The Development and Assessment of the Preliminary Model for Identifying Fish Habitat in Western Washington

Prepared by:
Robert H. Conrad – Northwest Indian Fisheries Commission
Brian Fransen – Weyerhaeuser Company
Steven Duke – Weyerhaeuser Company
Martin Liermann – National Marine Fisheries Service
Scott Needham – EarthRes.i

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Doug Sutherland - Commissioner of Public Lands

Cooperative Monitoring
Evaluation & Research
CMER 03-313

The Development and Assessment of the Preliminary Model for Identifying Fish Habitat in Western Washington

Prepared by:
Robert H. Conrad - Northwest Indian Fisheries Commission
Brian Fransen - Weyerhaeuser Company
Steven Duke - Weyerhaeuser Company
Martin Liermann - National Marine Fisheries Service
Scott Needham - EarthRes.i

September 2003

PAGE LEFT BLANK INTENTIONALLY

EXECUTIVE SUMMARY

New Forest Practices regulations were enacted as a result of a series of stakeholder negotiations which led to the Forest and Fish Report (FFR) in 1997. Changes to the Forest Practices regulations that were outlined in the Forest and Fish Report included direction to develop a model to classify fish-bearing waters using data available on a geographical information system (GIS). The model was to identify potential fish habitat and have an equal likelihood of over and under-classification of fish habitat. In addition, a performance target for model accuracy was adopted. Based on previous research, this habitat-based, water-typing model was to be developed using logistic regression analysis and GIS data which incorporated the results of field surveys that had been previously conducted.

The fish absent | fish present (FAFP) data used to estimate the logistic regression models for this project were all generated from end-of-fish points (EOFPs) placed on a Washington Dept. of Natural Resources (DNR) GIS hydraulic layer. Each EOFP was based on a field survey which followed specific protocols to identify a location on the stream that was designated as either last fish or last fish habitat. Potential EOFPs were submitted to DNR for error checking and initial screening. After approval by DNR, the EOFP was transferred from the DNR hydraulic layer to a 10m DEM-generated stream points network. An automated procedure was then used to classify points upstream of the EOFP as fish absent points and points downstream of the EOFP as fish present points. There were five physical attributes associated with each point on the DEM network: basin size, elevation, upstream gradient, downstream gradient, and mean annual precipitation.

EOFPs were submitted by 21 different groups which included state and tribal natural resource departments, timber companies, and environmental organizations. The EOFPs submitted to DNR were essentially "found" data meaning that collectively the data werenot gathered under a single, statistically-based sampling plan. The EOFPs did not represent a random sample of the stream systems in western Washington but instead tended to be concentrated in a few geographic locations. In total, 4,052 EOFPs were approved and used to generate 1,443,471 FAFP data points in the western Washington DEM network. These were the data available for our analyses.

The majority of the end-of-fish points were associated with smaller sub-basins at the upper end of stream systems. The physical attributes (basin size, elevation, gradient, and precipitation) associated with the FAFP points in these upstream reaches were more informative for building a logistic regression model than the physical attributes of lower mainstem FAFP points. Mainstem FAFP points were almost always Fish Present points and were typically characterized by very large basin sizes (relative to the other FAFP points), low elevations, and low gradients. These mainstem FAFP points could heavily influence the model building process and the assessment of model performance since they were rarely misclassified. An objective method of limiting the data used to estimate any logistic regression model (LRM) to those points associated with the upper ends of streams was needed as these data should be the most informative. Therefore, we divided the FAFP data into smaller subsets that were approximately defined by the DEM-generated stream network Strahler fourth order sub-basin (FOSB).

There were 867 FOSBs which contained at least one EOFP present in the 1.4 million FAFP points. For logistic regression model building, only FAFP data generated from EOFPs with end types 1 or 2 were used. End types 1 and 2 were "natural" end types, i.e., the reason identified for the change from fish presence to fish absence on the stream was that the stream became too small or the gradient too steep. The remaining EOFP end types were considered "unnatural" end types, i.e., the location of these EOFPs was unlikely to be related to the values of the five physical attributes (e.g., because of a dam). Because the data set that was available for model building was extremely large, a split sample validation method was used to develop the logistic regression models and evaluate their performance.

The FAFP data were partitioned into three groups: (1) data from FOSBs that contained EOFPs with end types 1 or 2 only (model building data pool); (2) data from FOSBs which included at least one EOFP with an end type other than 1 or 2 (model validation data); and (3) data that were not part of a fourth order sub-basin (unassigned data).

The data available for model building were sub-sampled because we wanted the data set used for model estimation to have WRIAs represented proportionally to their contribution to the western Washington DEM-generated stream points network and also wanted to use a relatively small fraction of the FAFP points available to reduce statistical dependency among the data. A range of possible sample sizes was examined for the number of fish absent and fish present points to draw from the model building data pool to estimate the "final" logistic regression model. These analyses indicated that 4,000 points from each group provided stable estimates of the model coefficients. The final model included four of the five GIS physical attributes available (upstream gradient was not included).

When the model validation data were classified with the final logistic regression model, the means of the classification accuracies for the western Washington WRIAs with data were: 93% for fish absent points in FOSBs; 90% for fish present points in FOSBs; 91% for fish absent points in the unassigned group; and 99% for fish present points in the unassigned group.

Any assessment of logistic regression model performance should not be limited exclusively to classification accuracy based on known FAFP points in the western Washington 10m DEM network. Overall classification accuracy based on the available data is likely biased because the data are not sampled representatively across factors such as end type, boundary type, WRIA, and basin size. Other metrics of model performance are less susceptible to this problem, but are no substitute for a more thoughtfully sampled validation data set.

Classifying DEM network points of unknown fish status with the final logistic regression model does not result in a prediction of an end-of-fish point on the stream system. A procedure was developed that processed the strings of FAFP data representing streams and placed an EOFP in each string. The procedure processed a string of FAFP data from the downstream end and progressed upstream until a set of conditions were met at which point it "stopped" and placed an end-of-fish point in the string. Therefore, we called this the stopping rule. Similarly to the development of the logistic regression, a portion of the data was used to develop the stopping rule and a portion was reserved for validation of the stopping rule.

The stopping rule had three parameters: a cut point, a trigger size, and an upstream habitat block size. The cut point was a value between 0 and 1 which classified the LRM probabilities into predictions of fish absence and fish presence. The trigger was the size of the block of consecutive predicted fish absent points which indicated a possible EOFP. Upstream habitat block size was the size of a block of potential fish habitat upstream from a trigger which overrode the trigger. We used a grid search technique with a constraint for balance to find the stopping rule which minimized error while achieving a degree of balance on the strings. Balance was defined as the errors in the upstream direction being approximately equal to the errors in the downstream direction.

Error was defined as the distance (stream distance in ft) between a predicted EOFP and the actual EOFP on the 10m DEM network. Negative error occurred when the predicted EOFP was upstream of the true EOFP. Stopping rules examined during the grid search were applied to the modeling strings, and the median absolute error of each rule was computed. For the set of rules that met our balance constraint (-50 ft > 10% trimmed mean error > 50 ft), the rule with the minimum median absolute error was chosen as optimal.

Using the 2,008 strings reserved for stopping rule development, the optimal rule identified had a cut point of 0.4, trigger size of 1, and upstream habitat block size of 85. The median absolute error for this rule based on the modeling data strings was 367 ft.

This stopping rule was then applied to the 2,042 validation data strings. The distribution of errors for the validation data was centered at zero (32% of the errors were zero), the distribution of errors was not symmetrical but skewed as the model tended to place the predicted EOFP upstream of the true EOFP, and there were a number of outliers (some of which were extremely large). The median absolute error for the validation data strings was 327 ft and mean error was -140 ft. The 10% trimmed mean error (which removed 10% of the values from either end of the error distribution) was 37 ft.

The string form of the stream network data was convenient for implementing the optimization process but it had one major drawback; the assessment of model performance based on strings probably does not correctly characterize performance when the model is applied to the 10m DEM network. Specifically, the estimates of model precision (total error) and balance based on strings probably does not accurately represent what occurs on the stream network. There are two reasons for this. The first is that strings can overlap so that sections of a string that are in error can overlap with error sections in other strings and be counted two or more times. This multiple counting results in an overestimate of the total error associated with each EOFP prediction. Secondly, strings tend to overestimate total error because the string data contain a smaller proportion of lateral confluence EOFPs than are found in the stream network. Lateral confluences are side tributaries entering a main channel and are frequently associated with a gradient break and/or a large change in basin area. The LRM and stopping rule procedure predicted the EOFPs associated with lateral confluences very accurately (68% of the time they were identified with no error).

Therefore, we examined model prediction error by two other methods that avoided these problems but had other limitations. There were data available from one basin (Stillman Creek) which had been completely surveyed. We applied the model and stopping rule to the data from this basin and estimated the precision and balance for model error based upon the entire basin. We also developed a method for approximating the results that would be obtained by entire sub-basin surveys using existing data in fourth order sub-basins that had eight or more survey points. This method used a procedure to randomly assign errors to predicted EOFPs for which there were not survey data available.

The mean absolute error estimate from the FOSB random error assignment process was 445 ft/mi compared to 297 ft/mi for the Stillman Creek basin assessment. These mean absolute error estimates are analogous to correctly classifying about 92% of the points in the FOSBs and about 94% of the points in the Stillman basin. However, these assessments are based on a very limited sample of the western Washington stream network data.

The FOSB and Stillman basin assessments of balance in the model errors, however, were not similar. The FOSB error assessment process estimated a mean net under-prediction of 170 ft/mile. For the Stillman Creek basin assessment, however, there was a mean net over-prediction of -226 ft/mile.

The Fish Habitat model requested in the Forest and Fish Report is complete. The preliminary model described in this report (final logistic regression equation and stopping rule) and resulting assessments represent the best possible science given the limitations of the available data and guidelines specified in the Forest and Fish Report.

We recommend caution when interpreting the model results due to: (1) the unknown variability in the field survey protocols that were used to establish the "end-of-fish" points for this study; (2) the non-random nature in which the data that were the basis for the analyses were collected; and (3) the methods used to estimate the precision and balance of the model predictions were not ideal because entire basins were not covered by field surveys. The "found data" that were used for this project presented problems in model development and assessment that cannot be addressed at this time.

We are reasonably confident that, on average, the distance between a predicted EOFP and the "true" EOFP will be less than 500 ft. The assessment of precision for fourth order sub-basins (±445 ft per mile of stream classified) is likely an over-estimate of the error. The overall error would be smaller (<445 ft per mile) if larger order streams (e.g., mainstem channels) were included in the data used to make the estimate.

