbers in parentheses represent proportion of each predicted EOFP boundary type.

	Distance between DEM-portrayed prediction and field location			
	> 300 ft	200 - 300 ft	100 - 200 ft	< 100 ft
Mid channel	0	0	2 (0.08)	22 (0.92)
Lateral confluence	6 (0.04)	10 (0.06)	27 (0.17)	120 (0.73)
Tributary junction	2 (0.10)	0	8 (0.38)	11 (0.52)
	8 (0.04)	10 (0.05)	37 (0.18)	153 (0.73)

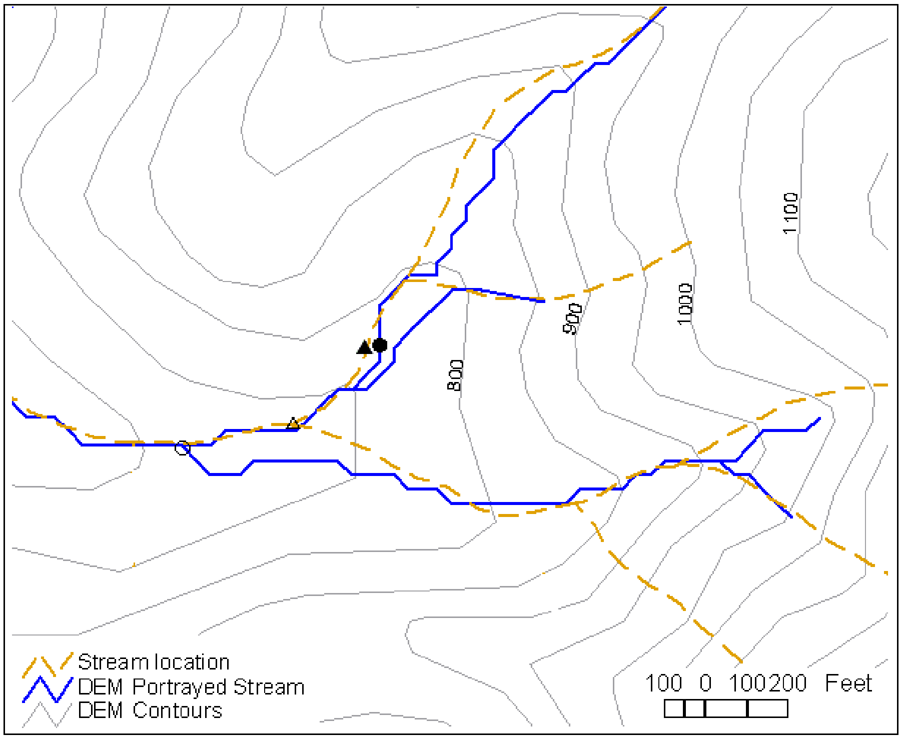

Figure 2. Schematic illustrating the discrepancies among mapped and field locations of stream channels, tributary junctions and predicted and observed EOFPs. The open circle identifies the DEM-portrayed EOFP prediction, whereas the open triangle represents the field located predicted EOFP. The closed triangle identifies the field location of the observed EOFP, whereas the closed circle identifies where the observed EOFP was mapped on the DEM network.

Discrepancies between field and DEM-portrayed EOFP predictions were less common with mid-channel point type (Table 1). Crews normally were able to field locate predictions by use of GPS readings or by determining the distance the point is situated from readily identified aerial photo reference (e.g. road crossing, tributary junctions, riparian openings, rock outcrops, ponds) or nearby GPS grab point. We recognize that in the absence GPS locations of prominent recognizable photo reference points near the prediction point, the risk of error in ground placement and mapping increases. In at least four cases, crews located the points by measuring distances along the channel valley bottom from tributary confluences, but had no clear, nearby reference points. In these situations, GPS locations were unavailable and photo references were limited to points situated a considerable distance from the predicted or observed EOFP. The greater the distance between reference points and the target location, the steeper and more rugged the terrain, and denser the overstory vegetation, the risk of error increases in locating and mapping EOFPs.

Prior to the pilot study, we expected to encounter many situations in which a mid-channel EOFP prediction corresponded to abrupt changes in channel gradient portrayed on the DEM channel network (Figure 3). We envisioned attempting to locate the predicted EOFP in the field at a point where channel profiles most accurately correspond to DEM estimated conditions. As described in the study protocol, we intended to identify and mark the mapped location in the field regardless of the correspondence between the DEM estimated and the field measured channel profile. An adjusted EOFP location would then be identified in the field at the gradient break feature that corresponds to the DEM portray channel profile. However, we did not encounter any such situation in the pilot study. We did not adjust field locations of mid-channel EOFPs to correspond with DEM generated gradient features.

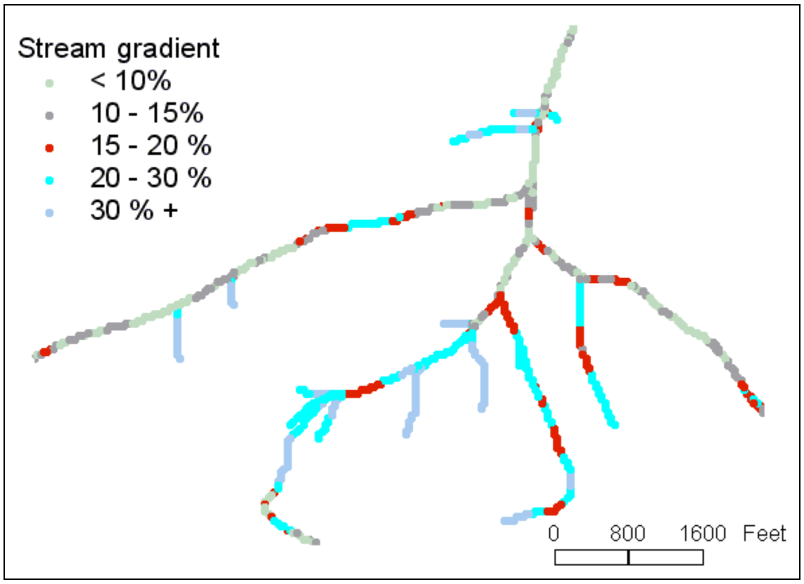

Figure 3. Channel profiles estimated by the DEM used to implement the water typing model. Each cell represents approximately 30 foot stream reach.

Comparison of GIS-Based Error Distances and Field-verified Error Distances

Assessment of model performance conducted by Conrad et al (2003) relied entirely on GIS-derived estimates of prediction error. The distance between the observed and predicted EOFPs were measured within the GIS framework using the strings of DEM-generated cells portraying the stream network. All field observed EOFPs were mapped within the DEM stream network accounting for mapping discrepancies described in the previous section. It is believed in some situations, this process can lead to substantial distances between model estimated and field measured error distances. Pilot results provide the opportunity to evaluate the relationship between GIS-based estimates and field measured error distances.

Seventy-seven observed EOFPs with an absolute error distance greater than zero, but less than 4000 feet were encountered. The difference between the absolute value of the field observed and the GIS estimated error distance varied slightly (mean=7 feet, median=2 feet), but did not differ significantly at $\alpha = 0.05$ (Wilcoxon Rank Test: P=0.09). The correlation between absolute field error distance and absolute map error distance is 0.9805 (Figure 4).

Map-estimated error distance exceeded field error distance in 64% of the cases where model prediction error was observed. The difference between map and field error measures was not strongly correlated with the absolute map estimated error (Spearman Rank Correlation = 0.1046). The largest difference between map estimated and field-verified error distances could be best explained by simple visual comparison between the DEM-portrayed stream network and high quality aerial photographs (Figure 2 for example). These discrepancies were caused largely by inaccurate DEM portrayal of channel confluences.

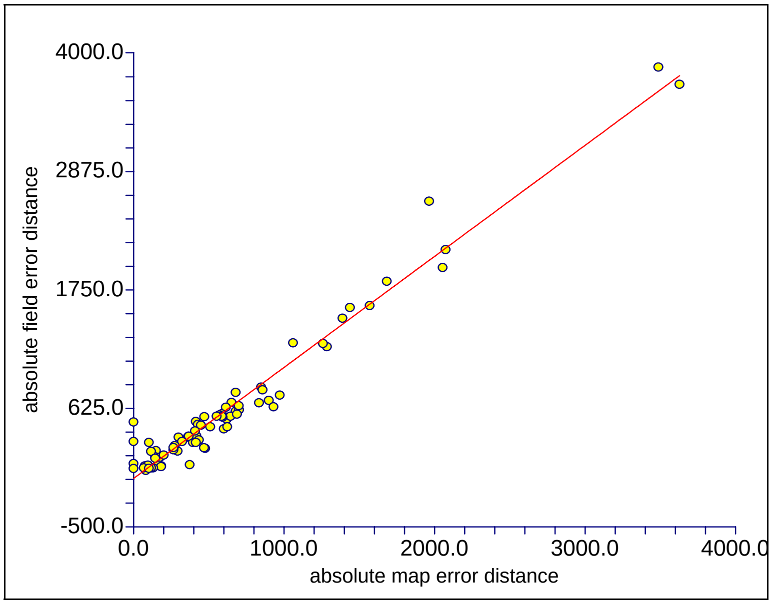

Figure 4. The straight line relationship of the absolute field error distance and absolute map error is represented as: absolute field error distance = (-40.7101) + (1.0537) absolute map error distance using the 77 observations in this dataset (R-Squared = 0.96).

<u>Reliability of the DEM to Depict Streams</u>

Inaccuracies and misrepresentations of stream locations derived through DEM analysis can be a confounding factor in assessing model performance. One source of error in model performance may be attributed to the mapping accuracy of the DEM-generated stream network, especially in depicting the existence and location of streams. The DEM may not accurately represent the abundance and distribution of channels observed in the field. Occasionally the DEM-generated stream network portrays channels where none actually exist in the field. Additionally, stream channels observed in the field may not be included in the DEM-generated stream network.

We investigated 158 predicted EOFPs at which there was no evidence of channelized surface flow or seepage (apparent sediment scour and transport, flotsam lines, stagnant or dry puddles) within 350 feet (usually at least twice this distance) of the map portrayed channel confluence. In a few situations intermittent reaches of seasonal, defined channels or more commonly wetland swales (nearly completely dry at time of survey) were observed further "upstream". The predicted EOFP associated with each of these channels was considered a non-defined channel (NDC) prediction. NDC predictions accounted for 35% of the model predicted EOFPs across 17 FOSBs.

Conversely, we encountered 124 field-verified EOFPs on streams not represented by the DEM stream network that flowed directly into otherwise mapped fish-bearing streams, and as such had no predicted EOFP. We identified each of these field-verified EOFPs as "non-mapped EOFPs". Fish were encountered in three streams not represented by the DEM stream network. Fish were found at a maximum distance of 300 feet upstream from the confluence with a larger, mapped channel. Ninety-six (96) of 124 field-verified "non-mapped" streams were depicted on the DNR Hydrolayer stream network; the remaining 28 were not shown on either the DEM or the DNR stream networks. No fish were observed in the 28 streams not represented by the DNR or DEM stream networks.

Three predicted EOFPs across were excluded from further analysis due to substantial inaccuracies in the DEM stream network mapping. In one FOSB, the DEM depicts a portion of sample basin as draining a relatively small drainage area. The DEM indicates a steep basin to the north and west as draining into an tributary branch, when in fact the steep basin is the headwaters of the sample basin (Figure 5). The DEM mapping resolution was insufficient to distinguish drainage patterns and individual channels within the an extensive wetland complex situated between the two portrayed basins. The mainstem of the sample basin in fact is fed by at least two other FOSBs situated higher in the drainage system that contain several terminal and lateral EOFP predictions. The DEM-portrayed the upstream extent of the sample basin nearly 2 miles below its actual headwater reaches, where several field observed EOFPs were surveyed. These observed EOFPs were excluded from the analysis because they are situated within an entire differently basin that was not selected as a sample FOSB. The other FOSB consisted of a basin dominated entirely by wetland channels and marshes located along flat, glacial-fluvial terraces. Here again, a DEM-generated stream network poorly portrayed both extent and location of stream channels. Two other predicted EOFPs were excluded from the field survey and analysis due to access issues.

<u>Accuracy of the Digital Elevation Models in Estimating Channel Gradients</u>

This Study Plan includes evaluation of the correspondence between predicted and field observed gradient features in order to evaluate if prediction error may in part be explained by poor DEM resolution. In some situations, digital elevation models alone may be insufficient to identify specific geomorphic features responsible for the upstream limits of fish distribution. The current water type model relies on stream gradients as portrayed by DEM averaged over a 100 m reach. To evaluate the accuracy of the DEM in portraying channel slope, we compared field measured reach average channel gradients with the DEM estimated reach average gradients at 215 EOFPs. We recognized that the DEM stream network may not accurately locate channel position within the valley bottom and the precise location of the EOFPs as represented on the DEM map may not lie directly in the channel. Given the discrepancies between the stream locations portrayed by the DEM maps and actual field locations, this assessment is intended as a general comparison of map and field observed channel gradients

Field observed and DEM estimated 100 m reach average channel gradients measures did not differ (Wilcoxon Rank Test: P=0.21). On average, the field measured and the DEM estimated channel gradient varied only slightly (mean difference = 0.8 %). To provide information on the ability of the DEM to identify specific features (steep channel inflections or waterfalls) that limit the upstream distribution, we also compared field observed gradient measures with the DEM gradient estimates immediately upstream of 223 EOFPs. On average, the field observed channel gradient and the DEM gradient estimate (estimated for approximately 10 m reach) varied only slightly (mean difference = 1.6 %); although the two gradient distance measures did differ significantly (Wilcoxon Rank Test: P=0.09). In general, the DEM generated gradients immediately upstream of an EOFP slightly underestimated the field observed gradients.