We are <u>not</u> confident in the assessment of the balance of model errors because of the conflicting results from the FOSB error assessment (+170 ft per mile of stream) and Stillman basin error assessment (-226 ft per mile of stream). More complete basin survey data similar to that from Stillman Creek are needed to more accurately assess the balance of model errors. Additional complete basin survey data should be collected to ensure that there is a broad representation of the western Washington landscape.

Recommendations

We recommend that the preliminary logistic regression model and stopping rule be implemented with consideration of the following:

- Where site-specific field survey information exists, it should be considered in addition to the model predictions.
- More field surveys need to be conducted to provide a more reliable assessment of model performance and to update and improve the model.
- Alternative data sources and statistical modeling approaches should be explored to improve and streamline future stream-typing models.
- Situations that might not be appropriate for inclusion in the next generation of modeling data should be reviewed and a protocol for exclusion of data developed.

PAGE LEFT BLANK INTENTIONALLY

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

INTRODUCTION

Washington State forest practices regulations for the protection of water bodies containing fish habitat are intended to provide for the maintenance and/or restoration of healthy fish populations and their associated habitats while minimizing economic impacts and operational restrictions on those subject to the regulations. Under these regulations, streams, lakes, and ponds with fish habitat receive different protections than those water bodies without fish habitat. For example, riparian buffer requirements adjacent to waters with fish habitat are tailored to provide riparian functions and the input of organic materials necessary to naturally maintain productive habitat. Riparian buffers adjacent to fish habitat include large areas of productive forestland. Thus, it is important that buffer prescriptions are efficiently and accurately applied to affected water bodies and provide the intended protection to fish habitat while minimizing unnecessary impacts to landowners. Accurate identification of fish habitat also is important in determining where fish passage structures are required at road crossings. Passage structures that provide unrestricted passage for fish are often expensive to install. It is possible to protect the ecological function of streams without fish habitat with less expensive drainage structures. Balancing resource protection and timber management considerations requires accurate methods of identifying fish habitat.

Fish habitat can be identified by either directly surveying a water body to determine fish presence or absence or by inference based upon the physical characteristics of the habitat. Surveys usually are conducted using electro-shocking equipment, which is a very efficient and effective method of capturing fish. Inference of fish habitat from factors such as the size and slope of the water body rely on previously established empirical relationships between habitat containing fish and associated physical characteristics of the water body. Each method has limitations and advantages.

Electro-fishing surveys use an electro-shocker to apply an electric current to the water which attracts and temporarily stuns fish. These surveys provide a very reliable method for verifying fish presence or absence and are widely employed. However, performing these surveys is very expensive and time consuming. Years of survey effort would be required to classify all waters of the state by direct survey. Fish may not occupy all available habitat at all times, leading to potential survey classification errors if fish are not occupying the habitat at the time of the survey. Surveys performed in areas where populations are at low abundance may fail to detect fish. Some water bodies are difficult or impossible to survey due to their size, complexity, or accessibility. The presence of fish during a survey is conclusive evidence of fish habitat presence, although absence of fish does not necessarily demonstrate the absence of fish habitat.

Inference of the presence or absence of fish habitat from the physical characteristics of the water body presents different advantages and disadvantages. As this approach relies on characteristics such as the size or gradient of the water body to determine suitability for fish, temporarily unoccupied habitat can be classified. Using widely available topographic information, opportunities exist for rapid classification of fish habitat across large areas using computer-based geographical information systems (GIS) or other tools which can greatly reduce the costs compared to on-the-ground surveys. However, inference of fish habitat may be less reliable than direct survey of fish presence, leading to higher classification error and potential bias in over or under classification. Factors other than physical characteristics may influence habitat suitability for fish. Above migration barriers, physical habitat may be consistent with suitable fish habitat. Whether or not fish are present within the habitat above natural barriers can not be reliably determined by physical characteristics alone.

The shortcomings of both methods indicate the need for more reliable and efficient approaches to classifying water bodies. Ideally, the classification system would:

- Include a spatially explicit framework to communicate, incorporate, and archive stream classification information,
- Be sufficiently accurate to ensure that protection is applied where it will provide benefits to fish and their habitats as intended,
- Not be excessively over or under-inclusive which could subject landowners to unnecessary economic hardship or fail to provide protection to public resources,
- Allow water bodies to be classified rapidly across large areas to facilitate accurate, efficient, and uniform application of the forest practices regulations,
- Be capable of classifying all waters of the state governed by the regulations.

These needs have led to the development of a GIS-based fish habitat model. This model is intended to be used for the identification of fish habitat across Washington State, with a goal of reducing or eliminating the need for on-the-ground surveys as the only method of identifying fish habitat. This report provides background information on why this approach was selected and documents the process employed to produce and assess the prescribed model. Building on the work of others, field survey information was used to derive a preliminary empirical model to predict the location of the upper extent of fish habitat in streams in western Washington. The resulting model was assessed against independent field survey data to provide a characterization of error, error balance, and the influence of regional and other factors.

BACKGROUND

In the 1970s, the first forest practices regulations intended to provide for the protection of public resources on managed forestlands were adopted in Washington State. Regulations intended for the protection of aquatic resources relied on a map-based framework that assigned classifications to water bodies covered by the new regulations. The map framework was managed by the Washington Department of Natural Resources (DNR), and provided the spatial framework for rules governing road construction, timber harvest, and chemical application near water bodies. The classification system assigned a "type" to water bodies that determined requirements for public resource protection including management of the adjacent forestland. Stream type categories were based on the beneficial uses of public waters, one of which was use by fish.

When the first maps were generated, the identification of fish-bearing water bodies was based on extremely limited knowledge of the upper limits of fish distribution. The presence of fish in the majority of headwater streams across the state remained largely unverified because the focus for fish protection at that time was primarily on the commercially or recreationally important fish species occurring farther downstream. In water bodies with unknown fish presence status, landowners and resource professionals worked together to assign fish-bearing status to most streams after cursory review of maps and aerial photos or from physical characteristics observed at the site. Professional judgment was used to assign the classification to most water bodies. Following the initial round of classification, maps and default physical criteria presuming fish presence were adjusted periodically as more reliable information became available.

As several fish populations were proposed for listing under the Endangered Species Act in the early 1990s, interest in the protection of fish and fish habitat increased. Funding became available to evaluate fish protection measures and potential causes for declines in fish abundance. State agencies and tribal organizations initiated programs to evaluate the accuracy of the existing classification maps. Technical reports produced by the Quinault Indian Nation (Baxter and Mobbs 1992; Baxter 1993; and Mobbs et. al. 1995) and the Point No Point Treaty Council (Bahles and Ereth 1994) documented a high level of inaccuracy in the existing maps. Their findings suggested that more than 70% of the boundaries between fish-bearing and non fish-bearing streams were incorrectly identified, with nearly all the errors resulting in significant under-classification (placement too far downstream) of the actual upper extent of fish-bearing streams. Due to the prevalence of under-classified streams, a strong incentive was created for tribal groups, agencies, and landowners to expand survey programs and initiate revisions to the classification maps. As a result, survey crews were deployed across a large portion of the lands covered by the Forest Practices regulations to correct classification errors in fish-bearing water bodies.

The Forest Practices Board adopted a temporary Emergency Rule for water typing in 1996 which provided an interim solution until a more permanent, science-based rule could be developed (FPB 1996a). To immediately address the well-documented and widespread under-classification problem, the new rule called for a presumption of fish presence on streams greater than two feet wide and less than a 20% gradient. The default criteria were based on limited field data collected at the upper limit of fish distribution. During rule making, the regulators intended to derive reliable default physical criteria for presumption of fish presence consistent with a " shared risk " approach. A reduction in electro-fishing survey effort was agreed to as a goal of the new rule due to concerns about survey impact to the fish and the high cost of extensive survey effort. If successfully achieved, accurate and balanced default criteria should present an equal likelihood of over and under-classifying the actual extent of fish-bearing streams, thereby reducing the incentive for field surveys. However, uncertainties surrounding the reliability of the default criteria resulted in the retention of the opportunity for interested parties to verify fish presence or absence by performing field surveys. A standard protocol for verifying fish absence using field surveys was developed (FPB 1996b).

Surveys performed following the adoption of the Emergency Rule found that the default physical criteria for presumption of fish presence over-estimated the actual extent of fish-bearing streams in many areas. The resulting misclassification of non fish-bearing streams as fish-bearing created an incentive for landowners to perform more surveys. Large-scale survey efforts, primarily sponsored by landowners, were initiated to increase the classification accuracy for fish-bearing streams. Agency and tribal organizations were forced to allocate scarce resources to handle the large volume of landowner survey information submitted for their review and concurrence. All stakeholder groups involved in performing surveys or reviewing the survey results were eager to explore alternatives that would reduce the need for extensive survey effort.

To develop the permanent rule to replace the Emergency Rule, an ad hoc Water Typing Committee was formed within the Timber, Fish, and Wildlife (TFW) program. TFW was a multi-stakeholder program intended to develop science-based recommendations for changes in Forest Practices regulations and resolve conflicts among stakeholders with varying interests in the management of forested landscapes. Participants in the committee included technical representatives of government agencies, tribal groups, environmental organizations, and landowner representatives.

This process occurred during a time of significant and rapid improvement in technical information and software tools. As a result of the extensive fish surveys being performed, abundant field survey information was available for many areas of the state. Advances in GIS technology provided opportunities to evaluate resource protection and economic performance of alternative water typing systems across large geographic areas. Digital Elevation Models (DEM) produced by the U. S. Geologic Survey became widely available, allowing for consistent and reliable characterization of the physical landscape. For the first time since the implementation of forest practice regulations governing fish-bearing water bodies in the 1970s, the tools and data were available to develop and assess a data-driven classification system for use across the entire state.

Previous Studies Predicting Fish Absence|Presence Using Logistic Regression Analysis

During the mid 1990s, studies performed by other researchers demonstrated the utility of empirically-derived models using physical habitat characteristics to predict the presence or absence of fish. These models relied primarily on the relationships between fish habitat suitability and the size, slope, and elevation of streams. Often the statistical model used for these studies was logistic regression analysis. Logistic regression analysis is a statistical procedure that has been widely used to develop models to predict the absence|presence of salmonids in stream and river systems based upon site or reach specific biotic and physical characteristics. Some of these studies are summarized below.

Rieman and McIntyre (1995) used logistic regression models (LRMs) to examine the relationship between the absence|presence of bull trout (*Salvelinus confluentus*) in Pacific Northwest streams and three physical characteristics measured on site: stream width, channel gradient, and the estimated size of potential bull trout habitat present. Stream width and habitat patch size were found to be significant factors in the models. They did not estimate the predictive classification accuracy of the LRMs they developed.