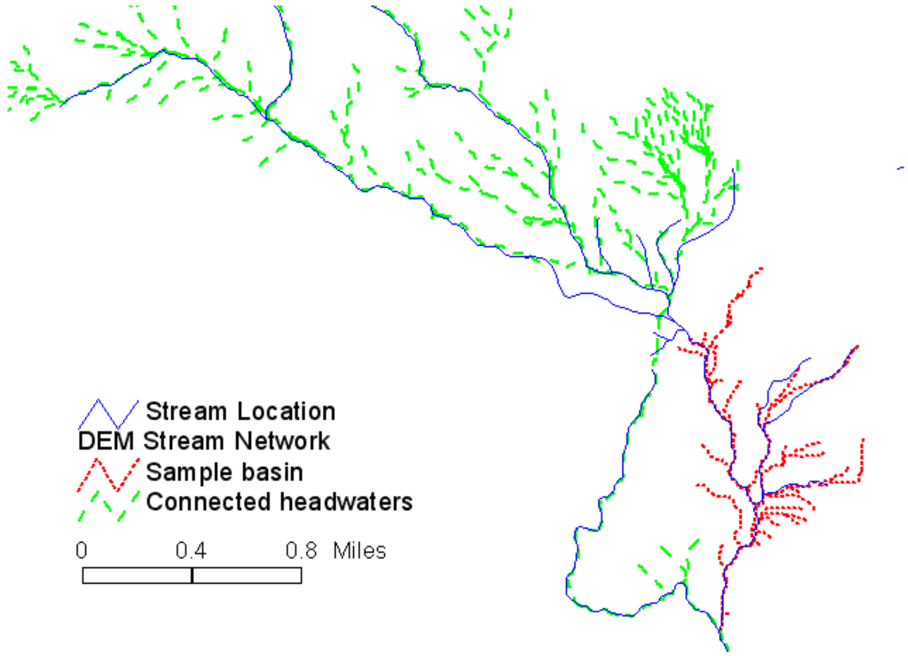

Figure 5. Discrepancies between the DEM portrayal and field locations of stream network. The solid blue lines represents the approximate location of primary tributary trunks only.

At both the 10 m and 100 m reach scale, it would appear that the channel gradient estimated between individual DEM points (approximately 10 m length) or groups of points could identify channel inflections at a scale as small as 10 m. However, discrepancies between the field observed and DEM generated gradients were most pronounced at EOFPs where steep cascades, chutes, or waterfalls effectively prohibited upstream movement of fish (mean difference =6.5%, n = 25). Average field measures exceeded 20% upstream of these points, whereas DEM generate estimates were around 13%. The change in channel slope immediately upstream of an EOFP was substantial, averaging over 13% as measured in the field. However, the change in channel slope portrayed by the DEM 100 m reach estimates averaged only 6%, suggesting that DEMs are not always accurate at identifying the severity of short, abrupt inflections in channel slope. Nevertheless, in nearly all of these cases, an abrupt change in channel slope was identified by the DEM, albeit not as severe as observed in the field.

Model Performance Measures

Model performance can be summarized in a variety of ways and at more than one level of resolution. Under the validation Study Plan, we have proposed to measure model performance at both the EOFP scale and at the basin scale. Results pertaining to model performance under this section are based on the 15 randomly selected basins only. The two additional Coastal Basins were excluded because they were not randomly selected for sample, but rather were intentionally surveyed as a follow up to an earlier investigation.

<u>Accounting for DEM Inaccuracies in Model Performance Measures</u>

Given the mapping inaccuracies previously described, inspection of the pilot data suggest that rules for inclusion of specific data into the analysis of overall model performance may be warranted. It is recognized that the DEM may not accurately represent the abundance and distribution of stream channels observed in the field. Occasionally the DEM-generated stream network portrays stream channels where none actually exist or misses channels that could be used by fish. As discussed above, inclusion of NDC predictions and the "correct" classification of non-fish habitat predictions situated "upstream" can provide what some would argue a false sense of model success.

We initially report on the precision and balance of errors based on the field measured error distance associated with all predicted EOFPs. The initial summary includes the predicted EOFPs situated on DEM depicted stream channels that do not exist in the field. Prediction error distance (ΔEOFP) of these predicted points on "non defined channel" (NDC) predictions are recorded zero, as there are obviously no fish "upstream" of the EOFP along DEM predicted stream.

Secondly, we present prediction error summaries of model points associated with defined channels only. That is, the 158 "NDC predictions" are excluded. Although this summary provides a more accurate representation of model performance on the ground, there remains another 124 field observed EOFPs associated with streams not accounted for by the DEM stream network.

To account for streams encountered that are not depicted on the DEM network, we provide a third measure of model performance that includes measured field error distances for each non-mapped EOFP. We tabulated each of these non-mapped EOFPs as a lateral confluence boundary type (i.e. the EOFP prediction would be at the confluence of the non-mapped channel with a mapped channel). Fish encountered in the non-mapped stream upstream of the confluence constituted an underestimation of fish use, and the distance between the confluence and the observed EOFP was recorded as a positive value.

In addition to accounting for mapping inaccuracies, we present a fourth measure of model performance to account for lateral confluence EOFP predictions that are situated upstream of observed mid-channel EOFPs. Under the rules for measuring error distances (Section 4.1.1 of the Study Plan), if no fish are observed in the lateral tributary upstream of the predicted point, then the error distance is zero regardless of whether the observed EOFP on the mainstem is situated upstream or downstream of the lateral tributary confluence. Clearly, the distance between the predicted EOFP and the observed EOFP on the primary trunk channel represents prediction error. However, if the observed EOFP on the primary trunk is situated downstream of lateral confluence predictions, the model incorrectly establishes a lateral confluence EOFP prediction at the mouths of each of the low order tributary channels entering the primary trunk. Each of these predicted lateral confluence EOFPs would conceivably have an associated over-prediction error (i.e. the distance between the tributary confluence, or lateral confluence point, and observed mid-channel EOFP situated downstream on the primary trunk). Such an accounting system would result in duplicate counts of model error where predicted lateral confluence EOFPs are situated upstream of terminal observed EOFPs. To avoid duplicate counts of model error, error distance associated with each of these lateral confluence predictions upstream of the observed EOFP is considered zero. Although the modeled lateral confluence EOFP correctly predicts fish absence for the tributary, including these points as correct predictions could be misleading, as fish are not present immediately downstream of the point. Model performance summaries that exclude lateral confluence EOFP predictions situated upstream of the observed fish distributions demonstrate the influence of these "correct" lateral EOFP predictions on overall model performance. In the following section, we provide the four different measures of model error.

Precision and Balance of Prediction at the EOFP Scale

Although the majority of model predictions were correct (error distance = 0), the distribution of error distances were skewed (Figure 6). Large negative errors (overestimates of upstream fish habitat) were more common then very large positive errors (underestimates of upstream fish habitat)[Figure 7]. Large negative errors were commonly associated with end of fish points situated downstream of reaches of sustained steepness and/or waterfalls, although suitable habitat with moderate to large drainage areas occurred upstream.

The model performed fairly well when applied to the randomly selected sample fourth order sub-basins; it correctly located the upstream extent fish use at 75% of the predicted EOFPs associated with defined channels across 15 basins (Table 2). On average, the model overestimated upstream fish use by 185 feet (based on error distances associated with predicted EOFPs associated with defined channels), although the number of underestimates (31) slightly exceeded the number overestimated predictions (26) .

Influence of EOFP Boundary Types

Boundary type designations describe the relationship of the EOFP to confluences with other streams. Similar to Conrad et al. (2003), we found the EOFP boundary type designation accounts for variation in model performance (Table 3) Prediction error is lower and less variable for lateral confluence EOFP predictions as compared to the mid-channel and tributary junction predictions (collectively referred to as terminal predictions) [Figure 8]. Lateral confluences are side tributaries entering a main channel and are frequently associated with a large change in basin area and/or a gradient break. Over 90% of lateral confluence predictions investigated were correct (ΔEOFP = 0). Terminal predictions accounted for only 11% of the predicted points, yet 92% of total error distance was associated with the terminal predictions.

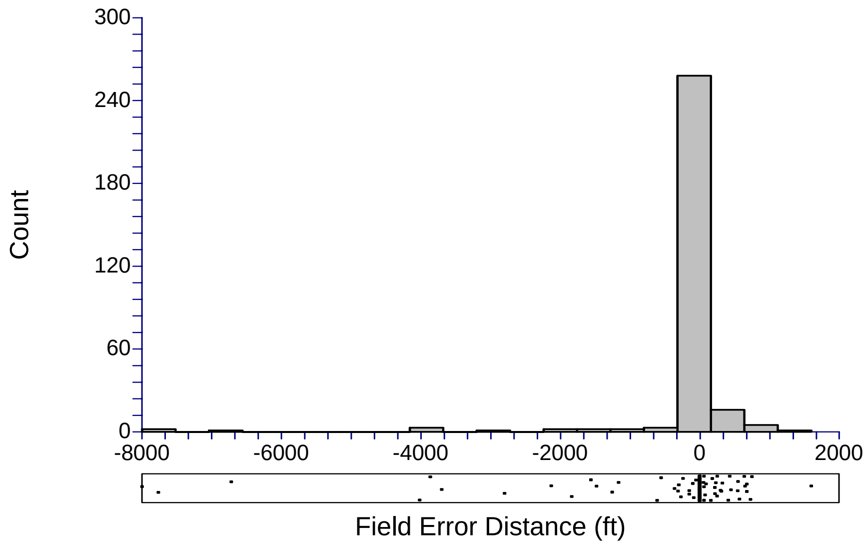

Figure 6. Frequency distribution of error distances of field-verified end of fish point predictions. All field observed error is accounted for in the illustration, which excludes NDC predictions but includes non-mapped EOFPs.

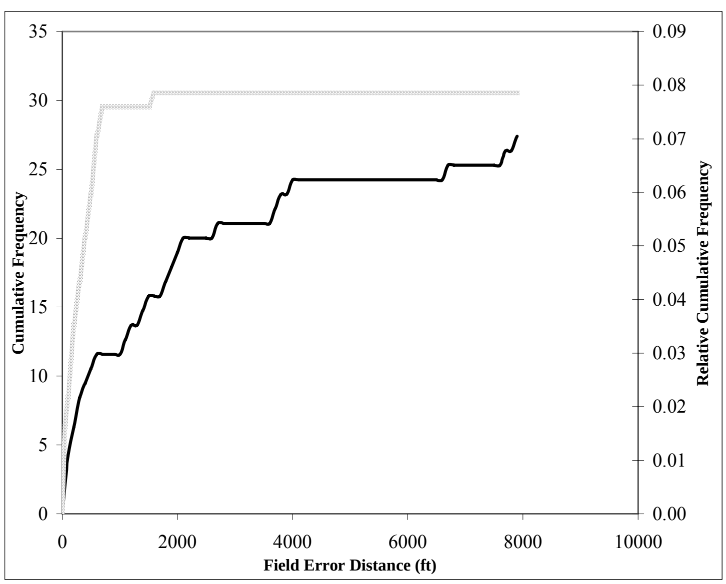

Figure 7. Cumulative frequency polygon of the field error distance associated with underestimated (dotted line) and overestimated (solid line) predictions. The relative cumulative frequency is in relation to the total number of predicted EOFPs investigated (87 percent had a field error distance of zero and are not included on the graph). All field observed error is accounted for in the illustration, which excludes NDC predictions but includes non-mapped EOFPs.

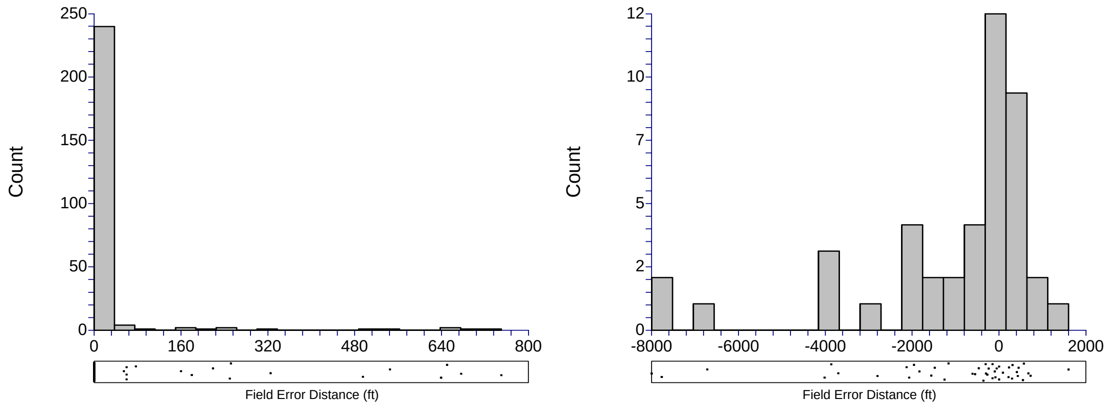

Figure 8. Frequency distribution of error distances of field-verified lateral confluence end of fish point predictions (left side) and terminal end of fish point predictions (right). All field observed error is accounted for in the illustration, which excludes NDC predictions but includes non-mapped EOFPs.