Watson and Hillman (1997) also used LRMs to examine the relationship between 16 different physical and biotic characteristics measured at the site level and the occurrence of bull trout in Pacific Northwest streams. They developed significant LRMs relating the absence|presence of bull trout in different habitat types to a suite of different physical and biotic characteristics. The predictive abilities of these LRMs were not assessed.

Using data from stream reaches located on the eastern slope of the Canadian Rocky Mountains, Paul and Post (2001) developed LRMs to relate the absence|presence of four salmonids [bull trout, cutthroat trout (*Oncorhynchus clarki*), brook trout (*S. fontinalis*), and rainbow trout (*O. mykiss*)] to nine physical habitat characteristics measured at the site scale. Elevation and bank height were the only characteristics found to be significantly related to fish presence. They did not estimate the predictive classification accuracy of the LRMs they developed.

LRMs were successfully used by Kruse et al. (1997) to develop models to predict the absence|presence of Yellowstone cutthroat trout (*O. clarki bouvieri*) in the Greybull-Wood river drainage in northwestern Wyoming. Models were developed using data from sites in the Greybull River and three physical characteristics (channel slope, elevation, and wetted stream width) measured at the site scale. Independent data from the Wood River system was used to assess the predictive accuracy of the LRMs they developed: overall classification accuracies ranged from 74% to 91% for the models they developed.

Porter et al. (2000) used LRMs to develop predictive (absence|presence) models for 14 species of fish in the Blackwater River drainage (British Columbia) using four map-based habitat characteristics (drainage area, watershed gradient, reach gradient, and elevation). They developed significant ($P < 0.02$) LRMs for chinook salmon (*O. tshawytscha*) and rainbow trout. Overall self-classification[1] accuracies for these two models were 73% for chinook and 90% for rainbow trout. The addition of physical characteristics measured in the field from the sites used to estimate the LRMs did not improve the performance of these models.

Preliminary Studies Using Logistic Regression Models for Water Typing:

Logistic regression models were examined as one possible method for classifying streams by the Water Typing Committee. Preliminary GIS-based models to predict fish habitat were generated and assessed (Fransen et. al. 1997). The GIS modeling effort took advantage of emerging technology to produce a modeling platform capable of characterizing physical characteristics for stream systems at the landscape scale. Physical characteristics such as basin area, stream gradient, elevation, and mean annual precipitation could be generated and rapidly assigned to GIS points representing entire watersheds. The GIS model platform

[1] Model classification accuracy estimated from the same data used to estimate the model parameters.

largely automated the data generation and model application process. These physical data were then related to information on fish absence|presence collected during hundreds of field surveys. Models developed from these data could be rapidly and consistently applied to the stream network and evaluated against independent survey information.

The modeling process developed included the following basic steps:

- Generate a GIS-derived stream network from a USGS DEM network.
- Assign physical characteristics such as basin area, channel gradient, elevation, and precipitation from the DEM and other coverages to points along the stream network.
- Assign a fish presence or absence classification to all points with known fish status (based upon field surveys).
- Extract logistic regression modeling data using the points with known fish status.
- Generate logistic regression models from the modeling data.
- Apply the model and assign a probability of fish presence to all network points.
- Develop a heuristic rule to identify the point representing the likely upper extent of fish habitat.
- Apply the resulting model and heuristic rule to a GIS stream network.
- Assess the results against independent survey data.

<u>The Forest and Fish Report</u>

In 1997, the development of new Forest Practices regulations by a wide range of stakeholder groups was underway as a component of the "Forestry Module" of the Governorís statewide salmon recovery strategy. The products of these negotiations were eventually known as the Forest and Fish Report (FFR). During this process, several key elements for an improved water typing system provided the sideboards within which the performance of a new stream classification system would be evaluated. The concept of "shared risk" would provide the basis for determining how the burden of over and under-classification error in the new stream classification system would be distributed. Language calling for determination the identification of fish habitat rather than the presence of fish was adopted. Water bodies were to be classified with an equal likelihood of over and under-classification of fish habitat and a performance target for model accuracy was adopted.

The results from the preliminary GIS fish habitat modeling work were supported by the FFR negotiators and that approach was identified as the preferred option for water typing under the new rules. Changes to the Forest Practices regulations that were outlined in the Forest and Fish Report included direction to develop a GIS-based fish habitat model for classification of fish-bearing waters and the end of electro-shocking as a tool for routinely classifying fish habitat. The Forest Practices Board adopted the new rules in 1999.

Project Initiation and Data Accumulation

Following the adoption of the Forest and Fish Report as the foundation for new Forest Practices regulations, work began on the prescribed fish habitat model. The model would be a collaborative effort of the state, federal, tribal, and landowner groups within the Forest and Fish process. The ad hoc Water Typing committee was renamed the "In-stream Scientific Advisory Group" (ISAG), and assimilated into the Cooperative Monitoring, Evaluation, and Research group (CMER) which is the technical body for the FFR structure.

The Forest Practices Board and Forest and Fish policy group instructed ISAG to develop the model as rapidly as possible so that continued electro-fishing effort could be reduced or eliminated. The urgent need for the model along with concerns about impacts to fish from continued electro-fishing surveys and possible "take" of ESA listed species precluded the execution of a carefully designed and statistically valid field data collection effort. Guidance provided to ISAG was to generate the model using existing survey data to facilitate rapid model implementation. Subsequent refinement of the model would occur as part of a carefully designed research effort employing new survey data collected with a statistically valid sampling design and consistent survey protocol.

The existing survey data available were collected under similar protocols, all requiring at least a 1/4 mile distance of electro-fishing survey effort above the location of the uppermost fish to establish fish absence. Two primary factors influenced survey protocol variability.

- The regulations governing the survey protocol changed somewhat during the years of data collection with the inclusion of non-salmonid fish species after 1998. Prior to that time, only salmonid fish species were considered under the regulatory definition.
- There were differences in the application of the protocol among the various organizations performing the surveys. Some groups identified the precise location of the upper limit of fish distribution at the time of the survey, while others estimated the upper limit of additional available habitat beyond the uppermost fish when they felt fish were not occupying the full extent of available habitat during the survey.

As a result of these protocol differences, four categories of survey data were established:

1. Last Salmonid (LS) - Surveys conducted prior to 1999 which identified the upper limit of salmonid species.
2. Last Salmonid Habitat (LSH) ñ Surveys which estimated the extent of available habitat at or above the LS location to identify the upper limit.
3. Last Fish (LF) ñ Surveys which identified the uppermost extent for fish of any species.
4. Last Fish Habitat (LFH) ñ Surveys which estimated extent of available habitat at or above the LF location to identify the upper limit.

The variability in the survey protocols and the non-random sampling approach for collection of the available survey information presented obvious challenges for model development and assessment. However, the large number of readily available survey data collected by a diverse group of stakeholders offered opportunities as well. In addition, the policy requirement for a model to be produced as quickly as possible using the available data was firm. A preliminary evaluation of protocol differences suggested that variability in survey protocol and the inclusion of all fish species did not usually result in large differences in the location of the identified habitat classification break. It was decided that the survey information would be pooled to produce a preliminary model. A post hoc assessment of the influence of protocol differences and non-random data collection would be conducted when the modeling was completed. A more rigorous sampling scheme employing a single data collection protocol would be used in the next modeling iteration. Work began on the preliminary model in 2000.

The modeling work for the state was divided along the crest of the Cascade Mountains into an eastside and a westside effort. Westside modeling work began immediately using the available survey data. As the westside model was developed and tested, data for the eastside would be collected using a more deliberate sampling approach and a consistent survey protocol.

A call for existing westside stream survey data was made by ISAG which solicited maps and field notes from all interested parties for use in the development and testing of the new system. The Washington Department of Natural Resources was assigned responsibility for model implementation tasks, and for maintaining and updating the classification maps. A contractor was hired to develop the GIS framework required to generate the data required to develop, evaluate, and apply the model. Timelines for completion of tasks were developed. A group of statisticians was formed with representation from landowner, federal, and tribal organizations to conduct the model development and testing.

Concurrently, ISAG initiated several research projects intended to address key technical uncertainties regarding the reliability of electro-fishing surveys as the primary tool employed to identify fish habitat. Research was initiated to characterize temporal variability in the location of the uppermost fish, determine the influence of forest management on the location of the uppermost fish, assess the influence of annual precipitation patterns on fish distribution, and measure sampling efficiency of the electro-fishing surveys (Cole and Lemke 2002). ISAG also initiated the development of a statistically sound study plan to guide collection of the data for the next model iteration. Within this effort, the preliminary model will be evaluated as new survey data are collected. Together, this research will greatly advance our understanding of important issues affecting the classification of fish habitat and the opportunities for improvement of the preliminary model. Refinement of the fish habitat model will continue as part of the adaptive management process within the Forest and Fish program.

The following sections of the report summarize the work of ISAG, the Statistical Sub-group, DNR, and our contractors in developing and evaluating the performance of the preliminary Forest and Fish GIS-based fish habitat classification model.

LOGISTIC REGRESSION MODEL DEVELOPMENT

This section of the report describes the statistical methods and data used for the development of the logistic regression models. It includes summary descriptions of the data used for model development and the results of models developed during the process of arriving at a " final " logistic regression model.

METHODS

Binary Logistic Regression Analysis

Binary logistic regression analysis[2] is similar to the linear regression analysis that is familiar to many people. One major difference between the two methods is that the dependent (response) variable for logistic regression can assume only two values, i.e., an event does not occur or an event occurs. Or for the binary logistic regression models used in this report, fish are absent or fish are present. Typically, the dependent variable is coded 0 to indicate absence (the event not occurring) or 1 to indicate presence (the event occurring).

Similarly to linear regression analysis, logistic regression analysis is used to relate the dependent variable to one or more independent (explanatory) variables. In logistic regression analysis the independent variables can be either continuous or discrete. The logistic distribution is then used to describe the expectation of Y (the dependent variable) given X (the set of independent variables). There are two reasons for choosing the logistic distribution to relate X and Y (Hosmer and Lemeshow 1989):

- " it is an extremely flexible and easily used function" , and
- " it lends itself to a biologically meaningful interpretation" .

The " biologically meaningful interpretation " provided by the logistic regression function is an estimate of the probability of an event occurring [$P(Y)$] or, for our model, the probability of a fish being present.