Influence of Prominent Waterfalls and Steep Reaches

Considerable variation in prediction error is a direct result of several large overestimates of fish use observed within only three of the sample FOSBs. Prediction error measures of the terminal EOFPs in these three basins greatly influence the estimated average. Steep, precipitous waterfalls or extended cascade reaches that have apparently prohibited fish from colonizing upstream reaches are situated near or just downstream of the lower end of these basins. Six predicted EOFPs in these three basins alone account for 46% of the total absolute prediction error observed in the field.

Our findings suggest that topographic maps, DEM estimated stream gradients, and local knowledge of the location of large waterfalls might be a useful tool to screen for potentially large overestimate errors. This initial screen is likely to account for a substantial portion of model error across the study area. For example, model performance improves substantially when applied only to the 12 randomly selected fish-bearing basins (Table 4). Mean absolute error distance on terminal EOFPs (mid-channel and tributary junctions) reduces by nearly 40%. Although the proportion of correctly located EOFPs remains at 87%, the model on average overestimates upstream fish use by 39 feet as compared to 128 feet with the screened basins included. Nearly all of the error in 20% of the sample basins is related to steep gradient features that are readily apparent on the DEM generated stream profiles.

Table 2. Summary statistics for absolute and net field error distances (feet) by percentiles. Bounds on the error of estimation are calculated following cluster sampling procedures described by Scheaffer et al. (1996).

	Count	Mean	Confidence Bound	Percentile											
				100	95	90	85	75	50	25	15	10	5	1	0
All predicted EOFPs included															
Number of Points				369	351	332	314	277	185	92	55	37	18	4	0
Net Field Error Distance	369	-114	78	1600	240	0	0	0	0	0	0	0	-414	-4827	-7997
Absolute Field Error Distance	369	176	60	7997	702	270	27	0	0	0	0	0	0	0	0
Only predicted EOFPs associated with defined channels included															
Number of Points				227	216	204	193	170	114	57	34	23	11	2	0
Net Field Error Distance	227	-185	129	1600	442	164	0	0	0	0	0	-168	-1725	-7472	-7997
Absolute Field Error Distance	227	287	94	7997	1741	616	308	0	0	0	0	0	0	0	0
Predicted EOFPs associated with defined channels included plus all field verified non-mapped EOFPs															
Number of Points				299	284	269	254	224	150	75	45	30	15	3	0
Net Field Error Distance	299	-142	113	1600	310	0	0	0	0	0	0	0	-1164	-6718	-7997
Absolute Field Error Distance	299	219	92	7997	1256	430	180	0	0	0	0	0	0	0	0
Predicted EOFPs associated with defined channels plus all field verified non-mapped EOFPs; Lateral EOFP predictions situated upstream of field verified fish distribution excluded															
Number of Points				229	218	206	195	172	115	57	34	23	11	2	0
Net Field Error Distance	229	-185	79	1600	440	160	0	0	0	0	0	-239	-1698	-7451	-7997
Absolute Field Error Distance	229	286	170	7997	1718	610	310	27	0	0	0	0	0	0	0

Table 3. Summary statistics for net and absolute field error distance (feet) by end of fish point boundary types. Bounds on the error of estimation are calculated following cluster sampling procedures described by Scheaffer et al. (1996).

		Net Field Error							Absolute Field Error						
Last Fish Point Boundary Type	Number of Points	Mean	Conf. Bound	Percentile					Mean	Conf. Bound	Percentile				
All predicted EOFPs				95	75	50	25	5			95	75	50	25	5
Mid-channel and Tributary Junction	41	-1146	666	723	265	-270	-1898	-7660	1466	578	7660	1898	552	255	5
Lateral Confluence	328	15	11	30	0	0	0	0	15	11	30	0	0	0	0
Only predicted EOFPs associated with defined channels															
Mid-channel and Tributary Junction	41	-1146	666	723	265	-270	-1898	-7660	1466	578	7660	1898	552	255	5
Lateral Confluence	186	27	16	205	0	0	0	0	27	16	205	0	0	0	0
Predicted EOFPs associated with defined channels plus non-mapped observed EOFPs															
Mid-channel and Tributary Junction	43	-1102	662	717	230	-270	-1835	-7556	1407	1993	7556	1835	550	239	11
Lateral Confluence	256	19	10	63	0	0	0	0	19	10	63	0	0	0	0
Predicted EOFPs associated with defined channels plus all field verified non-mapped EOFPs; Lateral EOFP predictions situated upstream of field verified fish distribution excluded															
Mid-channel and Tributary Junction	43	-1102	662	717	230	-270	-1835	-7556	1407	584	3955	1545	446	203	0
Lateral	186	27	8	205	0	0	0	0	27	8	205	0	0	0	0

Table 4. Measures of model performance by end of fish point boundary types within basins where no prominent waterfalls or extended very steep gradients exclude fish from the FOSB.

		Net Field Error			Absolute Field Error			
Last Fish Point Boundary Type	Number of Points	Mean	Median	COV	Mean	Median	COV	Proportion of Total Error
Predicted EOFPs associated with defined channels plus non-mapped observed EOFPs of fish-bearing sample basins only								
Mid-channel	19	-556	-150	-2.3	884	430	1.2	0.51
Lateral	211	24	0	4.5	24	0	4.5	0.15
Tributary Junction	15	-269	-53	-4.5	729	300	1.3	0.33
	245	-39	0	-12.4	133	0	3.5	

Influence of Ecoregion

The Study Plan uses stratified sampling to ensure that sample basins are spread throughout western Washington. The study population is divided into 6 geographic regions, or ecoregions, with relatively homogenous geologic and climatic conditions. Five of the six ecoregions were investigated under this pilot study. Considerable variation was observed in both absolute and net field error distances across the five regions (Table 5). Net field error distances for all boundary types combined differ significantly at $\alpha = 0.05$ between ecoregions (Kruskal Wallis: $P = 0.013$). The Kruskal-Wallis multiple-comparison Z-value test detect significant differences only between the Puget Sound and the Coastal Basins. Differences were not significant for absolute error of all boundary types combined (Kruskal Wallis: $P = 0.61$). Differences were also not significant for the net and absolute error of terminal predictions alone (Kruskall Wallis: $P = 0.06$ and $P = 0.25$, respectively).