This logistic regression model is written as

$$P(Y) = \frac{e^{B_0 + B_1 X}}{1 + e^{B_0 + B_1 X}}$$

which is equivalent to

$$P(Y) = \frac{1}{1 + e^{-(B_0 + B_1 X)}}$$

where B_0 and B_1 are coefficients estimated from the data, X is the independent variable, and e is the base of the natural logarithm. If there are multiple independent variables we can define Z as:

$$Z = B_0 + B_1X_1 + B_2\,X_2 + \text{Ö} \; + B_pX_p$$

[2] Binary logistic regression and logistic regression are used interchangeably.

where p is the number of variables. This resembles the common method of denoting a linear regression equation. The logistic model can then be written as

$$P(Y) = \pi(X_i) = \frac{1}{1+e^{-Z}} \qquad [1]$$

where $\pi(X_i)$ is the probability of occurrence given the set of independent variables X_i. The relationship between the independent variable(s) and the estimated probabilities described by the model is non-linear and the probability estimates will always be between 0 and 1.

In linear regression the parameters of the model are estimated using the method of least squares. In binary logistic regression the parameters of the model are estimated using the method of maximum likelihood. In a general sense, "the method of maximum likelihood yields values for the unknown parameters which maximize the probability of obtaining the observed set of data" (Hosmer and Lemeshow 1989). In maximum likelihood estimation, the likelihood function is constructed which expresses the probability of the observed data as a function of the unknown parameters. An iterative algorithm is then used to estimate the parameters which maximize the likelihood function. It is easier mathematically to work with the log likelihood function ($\ln[l(B)]$) which is defined as (Hosmer and Lemeshow 1989)

$$\ln[l(B)] = \sum_{i=1}^{n} \{ y_i \ln[\pi(X_i)] + (1 - y_i)\ln[1 - \pi(X_i)] \} \qquad [2]$$

where y_i = the observed value of the dichotomous dependent variable (0 or 1),

$\pi(X_i)$ = the probability of occurrence given the set of independent variables X_i (defined by equation 1), and

n = the number of independent observations used to estimate the model.

Estimating the Significance of the Model Coefficients:

Logistic regression uses the same principle as linear regression to test the significance of model coefficients. That is, observed values of the response variable are compared to predicted values obtained with and without the independent variable in question. In logistic regression, this comparison is based on the log likelihood function defined in equation 2. This can be expressed as

$$G = -2\ln\left[\frac{\textit{(likelihood without the variable)}}{\textit{(likelihood with the variable)}}\right] \qquad [3]$$

where the statistic, G, follows the chi-square distribution (Hosmer and Lemeshow 1989). If G is significant (P > 0.05 for this report) then the addition of the variable significantly improved the fit of the model; this is referred to as the likelihood-ratio test.

Model Estimation and Variable Selection:

We used the binary logistic regression procedures in SPSS version 11.0 (Norusis 1999) to conduct our LRM analyses. A forward, stepwise selection procedure based on the likelihood-ratio test was used to select the independent variables to include in the logistic regression models. Variable entry into the model at each step was based on the significance of the likelihood-ratio test comparing the likelihood of the model with the variable in the model to the likelihood of the model without the variable. The variable with the smallest significance level for the *G* statistic was entered into the model contingent on the significance being > 0.05. Once a variable had entered into the model, the significance of previously entered variables was assessed. The log-likelihood for each of the models with only one of the previously entered independent variables removed was computed. The change in -2LL[3] was then computed for each of these models compared to the model with the new variable and, if the observed significance level for the change was 7 0.10, the previously entered independent variable was removed from the model.

Logistic Regression Model Fit and Performance Measures:

Several statistics were used to assess the fit of the logistic regression models developed and to compare different models. One measure of model fit was -2 multiplied by the log likelihood function (equation 2). This statistic (-2LL) is a function of the number of observations used to estimate the logistic regression. When comparing LR models estimated using the same number of observations, smaller values of -2LL indicate a better fit to the data.

The Hosmer and Lemeshow goodness of fit chi-square statistic was also used to evaluate model fit and compare models (Hosmer and Lemeshow 1989). This statistic is estimated by dividing the observations used to estimate the LR model into ten approximately equal (in number of observations) groups based upon their values for the predicted probability of fish presence (equation 1). For each group two cells are generated, one with the observed and the predicted number of fish absent occurrences, and one with the observed and the predicted number of fish present occurrences[4]. Then a chi-square statistic is computed from the cells.[5] The significance of the resulting chi-square statistic can then be determined. If the observed significance level was 7 0.05 the null hypothesis was not rejected. This indicated that there was no difference between the observed values and those predicted by the LR model and the model appears to fit the data reasonably well. The value of the Hosmer and Lemeshow goodness of fit chi-square is proportional to the sample size therefore with very large sample sizes the statistic can be large (and significant) even if the model fits well (Norusis 1999).

---

[3] -2 multiplied by the natural log of the likelihood (equation 2).

[4] Observations with a predicted probability of fish presence < 0.50 were classified as fish absent points and observations with a predicted probability of fish presence 7 0.50 were classified as fish present points.

[5] The chi-square statistic is calculated as the (observed number ñ predicted number)^2/predicted number. The final chi-square statistic is the sum of these calculations across all cells. The degrees of freedom for this statistic are the number of groups minus two.

Logistic regression model performance was assessed using both self-classification accuracy and validation data that were reserved and not part of the model estimation process. Self-classification accuracy is calculated using the same observations that were used to estimate the LR model. The model observations were classified with the resulting model and two measures of classification accuracy were calculated; (1) the percent of fish absent points correctly classified as fish absent points and (2) the percent of fish present points correctly classified as fish present points.

Data that were excluded from the model estimation process were used to assess model performance, also. Fish presence probabilities for these points were estimated using the logistic regression equation being examined. Based upon the estimated probabilities, the points were then classified as fish absent or fish present using the same procedure used for self-classification (see footnote 4). The two measures of classification accuracy described above were then calculated.

Data Used for Logistic Regression Analyses

The fish absent | fish present (FAFP) data used to estimate the logistic regression models for this project were all generated from end-of-fish points (EOFPs)[6] placed on a Washington Department of Natural Resources GIS hydraulic layer. Each EOFP was based on a field survey which followed specific protocols to identify a location on the stream that was designated as either last fish (LF), last fish habitat (LFH), last salmonid (LS), or last salmonid habitat (LSH).

Potential EOFPs were submitted to DNR for error checking and initial screening. After approval by DNR, the EOFP was forwarded to EarthRes.i for processing and placement on a 10m DEM-generated stream points network. After the EOFP had been placed onto a point in the 10m DEM-generated stream network, an automated procedure was used to classify points upstream of the EOFP as fish absent (FA) points and points downstream of the EOFP as fish present (FP) points. Needham (2001) describes the GIS processes that were used for this project.

EOFPs were submitted by 21 different groups (Table 1) which included state and tribal natural resource departments, timber companies, and environmental organizations. The EOFPs submitted to DNR were essentially "found" data meaning that, as a group, the data were not collected under a single, statistically-based sampling plan. The EOFPs do not represent a random sample of the stream systems in western Washington but are what is often referred to as a "haphazard" sample. Because of the haphazard nature of the data collection, we felt it was very important to try to include as many EOFPs as possible to get a broad representation of the Western Washington landscape. Rather than being exclusive, i.e., looking for reasons to exclude data from the analysis, we tried to be inclusive and have as many EOFPs as possible available to generate FAFP points for model building.

[6] We refer to the points as end-of-fish points. This is a generic reference for convenience. The majority of the field survey points used in the analyses were Last Fish Habitat points. However, Last Fish, Last Salmonid, and Last Salmonid Habitat survey points were also used.

Table 1 Summary of the number (#) of end-of-fish points (EOFPs) contributed by each sponsoring agency, by point type (% is the percentage of the total number of EOFPs).

Sponsoring Agency		End-of-Fish Point Type[a]					Total
		Unknown	LF	LFH	LS	LSH	
Adopt-a-Stream	#				18		18
	%				0.4%		0.4%
Aquatic Tech	#		16	73	25		114
	%		0.4%	1.8%	0.6%		2.8%
Champion	#		3	76		3	82
	%		0.1%	1.9%		0.1%	2.0%
Crown Pacific	#			23			23
	%			0.6%			0.6%
WA Dept. Fish and Wildlife	#			1	33	21	55
	%			<0.1%	0.8%	0.5%	1.4%
WA Dept. Natural Resources	#		4	99	2	54	159
	%		0.1%	2.4%	<0.1%	1.3%	3.9%
Hoh Fisheries Dept.	#		22	366	1		389
	%		0.5%	9.0%	<0.1%		9.6%
International Paper Co.	#			8			8
	%			0.2%			0.2%
Longview Fibre Co.	#		5				5
	%		0.1%				0.1%
Merril and Ring	#			2			2
	%			<0.1%			<0.1%
Point No Point Treaty Council	#			1			1
	%			<0.1%			<0.1%
Olympic Environmental	#		1				1
	%		<0.1%				<0.1%
Olympic Resource	#		9	21			30
	%		0.2%	0.5%			0.7%
Port Blakely Tree Farm	#		9	82	15		106
	%		0.2%	2.0%	0.4%		2.6%
Plum Creek Co.	#		7	3		40	50
	%		0.2%	0.1%		1.0%	1.2%
Quinault Dept. Nat. Resources	#		182	2,214	89		2,485
	%		4.5%	54.6%	2.2%		61.3%
Scatter Creek	#			5			5
	%			0.1%			0.1%
Tulalip Dept. Nat. Resources	#		1		6	44	51
	%		<0.1%		0.1%	1.1%	1.3%
Washington Trout	#		20	144	27		191
	%		0.5%	3.6%	0.7%		4.7%
West Fork	#				15		15
	%				0.4%		0.4%
Weyerhaeuser Co.	#	2	255	2	3		262
	%	<0.1%	6.3%	<0.1%	0.1%		6.5%
TOTALS	#	2	534	3,120	234	162	4,052
	%	<0.1%	13.2%	77.0%	5.8%	4.0%	100.0%

[a] LF = Last Fish; LFH = Last Fish Habitat; LS = Last Salmonid; and LSH = Last Salmonid Habitat.

However, there was a group of end-of-fish points collected by one agency that was not used in the logistic regression model building process. Washington Trout had submitted a group of EOFPs collected using a survey protocol that existed prior to the establishment of the emergency stream typing rules in 1996. They (Washington Trout) requested that these EOFPs not be used in the analyses. Since they were the agency that had originally collected the data we excluded these EOFPs, and the FAFP points generated from them, from the model building process. This removed a total of 106 EOFPs from the original total of 4,158 EOFPs available for logistic regression model building[7]. A summary of these omitted EOFPs is provided in Appendix Table 1. There were still 191 EOFPs collected by Washington Trout used in the process to generate FAFP data for the logistic regression model building process (Table 1).