The mid-channel and tributary junctions predictions (terminal EOFPs) were more commonly underestimated in the Coastal and Interior Olympic ecoregions as compared to the other three regions. Nearly half (47%) of the terminal predictions underestimated fish habitat as compared to only 15% in the other three regions. Error in lateral confluence EOFP predictions was also more common in the Coastal and Interior Olympic ecoregions. Fish were observed upstream of 13% of the lateral confluence predictions within the Coastal and Interior regions as compared to only 3% of the lateral predictions in the other three regions. This trend became more apparent when the model was applied to the two additional "poor performing" coastal basins. Fish were observed upstream of 44% of the lateral predictions in these two basin, and lateral predictions on average underestimated fish use by 48 feet. Errors for four of the lateral confluence predictions exceeded 400 feet.

Table 5. Net field error distances (feet) of predicted end of fish points by boundary types for 15 randomly selected basins in five geographic regions in western Washington. Only predicted EOFPs associated with defined channels are included.

		Region					
Last Fish Habitat Point Boundary Type		**Coastal**	**Interior Olympic**	**Puget Sound**	**South Cascades**	**North Cascades**	**Total**
Mid-channel and Tributary Junction	No. Points	12	9	7	9	4	41
	Mean	-667	-456	-2908	-1174	-991	-1146
	Median	189	-149	-2798	-150	-918	-270
Lateral Confluence	No. Points	48	62	37	35	4	186
	Mean	47	42	0	9	0	28
	Median	0	0	0	0	0	0
Total	No. Points	60	71	44	44	8	227
	Mean	-96	-21	-463	-233	-496	-185
	Median	0	0	0	0	0	0

<u>Precision and Balance of Prediction at the Basin Scale</u>

The success of the water type model in accurately predicting the presence of fish habitat along the entire stream network of sample basins was examined as part of the Pilot Study. The overall classification success of the model was partitioned by deriving confusion matrices (Fielding and Bell 1997). A confusion matrix tabulates the observed and predicted presence/absence patterns and thus provides a summary of the number and direction of correct and incorrect classifications produced by the model (Table 6). First, the overall classification performance of the model is quantified as the percentage of cells where the model correctly predicts the presence/absence of fish (referred to as overall correct classification calculated as length of correct presence/absence predictions / total length of stream length classified). Second, the ability of the model to correctly predict fish presence was examined (referred to as model sensitivity rate and calculated as length of stream correctly classified as fish habitat / total length of stream predicted to contain fish habitat). Third, the ability of the model to correctly predict fish absence was determined (referred to as model specificity and calculated as length of stream correctly classified as non-fish habitat / total length of stream predicted to contain no fish). Table 6 provides an assessment of the balance of errors (i.e., model sensitivity and specificity rates) and describes the overall correct classification, sensitivity, and specificity rates.

On average, the model performed well when applied to the sample fourth order sub-basins. The model correctly predicted fish presence or absence in over 90% of stream miles depicted by the DEM Hydrolayer. The correct classification dropped slightly (down to 88%) when adjusted for non-defined channels being depicted, but model sensitivity and specificity rates remained unchanged. Likewise, the model correctly predicted fish absence (96% specificity rate, or correct absence predictions). That is, fish were observed in 4% of the stream length that was predicted not to contain fish habitat. However, the percent of correct presence predictions was considerably lower, with only 64% of the stream length predicted as fish habitat actually being inhabited by fish. In other words, the model routinely predicted fish habitat beyond the upstream extent of fish use. Conversely, the model less commonly predicted non-fish habitat on stream channels where fish were observed in the field.

Similar to the performance measures at the EOFP scale presented previously, the basin scale classification accuracy results illustrate that overestimates of fish use are more common than underestimates. Overestimates are less frequent in the Coastal and Interior Olympic Region, especially in basins where geomorphic conditions are not conducive to development of steep, bedrock-dominated reaches and waterfalls.

Use of the basin scale model performance measures as described above assumes that the DEM-generated stream network accurately represents the headward extent of the stream network. However, inaccuracies and misrepresentations of streams derived through DEM analysis make it difficult to determine the true lineal extent of stream channels in most sample basins, even with revisions made to account for the false portrayal of stream channels associated with lateral confluence EOFP predictions. It is not uncommon for DEM generated stream networks to portray stream channels where none actually exist in the field. Yet other situations occur where stream channels observed in the field are not included in the DEM derived stream network. Thus, model performance measures relying on DEM estimates of total stream length clearly contain inaccuracies that can only be rectified by verifying the length and headward extent of all mapped streams within each sample basin.

Given that mapping inaccuracies of the headward extent of the channel network mostly influences the classification of non-fish-bearing channels, we also calculated the model prediction accuracy for only the fish-bearing portion of the sample basins (includes the total stream length situated downstream of observed EOFPs). Although there may be minor discrepancies between the mapped and field observed stream lengths, we our confident that there is a readily apparent defined channel downstream of all field-verified EOFPs. On average, the model correctly predicted fish presence in 93% of the total length of confirmed fish bearing streams (Table 7). Similar to the basin scale performance measures, prediction errors were largely skewed towards overestimation of fish habitat, as the model on average overestimated the total length of fish habitat by 243%, although the overestimation values varied widely across sample FOSBs.

Model Performance Summary

A field model performance assessment was conducted on 369 predicted EOFPs within 15 randomly selected FOSBs. Mean net and absolute prediction field error distances were estimated as -114 $\pm$ 78 feet and 176 $\pm$ 60 feet, respectively. Eighty five percent of all of the DEM-portrayed predicted EOFPs were correct (ΔEOFP = 0). When only predicted EOFPs associated with defined channels were included in the computation of model performance, a total of 158 of the 369 EOFP predictions were excluded from analysis and the mean performance measures decline to -185 $\pm$ 129 feet and 287 $\pm$ 94 feet for net and absolute error, respectively. Seventy five percent of the predicted EOFPs associated with defined channels were correct. The model correctly classified fish absence or presence in over 90% of the DEM-portrayed stream network.