Information Associated with Each End-of-Fish Point:

There was a large amount of information associated with each end-of-fish point. Table 2 summarizes the information that was important for our analyses associated with each EOFP.

Physical Attributes Associated with Each Fish Absent | Fish Present Point:

There were five GIS-derived physical characteristics or attributes associated with each FAFP point (including EOFPs) placed on the 10m DEM network. They were:

1. basin size (number of acres in surrounding basin that drain through a point),
2. elevation in feet (based on 10m DEM network),
3. upstream gradient which is the average gradient measured over 100 m upstream of the point (calculated from 10m DEM network elevation information),
4. downstream gradient which is the average gradient measured over 100 m down-stream of the point (calculated from 10m DEM network elevation information), and
5. precipitation in inches (GIS derived estimate of average annual precipitation at the point based on Daly et al. [1998]).

These five physical attributes associated with each FAFP point were the variables available for the logistic regression model building process. The acronyms we use in this report for each of the physical attributes are: BASIZE for basin size, ELEV for elevation, UPGRD for upstream gradient, DNGRD for downstream gradient, and PRECIP for precipitation.

[7] Because our analysis methods were based on using FAFP data associated with fourth order sub-basins (see page 26), there were an additional five EOFPs collected by other agencies that were excluded because they were in the same fourth order sub-basin as one of the excluded Washington Trout EOFPs.

Table 2 Information associated with the fish absent | fish present data points. Information associated with items marked with an asterisk (*) was recorded for all points.

		Information Item	Description
a.	*	Unique ID	Unique ID, a unique identifying number
b.	*	X coordinate	the X coordinate of the WA South, State plane
c.	*	Y coordinate	the Y coordinate of the WA South, State plane
d.	*	Stream Order	stream order calculated from 10m DEM elevation information
e.		EOFP flag	"Y" for all end-of-fish points used to generate FAFP data
f.	*	Absence|Presence indictaor	"Y" for a Fish Present point or "N" for a Fish Absent point
g.		Sponsor	The name of the agency, organization, tribe, or company that conducted the field survey
h.		Survey Date	The date the field survey was conducted
i.		Protocol	Protocol used during the field survey
j.		Point Type	The point type represented under the specified protocol (LF, LFH, LS, LSH)
k.		Boundary Type	The location of the EOFP on the stream: A = Mid-channel, B = Confluence point (non-fish bearing stream laterally intersecting a fish-bearing stream), or C = Tributary junction (two or more non-fish bearing streams join to form a fish-bearing stream)
l.		Determination Method	Method used to detect EOFP: 1 = Field electro-shocking, 2 = Day snorkeling, 3 = Night snorkeling, 4 = Visual 0bservation, or 5 = Unknown
m.		End Type	Reason for placement of the end point: 1 = Natural end, 2 = Gradient related, 3 = Large woody debris, 4 = Road culvert, 5 = Mass-wasting event, 6 = Beaver dam or other non-permanent dam, 7 = Permanent dam, 8 = Water quality limiter, 9 = None, or 10 = Unknown

Unresolved and Conflicting End-of-Fish Points:

There were a number of end-of-fish points submitted to EarthRes.i that could not be used to generate FAFP data or that required a decision to determine which of a conflicting pair of EOFPs to use in the process for generating FAFP points. Appendix A gives a summary of the number of EOFPs that were affected and the details of the analyses that were conducted which informed this decision-making process. This resulted in the following:

- Unresolved end-of-fish points were single EOFPs whose recorded location could not be associated with a stream on the 10m DEM network during processing by EarthRes.i. These EOFPs were not included in the process used to generate FAFP points. There were 48 unresolved, single EOFPs.
- There were occasionally end-of-fish points that had been submitted which conflicted with each other, i.e., there were two or more EOFPs specified on the same stream. When there were more than two EOFPs in conflict these potential EOFPs were excluded from all analyses. A total of 42 EOFPs that potentially could have been used to generate FAFP data were omitted from consideration due to multiple (>2) conflicting EOFPs on a stream.
- The majority of the conflicting EOFPs involved a conflicting pair of points. We developed a protocol (described in Appendix A) to examine these pairs of conflicting EOFPs. From a total of 268 conflicting EOFP pairs, the protocol allowed 224 of them to be resolved and used in the process to generate FAFP points for the logistic regression model building process.

End-of-Fish Point Summary:

After the initial processing and resolution of conflicting EOFP pairs, there were 4,052 EOFPs available for generating FAFP data. The majority of these EOFPs (77%) were Last Fish Habitat (LFH) points (Table 1). However, 71% of these LFH end-of-fish points were provided by a single agency, the Quinault Dept. of Natural Resources. The next most common EOFP type was Last Fish (LF). About half of the LF end-of-fish points were collected by the Weyerhaeuser Co. Combined, Last Salmonid (LS) and Last Salmonid Habitat (LSH) EOFPs represented less than 10% of the total EOFPs.

Table 3 summarizes the distribution of the EOFPs among western Washington WRIAs. There were no EOFPs from either WRIA 2 (the San Juan Islands) or WRIA 12 (the Lake Washington watershed including Seattle). More than 75% of the EOFPs came from five WRIAs on the Pacific Coast of Washington (Figure 1): WRIA 20 (Soleduck-Hoh Rivers Basin), 21 (Quinault-Queets Rivers Basin), 22 (Lower Chehalis River Basin), 23 (Upper Chehalis River Basin), and 24 (Willapa River Basin).

Field surveys for the EOFPs were conducted between January 1990 and October 2000. The majority (72%) of the EOFPs were collected during field surveys conducted in 1995, 1996, and 1997 (Table 4). Less than 1% of the EOFPs came from surveys conducted prior to 1994.

Table 3 Summary of the number of end-of-fish points (EOFPs) from each Western Washington WRIA, by point type (% is the percentage of the total number of EOFPs).

	End-of-Fish Point Type						
WRIA	Unknown	**LF**	**LFH**	**LS**	**LSH**	**Total**	%
1			4	2	5	11	0.3%
2						0	
3&4[a]		9	23	2	2	36	0.9%
5			32	30	22	84	2.1%
6				4		4	0.1%
7		11	43	17	52	123	3.0%
8				18		18	0.4%
9		16			44	60	1.5%
10		33	76	6		115	2.8%
11		5	14		1	20	0.5%
12						0	
13		5	5	3		13	0.3%
14		2	5	4		11	0.3%
15		9	19		14	42	1.0%
16		1	5		19	25	0.6%
17		1	25			26	0.6%
18		10	23			33	0.8%
19		1	29			30	0.7%
20		23	412			435	10.7%
21		43	415	6		464	11.5%
22		54	410	30		494	12.2%
23		42	1,040	64		1,146	28.3%
24	2	101	402	5	3	513	12.7%
25		1	9			10	0.2%
26		97	102	40		239	5.9%
27		68	8			76	1.9%
28		2	9			11	0.3%
29			10	3		13	0.3%
TOTALS	2	534	3,120	234	162	4,052	100.0%

[a] Data for WRIAs 3 (Lower Skagit River Basin) and 4 (Upper Skagit River Basin) are combined.

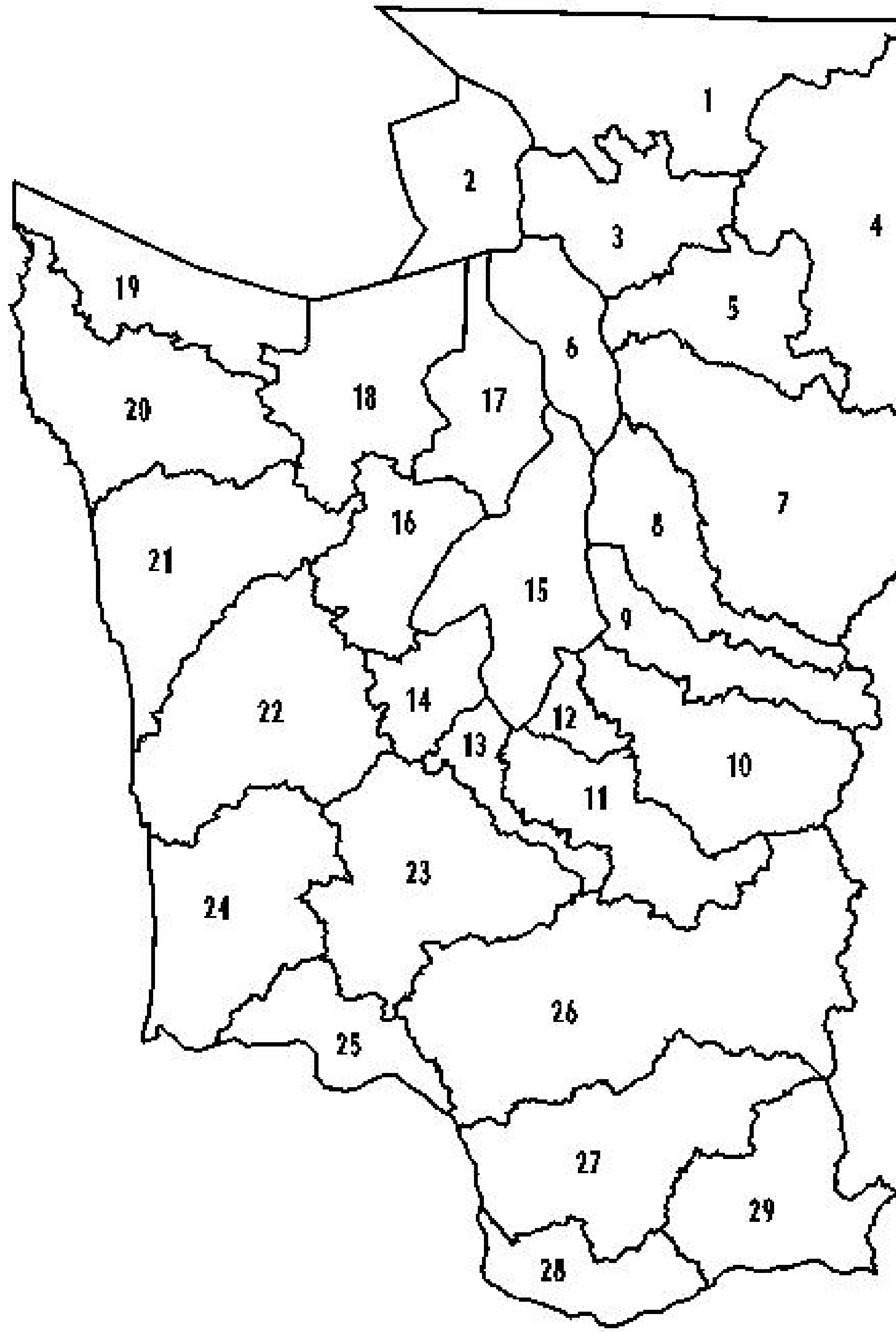


Figure 1 Schematic of Western Washington showing the location of the WRIAs pertinent to this study.

Table 4 Summary of the survey year for end-of-fish points (EOFPs), by point type (% is the percentage of the total number of EOFPs).

Year	End-of-Fish Point Type LF	LFH	LS	LSH	Total[a]	%
1990			1		1	<0.1%
1992			1		1	<0.1%
1993		3	7	4	14	0.3%
1994	23	326	18	5	372	9.2%
1995	91	680	69	4	844	20.8%
1996	64	1,104	16	24	1,208	29.8%
1997	201	539	52	52	844	20.8%
1998	44	139	60	32	275	6.8%
1999	32	181	8	1	222	5.5%
2000	79	148	2	40	269	6.6%
TOTALS	534	3,120	234	162	4,050	100.0%

[a] There was no survey date associated with the two points with an unknown end-of-fish point type.

Table 5 Summary of the number (#) of end-of-fish points (EOFPs) belonging to each boundary type, by point type (% is the percentage of the total number of EOFPs).

End-of-Fish Point Type		Boundary Type[a] A	B	C	Total
Last Fish	#	266	167	101	534
	%	6.6%	4.1%	2.5%	13.2%
Last Fish Habitat	#	872	1,410	838	3,120
	%	21.5%	34.8%	20.7%	77.0%
Last Salmonid	#	123	72	39	234
	%	3.0%	1.8%	1.0%	5.8%
Last Salmonid Habitat	#	84	53	25	162
	%	2.1%	1.3%	0.6%	4.0%
Unknown	#			2	2
	%			<0.1%	<0.1%
TOTALS	#	1,345	1,702	1,005	4,052
	%	33.2%	42.0%	24.8%	100.0%

[a] A = mid-channel boundary; B = confluence point where a non fish-bearing stream laterally intersects a fish-bearing stream; and C = tributary junction where two or more non fish-bearing streams join to form a fish-bearing stream.

All three boundary types were well represented in the EOFP data (Table 5). Although 42% of the EOFPs were associated with non fish-bearing streams laterally intersecting a fish-bearing stream (type B), there were also more than 1,000 EOFPs from both mid-channel EOFPs (type A) and tributary junctions where two or more non fish-bearing streams join to form a fish-bearing stream (type C). Figure 2 illustrates each of these boundary types.

The majority of the EOFPs were determined by electro-shocking during field surveys (Table 6). About 70% of the Last Fish EOFPs and 50% of the Last Fish Habitat EOFPs were determined by electro-shocking. Visual determination of the EOFP was used by 47% of the field surveys. The majority (52%) of the Last Fish Habitat EOFPs were visually determined.

The reason for the placement of almost 85% of the EOFPs (Table 7) was either a natural end[8] or a gradient-related end (end types 1 or 2). The placement of the remaining EOFPs was due to a blockage (man-made or natural) or other reason not specifically related to the five GIS-based physical attributes associated with each FAFP point.

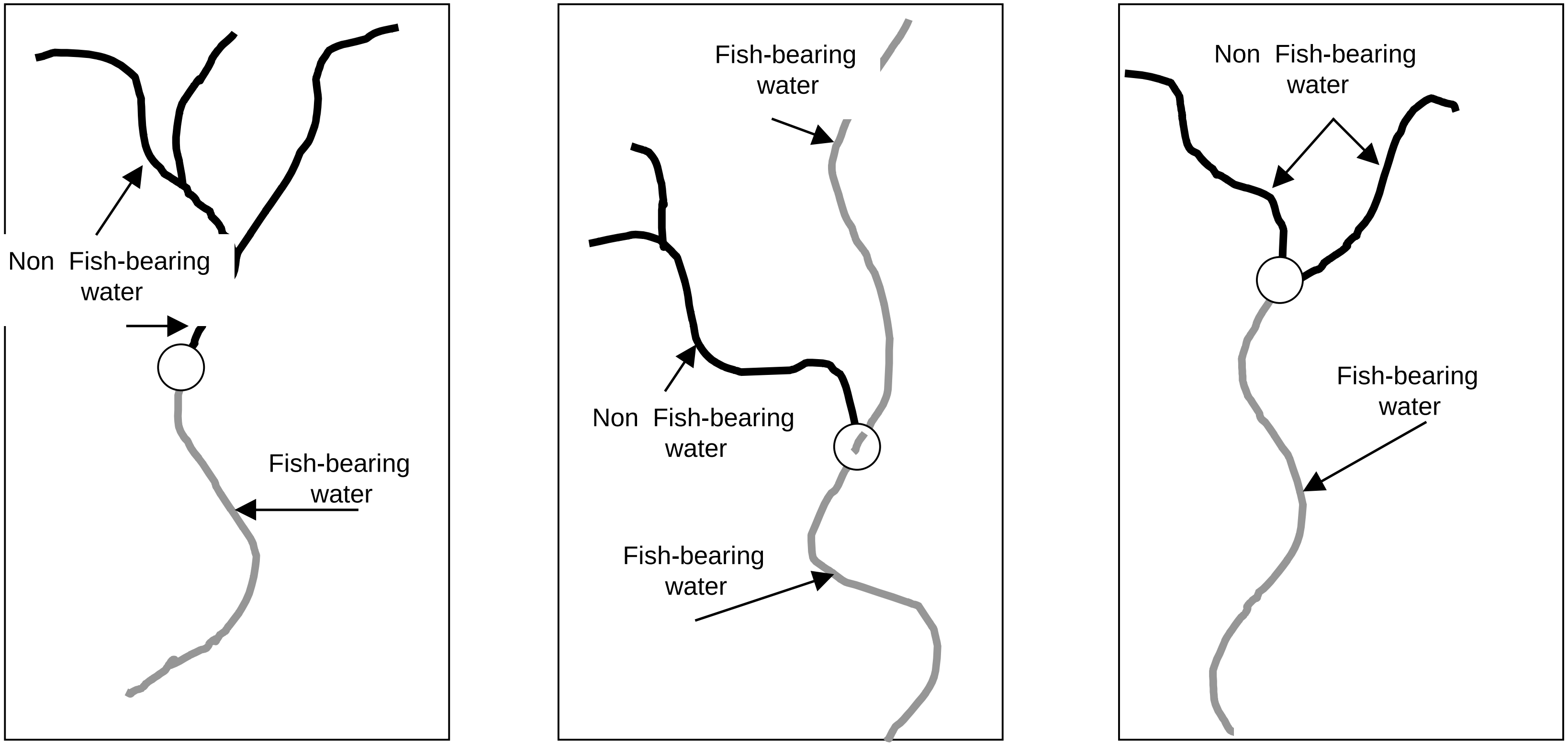
Figure 3 summarizes the point types (LF, LFH, LS, LSH), determination methods (electro-shocking or visual), and boundary types for the 3,435 EOFPs which had end types 1 or 2. Three EOFPs with end types 1 or 2 for which the determination method was unknown are not included in the summary figure.

Generation of Fish Absent | Fish Present Data:

All 4,052 EOFPs were used to generate fish absent | fish present data. Points on the 10m DEM above an EOFP were classified as fish absent points and points below an EOFP were classified as fish present points. There were 1,443,471 FAFP data points generated in the western Washington DEM network (Table 8). WRIAs 1 and 3&4 combined are under-represented in the FAFP data while WRIAs 22, 23, and 24 are over-represented relative to the proportion of total network points contributed by these WRIA to the western Washington DEM network[9] (Figure 4). Generally, the proportional contribution of the other WRIAs to the FAFP data is similar to their representation in the 10m DEM network (i.e., the proportional contribution of each of the other WRIAs to the FAFP data is within 3% of the WRIAís contribution to the 10m DEM network).

[8] A natural end was usually related to the width (size) of the stream decreasing to an extent that fish were no longer present.

[9] Total network points with WRIAs 2 and 12 omitted.



A. Mid-channel **B. Lateral Confluence** **C. Tributary Junction**

Figure 2 Schematic diagram illustrating the three boundary types between fish-bearing and non fish bearing waters.

Table 6 Summary of the number (#) of end-of-fish points (EOFPs) belonging to each determination method, by point type (% is the percentage of the total number of EOFPs).

End-of-Fish Point Type		Determination Method			Total
		Electro-shocking	Visual	Other	
Last Fish	#	372	160	2	534
	%	9.2%	3.9%	<0.1%	13.2%
Last Fish Habitat	#	1,500	1,617	3	3,120
	%	37.0%	39.9%	0.1%	77.0%
Last Salmonid	#	147	87		234
	%	3.6%	2.1%		5.8%
Last Salmonid Habitat	#	111	51		162
	%	2.7%	1.3%		4.0%
Unknown	#			2	2
	%			<0.1%	<0.1%
TOTALS	#	2,130	1,915	5	4,052
	%	52.6%	47.3%	0.1%	100.0%

Table 7 Summary of the number (#) of end-of-fish points (EOFPs) belonging to each end type, by point type (% is the percentage of the total number of EOFPs).