The majority of model EOFP predictions are situated at the mouths of tributary channels (lateral confluence EOFP type). Over 90% of lateral confluence predictions investigated were correct. Terminal predictions accounted for only 11% of the predicted points, yet the terminal predictions accounted for 92% of the total error distance. Net error in terminal predictions averaged -1146 $\pm$ 666 feet. Six terminal predictions situated upstream steep channel features readily apparent on the DEM generated channel profiles accounted for 46% of the total error measured.

The cluster sampling design used in this study surveys a representative allocation of EOFP boundary types and provides a more accurate and unbiased characterization of model performance then afforded by the withheld training data (Conrad et al. 2003). Because of the under representation of lateral confluence predictions in the withheld data set, only 25% of the predictions were correct as compared to 75% in the pilot study. However, care must be taken in interpreting these results. While the sample FOSBs were randomly selected under the pilot, it should be noted that all streams under Forest and Fish jurisdiction were not included in the sampling frame, nor were basin sizes taken into account. We restricted our sampling frame to basins solely under the jurisdictions of Forest and Fish Rules, and therefore excluded basins that are a mosaic of state, federal, private, and tribal ownership. Many FOSBs have a mix of ownership under Forest and Fish rules (i.e., state and private) and other rules (i.e., National Forests and National Parks). These basins are often situated in steeper, more dissected landscapes in which the DEM more accurately represents the drainage network, thereby producing a model that would conceptually predict fish habitat more accurately. Moreover, the sample selection process used under the pilot study did not account for basin or geographic region size and, as such, the pilot sample may not accurately represent the EOFP conditions as they occur across the landscape.

Table 6. Fish presence / absence classification accuracy for 15 fourth order sub-basins randomly selected for sample across five geographic regions of western Washington. The *Not Revised* classification includes stream length where no defined channels were evident in the field despite being depicted on the DEM stream network. The *Revised* classification excludes the non-defined channels in the classification accuracy calculations. Fish were effectively precluded from use of the Grays, Run, and Bush tributaries by a large waterfall at the mouth of the FOSB.

	Not Revised						*Revised*					
Basin	**Correct Classification**	**Correct Presence Calls**	**Correct Absence Calls**	**Error (ft) per mile of total stream length**	**Overestimate (ft) of fish habitat per mile of total stream length**	**Underestimate (ft) of fish habitat per mile of total stream length**	**Correct Classification**	**Correct Presence Calls**	**Correct Absence Calls**	**Error (ft) per mile of total stream length**	**Overestimate (ft) of fish habitat per mile of total stream length**	**Underestimate (ft) of fish habitat per mile of total stream length**
North River	0.89	0.91	0.87	599	184	415	0.87	0.91	0.85	661	203	457
Elkhorn	0.87	1.00	0.80	708	0	708	0.86	1.00	0.78	744	0	744
Grays	0.80	0.00	1.00	1049	1049	0	0.80	0.00	1.00	1051	1051	0
Coastal	**0.85**	**0.64**	**0.89**	**786**	**411**	**374**	**0.85**	**0.64**	**0.88**	**818**	**418**	**400**
Lentz	0.97	0.91	1.00	135	116	19	0.97	0.91	0.99	143	123	20
Newman	0.93	1.00	0.89	361	0	361	0.93	1.00	0.88	381	0	381
Pysht	0.92	0.67	1.00	399	399	0	0.75	0.67	1.00	1304	1304	0
Interior Olympics	**0.94**	**0.86**	**0.96**	**298**	**172**	**127**	**0.88**	**0.86**	**0.96**	**610**	**476**	**134**
Run	0.75	0.08	1.00	1295	1295	0	0.75	0.08	1.00	1295	1295	0
Skookumchuk	0.92	0.64	1.00	429	429	0	0.91	0.64	1.00	452	452	0
Thorndike	0.95	0.81	1.00	238	238	0	0.92	0.81	1.00	421	421	0
Puget Sound	**0.88**	**0.51**	**1.00**	**654**	**654**	**0**	**0.86**	**0.51**	**1.00**	**723**	**723**	**0**
Bush	0.81	0.05	1.00	988	988	0	0.81	0.05	1.00	988	988	0
NF Green	0.96	0.98	0.95	228	19	209	0.94	0.98	0.92	342	29	314
Nineteen	0.91	0.73	0.98	472	404	68	0.91	0.73	0.98	495	424	71
South Cascades	**0.89**	**0.59**	**0.98**	**563**	**470**	**92**	**0.88**	**0.59**	**0.97**	**609**	**480**	**128**
Bear	0.97	1.00	0.97	145	0	145	0.95	1.00	0.93	249	0	249
Deer	0.93	0.57	1.00	348	348	0	0.89	0.57	1.00	589	589	0
MF Snoq	0.91	0.28	1.00	452	452	0	0.88	0.28	1.00	651	651	0
North Cascades	**0.94**	**0.62**	**0.99**	**315**	**267**	**48**	**0.91**	**0.62**	**0.98**	**496**	**413**	**83**
Grand Average	**0.90**	**0.64**	**0.96**	**523**	**395**	**128**	**0.88**	**0.64**	**0.96**	**651**	**502**	**149**

Table 7. Fish presence classification accuracy of the portion of the stream network inhabited by fish in 15 fourth order sub-basins randomly selected for sample across five geographic regions of western Washington. Stream lengths are presented in miles.

Basin	Total length field verified as fish present	Total length modeled as fish habitat	Correct classification of fish bearing channels	Total length of field verified fish absence incorrectly predicted as fish habitat	Proportional overestimate of fish habitat relative to length of confirmed fish bearing stream
Elkhorn	2.11	1.52	72%	0.00	0.00
Grays	0.00	1.86	100%	1.86	-
North River	4.10	3.69	82%	0.33	0.08
Coastal	**2.07**	**2.36**	**85%**	**0.73**	**0.04**
Lentz	1.40	1.51	98%	0.13	0.09
Newman	5.88	5.00	85%	0.00	0.00
Pysht	1.40	2.10	100%	0.70	0.50
Interior Olympics	**2.90**	**2.87**	**94%**	**0.28**	**0.20**
Run	0.19	2.41	100%	2.21	11.52
Skookumchuk	1.24	1.94	100%	0.70	0.56
Thorndike	3.53	4.39	100%	0.85	0.24
Puget Sound	**1.65**	**2.91**	**100%**	**1.25**	**4.11**
Bush	0.07	1.35	100%	1.27	17.39
NF Green	1.71	1.44	83%	0.03	0.02
Nineteen	2.92	3.74	94%	0.99	0.34
South Cascades	**1.57**	**2.18**	**92%**	**0.76**	**5.92**
Bear	1.32	1.15	87%	0.00	0.00
Deer	0.27	0.47	100%	0.20	0.75
MF Snoq	0.18	0.64	100%	0.46	2.52
North Cascades	**0.59**	**0.75**	**96%**	**0.22**	**1.09**
Grand Average	**1.76**	**2.21**	**93%**	**0.65**	**2.43**

Sample Size Requirements for Further Study

Pilot results provide the opportunity to assess sample size requirements to achieve a range of confidence bounds on the estimate of the "true" error of the model predictions. The maximum sampling error that the State is willing to tolerate affects the number of sample basins ultimately surveyed under the Study Plan.