Reason for placement of the end point		End-of-Fish Point Type					Total
		LF	LFH	LS	LSH	Unknown	
Natural End	#	119	1,800	108	90		2,117
	%	2.9%	44.4%	2.7%	2.2%		52.2%
Gradient related	#	146	1,065	53	57		1,321
	%	3.6%	26.3%	1.3%	1.4%		32.6%
Large woody debris	#	77	44	7	1		129
	%	1.9%	1.1%	0.2%	<0.1%		3.2%
Road culvert	#	91	6	22	7		126
	%	2.2%	0.1%	0.5%	0.2%		3.1%
Mass-wasting event	#	16	18	2	1		37
	%	0.4%	0.4%	<0.1%	<0.1%		0.9%
Beaver dam or other	#	9	19	3			31
Non- permanent dam	%	0.2%	0.5%	0.1%			0.8%
Permanent dam	#		1				1
	%		<0.1%				<0.1%
Water quality limiter	#	1	2				3
	%	<0.1%	<0.1%				0.1%
None	#	12	6	12			30
	%	0.3%	0.1%	0.3%			0.7%
Reason unknown	#					2	2
	%					<0.1%	<0.1%
TOTALS	#	534	3,120	234	162	2	4,052
	%	13.2%	77.0%	5.8%	4.0%	<0.1%	100.0%

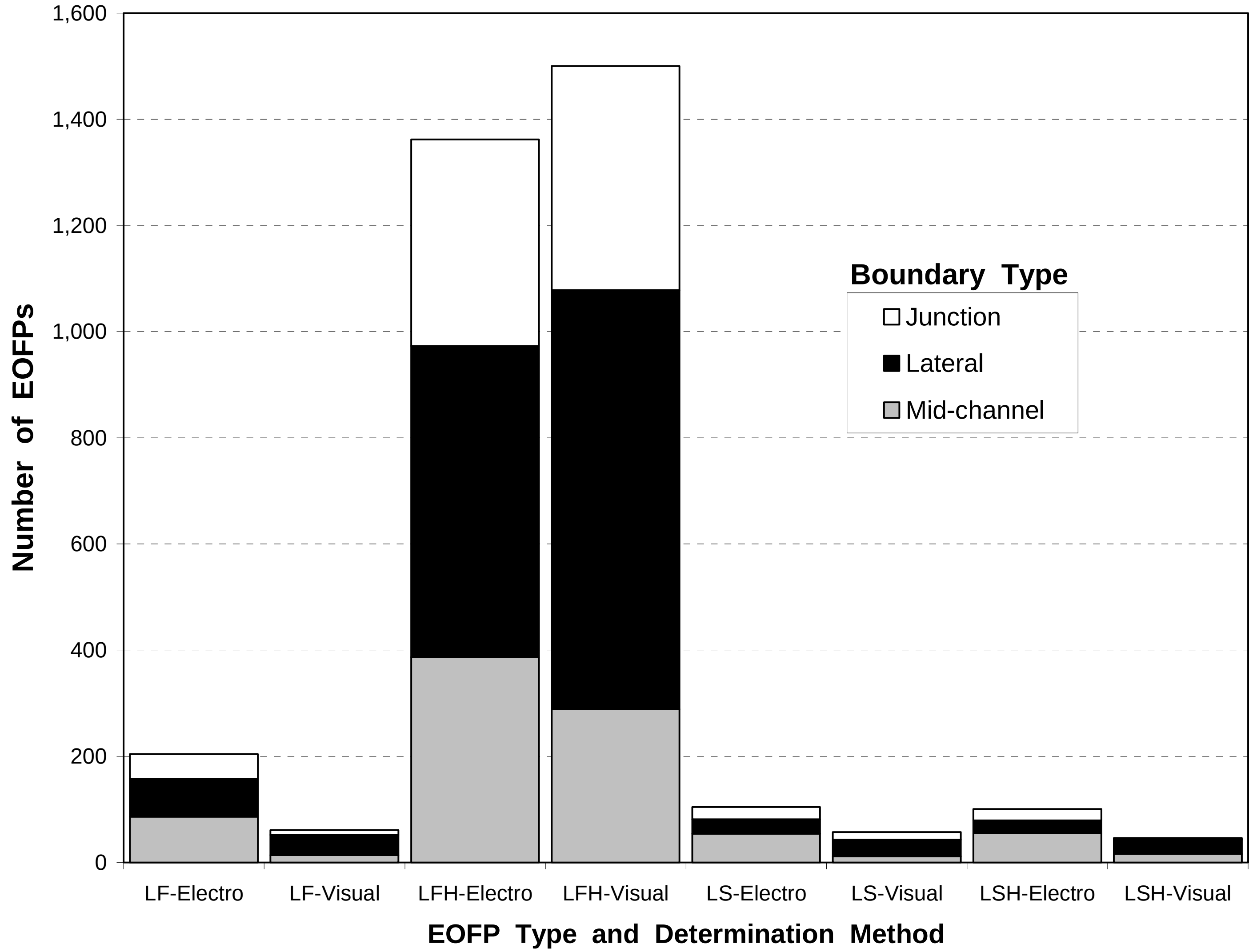


Figure 3 Bar chart summarizing the distribution of the 3,435 end-of-fish points (EOFPs) with end types of 1 or 2, by point type, boundary type, and determination method.

Table 8 Proportional contribution of each WRIA to the western Washington 10m DEM stream network used for the analyses in this report. Proportional contribution summarized for network miles, total 10m DEM network points, and total Fish Absent|Fish Present (FAFP) data points available for logistic regression analyses.

WRIA	Area (sq. miles)	% of Total	10m DEM Miles	% of Total	10m DEM Points	% of Total	FAFP Points	% of Total
1	1,271.87	5.0%	16,968	7.4%	2,292,227	7.4%	25,564	1.8%
2	183.74	0.7%		0.0%		0.0%		0.0%
3&4	2,994	11.8%	35,124	15.3%	4,741,618	15.3%	77,607	5.4%
5	703.29	2.8%	6,763	2.9%	902,332	2.9%	56,696	3.9%
6	220.41	0.9%	2,103	0.9%	287,198	0.9%	4,510	0.3%
7	1,868.24	7.3%	16,757	7.3%	2,275,979	7.3%	74,296	5.1%
8	641.15	2.5%	5,721	2.5%	769,851	2.5%	12,244	0.8%
9	544.26	2.1%	4,779	2.1%	644,297	2.1%	44,398	3.1%
10	1,040.27	4.1%	9,062	3.9%	1,222,644	3.9%	57,920	4.0%
11	766.60	3.0%	6,763	2.9%	917,815	3.0%	24,796	1.7%
12	155.94	0.6%		0.0%		0.0%		0.0%
13	269.81	1.1%	2,220	1.0%	299,835	1.0%	14,841	1.0%
14	332.73	1.3%	2,756	1.2%	368,699	1.2%	4,097	0.3%
15	679.36	2.7%	6,296	2.7%	855,096	2.8%	16,987	1.2%
16	603.70	2.4%	4,414	1.9%	593,142	1.9%	6,729	0.5%
17	406.46	1.6%	3,629	1.6%	489,839	1.6%	17,846	1.2%
18	701.75	2.8%	6,397	2.8%	857,558	2.8%	72,088	5.0%
19	369.02	1.4%	2,903	1.3%	389,274	1.3%	12,189	0.8%
20	1,202.00	4.7%	9,772	4.2%	1,306,319	4.2%	84,912	5.9%
21	1,170.35	4.6%	10,294	4.5%	1,376,830	4.4%	76,144	5.3%
22	1,322.44	5.2%	10,818	4.7%	1,447,494	4.7%	164,903	11.4%
23	1,291.94	5.1%	10,081	4.4%	1,347,776	4.3%	232,347	16.1%
24	1,011.32	4.0%	7,267	3.2%	981,348	3.2%	142,155	9.8%
25	503.73	2.0%	3,499	1.5%	469,381	1.5%	10,700	0.7%
26	2,493.56	9.8%	22,014	9.6%	3,004,165	9.7%	128,750	8.9%
27	1,307.38	5.1%	11,277	4.9%	1,540,370	5.0%	56,067	3.9%
28	493.69	1.9%	4,355	1.9%	595,704	1.9%	14,193	1.0%
29	900.54	3.5%	8,173	3.6%	1,101,370	3.5%	10,492	0.7%
TOTALS	25,449.75		230,205		31,078,161		1,443,471	

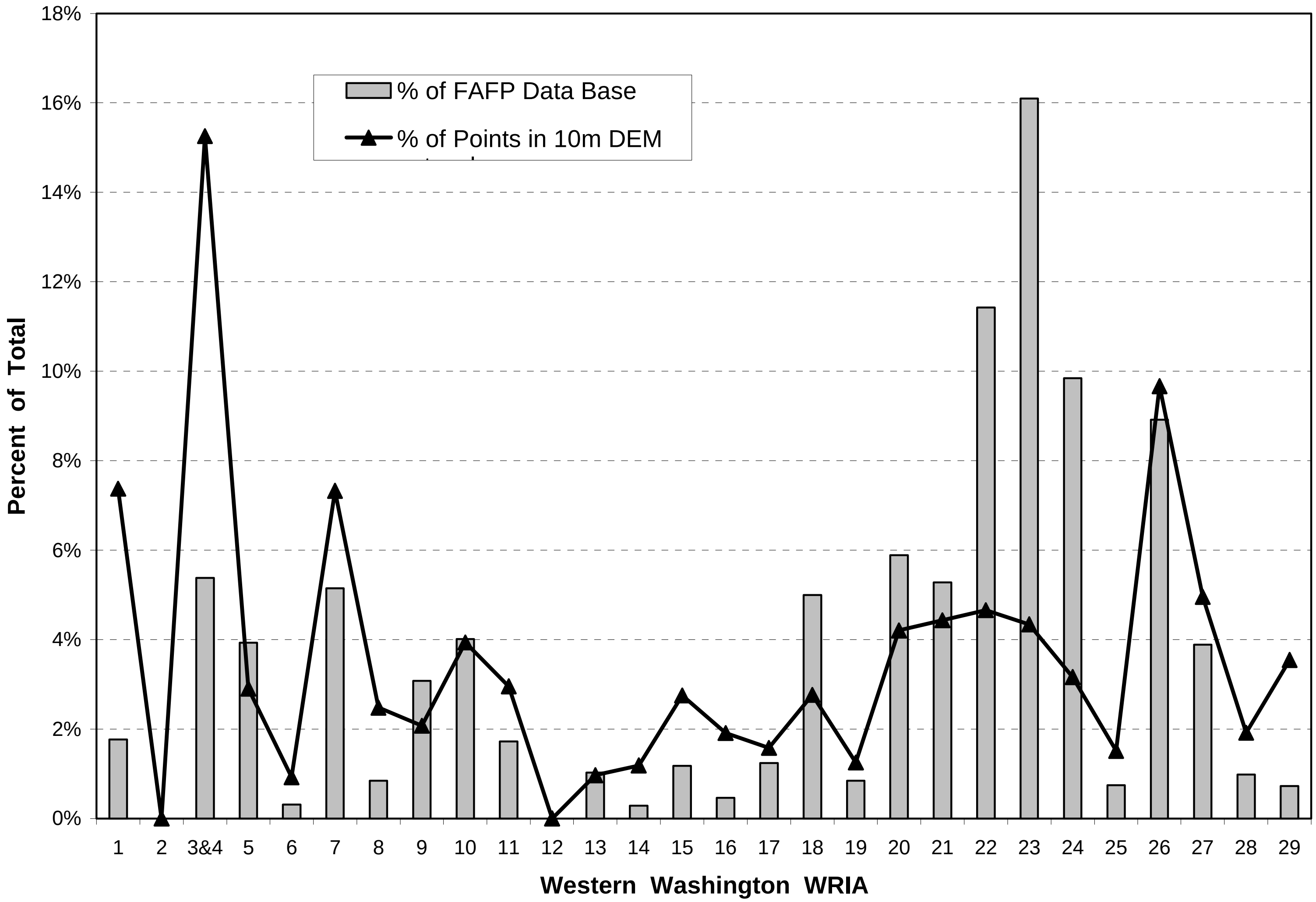


Figure 4 Percentage contribution of each western Washington WRIA to the total number of points in the 10m DEM network and the percentage contribution of each WRIA to the total fish absent|fish present (FAFP) data base.