By surveying all of the predicted EOFPs within each sample basin, the study will attain a representative allocation of EOFP boundary types. Thus, as exemplified by the pilot results, overall model performance (over 75% of the EOFP accurately depicting the upstream end of fish with an overall correct classification rate of 93% for fish-bearing channels) will appear especially robust. This is due to the large proportion of lateral confluence predictions, which have less variability (small variance). In addition, a large proportion of the drainage network is comprised of small, and often steep, headwater channels typically determined correctly as non-fish habitat.

Inclusion of the lateral EOFPs into the assessment provides a complete view of model performance. However, by including the lateral predictions into the computation of sample size requirements, an under sampling of terminal predictions could result. The lateral confluence EOFPs, which far outnumber the mid-channel and tributary junction EOFPs (collectively referred to as terminal EOFPs), have less variability (small variance) and therefore would require a much smaller sample size to attain the same level of precision (Table 8). It is generally accepted that sample size requirements should be based on the performance indicator that exhibits the most variability (larger variance).

Table 8. Summary statistics for net and absolute field error distance (feet) by end of fish point boundary types. Bounds on the error of estimation are calculated following cluster sampling procedures described by Scheaffer et al. (1996) with a confidence coefficient of at least 0.75 regardless of sampling distributions. The confidence coefficient approaches 0.95 as the sampling distribution approaches normality. The sample size requirements represent the number of cluster (sample basins) needed to estimate the mean prediction error with the specified bounds on the error of estimation.

			Net Field Error						Absolute Field Error					
					Target bounds for error of estimate						Target bounds for error of estimate			
Last Fish Point Boundary Type	Number of Points	Points Per FOSB	Pilot Mean	Min 75% Confidence Bound	50	100	200	300	Mean	Min 75% Confidence Bound	50	100	200	300
Only predicted EOFPs associated with defined channels					*Estimated sample size required*						*Estimated sample size required*			
Mid-channel and Tributary Junction (15% Trimmed)	35	3	-534	348	581	145	36	16	910	235	264	66	16	7
Mid-channel and Tributary Junction	41	3	-1146	678	2760	690	173	77	1466	588	2076	519	130	58
Lateral	186	12	27	17	2	0	0	0	27	17	2	0	0	0
Lateral and Terminals Combined	227	15	-185	129	105	26	7	4	287	94	54	14	3	1

The Study Plan uses procedures outlined by Scheaffer et al. (1996) for determining appropriate sample size (number of clusters or sample basins) needed to estimate the mean ΔEOFP with a desired level of precision. These procedures are especially useful for these data, as they do not rely on a sampling value being taken from a normal distribution. From both this pilot data and the withheld data set used by the ISAG Statistical Team, it is readily apparent that the bulk of the ΔEOFP associated with the lateral confluence boundary types are zero and the distribution of all of the ΔEOFP values is skewed and leptokurtic. Error distance for over 87% of the EOFP predictions are zero, and no data transformations can effectively normalize the distribution when laterals are included. All the preceding discussions on model performance in the previous sections and the sample size estimates have been based on the use of non-transformed data.

However, a variety of data transformations can be used to improve normality of the frequency distribution for terminal error predictions, each resulting in slightly different estimates of average model performance. For instance, a log (x + 1) transformation of the absolute field error distances was used to improve normality for the 41 terminal predictions. By using the transformed data to calculate means and confidence bounds and presenting the back transformed estimates, the pilot data for the 15 basins estimates the mean absolute prediction error of terminal points as 722 $\pm$ 5 feet. The justification for transforming the data becomes more apparent if the data are used for hypothesis testing via parametric statistical methods such as those described in the Study Plan to investigate factors that contribute to model error. Even when analytical methods are robust to violations of assumptions, it is possible to reduce the violations by appropriately transforming the original data.

In summary, reasonable precision levels (e.g., bounds on error estimates targeted at 100 feet) for net prediction error can be attained with as few as 26 sample basins when all EOFP types (lateral and terminal EOFPs) are examined. If only non-transformed, terminal EOFPs are included in the calculations, at least 77 basins are needed to achieve a bound on the estimation of net prediction error of 300 feet. However, if only terminal EOFPs are considered, then data transformations to achieve normality would substantially reduce sample size requirements.

Further Study Implications

The pilot study includes a representative allocation of EOFP boundary types and provides a more accurate and unbiased characterization of model performance then afforded by the dataset assessed by Conrad et al. (2003). However, because of limitations placed on the sampling frame, the pilot sample may not accurately represent the EOFP conditions as they occur across the landscape.

Using the stratified cluster sample design with probability proportional to size approach described under the validation Study Plan, all basins and attendant predicted EOFPs under the Forest and Fish rules would have a chance of being selected during the sampling process, regardless of basin size. In addition, any higher order stream channel situated between two identified sample basins will be included as part of the potential sample basin situated upstream. Therefore, the sampled population will represent the distribution of EOFP boundary types as they occur on the landscape. Using this approach, we can compute an overall estimated mean, median, and bounds on the prediction error estimates. This unbiased and representative approach to sampling will provide data on which to explore the factors that affect model performance as described under the Study Plan. As the pilot data suggest, further examination of model sensitivity to channel gradient (especially in larger streams) and differential model performance across the geographic / geologic regions may be warranted.

References

Cochran, W. G. 1977. Sampling techniques. Third edition. John Wiley and Sons, Inc., New York, NY.

Conrad, R.H., B. Fransen, S. Duke, M. Liermann, and S. Needham. 2003. The development and assessment of the preliminary model for identifying fish habitat in western Washington. Draft report prepared for Washington Dept. Natural Resources Instream Scientific Advisory Group. 150 pp.

Fielding, A. H., and J. F. Bell. 1997, A review of methods for the assessment of prediction errors in conservation presence/absence models: Environmental Conservation. 24,38–49.

Scheaffer, R. L., W. Mendenhall, and L. Ott. 1996. Elementary survey sampling. Fifth edition. PWS-KENT Publishing Company, Boston, MA.

Thompson, S. K. 1992. Sampling. John Wiley and Sons, Inc., New York, NY.