Definition of a Fourth Order Sub-Basin:

The majority of the end-of-fish points are associated with smaller sub-basins at the upper end of stream systems. The physical attributes (basin size, elevation, gradient, and precipitation) associated with the Fish Absent|Fish Present points in these upstream reaches are more informative for building a logistic regression model than the physical attributes of lower mainstem FAFP points. Mainstem FAFP points are almost always Fish Present points and are typically characterized by very large basin sizes (relative to the other FAFP points), low elevations, and low gradients. These mainstem FAFP points can heavily influence the model building process and the assessment of model performance since they are rarely misclassified. For example, almost half (49%) of the 817,985 Fish Present points have basin sizes greater than 7,200 acres. However, none of the 625,486 Fish Absent points in the database have a basin size greater than 7,103 acres. We wanted an objective method of limiting the data used to estimate any LRM to those points associated with the upper ends of streams as these data should be the most informative. The approach we decided upon was to divide the FAFP data into small subsets that were approximately defined by a DEM-generated stream network Strahler fourth order sub-basin (FOSB).

There is a major benefit to this approach when we evaluate the performance of the logistic regression models that are developed. By dividing the data into small subsets and selecting a sub-sample of these FOSBs to use in the LRM building process, the FOSBs not used in the model provide a quasi-independent set of data to assess model performance. The data in the non-model FOSBs are generated from EOFPs that were not used to generate the FAFP data used to estimate the model. Therefore, the data from non-model FOSBs are very similar to the data that would be generated from a post-analysis field survey which specified an end-of-fish point(s) on a system. The FAFP data generated from this "new" EOFP can then be used to validate (assess the performance) of an existing LRM.

We defined fourth order sub-basins based on the change in stream order from 4 to 5 in a string of FAFP points on the 10m DEM network. The fourth order sub-basins were consecutively numbered within each WRIA (Figure 5). All FAFP data that were not part of a FOSB in a WRIA were allocated to a separate group labeled 9999. These are subsequently referred to as <u>unassigned</u> FAFP data. It is important to note that not all Fish Absent points or all end-of-fish points are associated with a FOSB.

Results of Grouping Data into Fourth Order Sub-basins:

There were 895 FOSBs defined within the 1,443,471 FAFP points summarized in Table 8[10]. There were 28 FOSBs removed from this total because they did not contain an EOFP which left 867 FOSBs which contained at least one EOFP[11]. Table 9 summarizes the number of

---

[10] There were originally 935 fourth order sub-basins (FOSBs) defined in the data. Forty (40) FOSBs which contained at least one EOFP from a Washington Trout survey using the "PRE" protocol were removed.

[11] The FAFP data from these 28 FOSBs were used in the logistic regression model validation process, however.

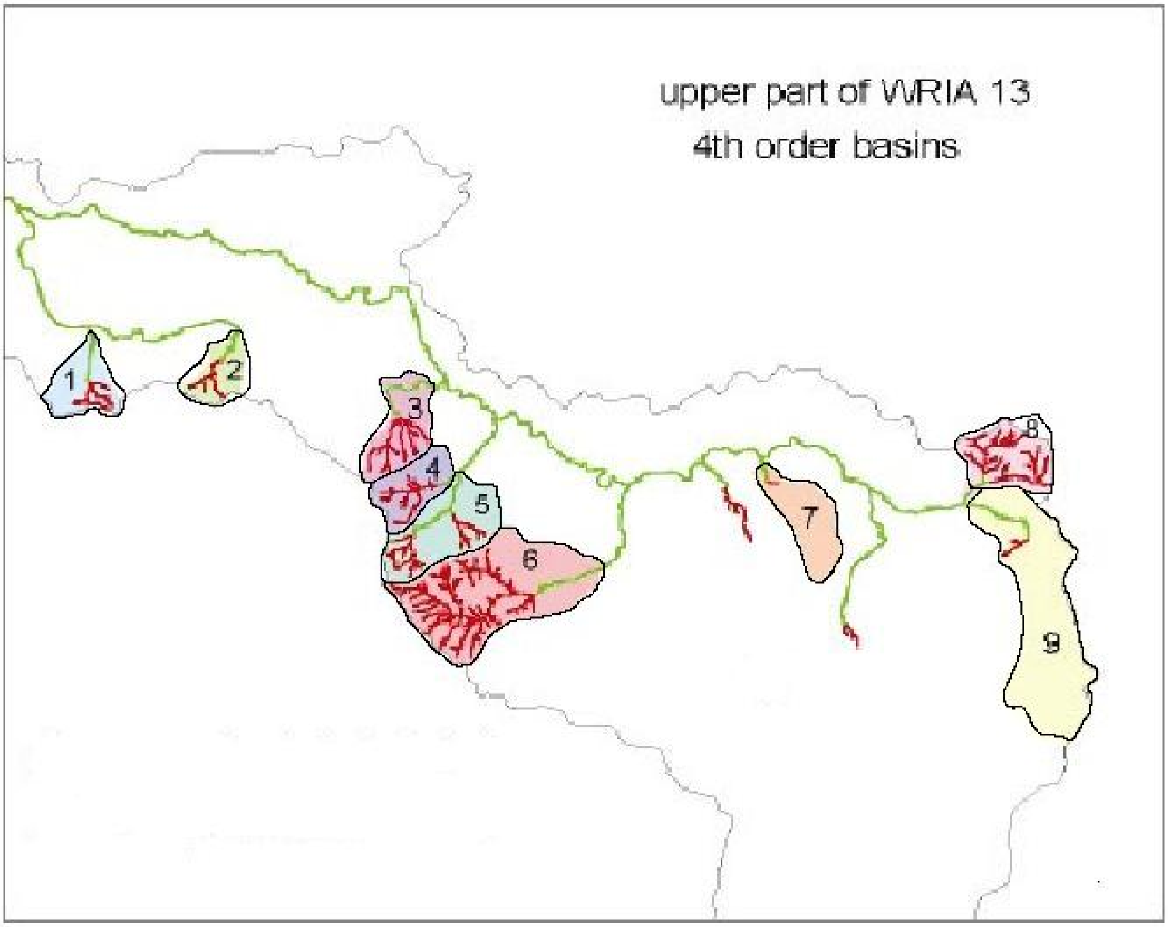


Figure 5 Schematic illustrating the delineation and numbering of fourth order sub-basins (FOSBs) in WRIA 13.

Table 9 Distribution of fourth order sub-basins (FOSBs) among western Washington WRIAs. The numbers of end-of-fish points (EOFPs), fish absent (FA) points, and fish present (FP) points in the FOSBs for each WRIA are shown, also.

WRIA	%[a] of Total	# of FOSBs	Percent of Total	# of EOFPs	Number of Points in FOSBs FA	FP	Total	Percent of Total
1	7.4%	6	0.7%	8	6,456	828	7,284	1.1%
2	0.0%		0.0%					0.0%
3&4	15.3%	17	2.0%	18	30,352	2,431	32,783	4.7%
5	2.9%	38	4.4%	60	18,464	8,678	27,142	3.9%
6	0.9%	2	0.2%	2	3,540	168	3,708	0.5%
7	7.3%	50	5.8%	88	13,132	11,813	24,945	3.6%
8	2.5%	8	0.9%	10	3,828	1,360	5,188	0.8%
9	2.1%	17	2.0%	23	17,688	1,707	19,395	2.8%
10	3.9%	28	3.2%	64	13,138	5,383	18,521	2.7%
11	3.0%	8	0.9%	19	3,339	1,602	4,941	0.7%
12	0.0%		0.0%					0.0%
13	1.0%	9	1.0%	11	5,465	1,148	6,613	1.0%
14	1.2%	4	0.5%	11	767	855	1,622	0.2%
15	2.8%	15	1.7%	34	4,337	3,495	7,832	1.1%
16	1.9%	7	0.8%	17	2,903	1,622	4,525	0.7%
17	1.6%	8	0.9%	14	9,065	1,270	10,335	1.5%
18	2.8%	16	1.8%	17	38,186	278	38,464	5.6%
19	1.3%	9	1.0%	19	2,389	1,572	3,961	0.6%
20	4.2%	62	7.2%	325	25,282	20,987	46,269	6.7%
21	4.4%	53	6.1%	289	20,962	21,253	42,215	6.1%
22	4.7%	115	13.3%	346	34,072	39,830	73,902	10.7%
23	4.3%	148	17.1%	850	86,825	43,006	129,831	18.8%
24	3.2%	117	13.5%	387	36,151	40,143	76,294	11.0%
25	1.5%	5	0.6%	10	1,789	1,048	2,837	0.4%
26	9.7%	84	9.7%	182	42,290	16,674	58,964	8.5%
27	5.0%	27	3.1%	49	21,673	3,750	25,423	3.7%
28	1.9%	6	0.7%	10	10,223	2,148	12,371	1.8%
29	3.5%	8	0.9%	11	5,084	953	6,037	0.9%
TOTALS		867		2,874	457,400	234,002	691,402	

[a] Percentage of total points in the western Washington 10m DEM network with WRIAs 2 and 12 omitted.

FOSBs in each WRIA and compares the proportional representation of the FOSBs across WRIAs to the proportional representation of points by WRIA in the western Washington 10m DEM network. Table 9 also summarizes the number of EOFPs in the FOSBs for each WRIA and the total number of fish absent and fish present points. The proportional distribution of FOSBs, and the FAFP points in them, is similar to that observed previously for all FAFP data. The same WRIAs identified previously are over- and under- represented. These 867 FOSBs contain 2,874 (71%) of the total 4,052 end-of-fish points available. There were from 1 to 29 surveyed end-of-fish points in each FOSB (average of 3.3 EOFPs per FOSB).

Table 10 summarizes the numbers of end-of-fish points, fish absent points, and fish present points belonging to the unassigned group (not in a fourth order sub-basin) in each WRIA. The proportional distribution of FAFP points is again similar to that observed previously for the entire FAFP data set. The same WRIAs identified previously are over- and under-represented. There are 1,178 EOFPs (29% of the total) in the unassigned group.

Selected Approach to Logistic Regression Model Building

There were 1,443,471 FAFP points available for building binary logistic regression models to predict fish absence|presence. We conducted a large number of exploratory LRM building exercises with these data. Based on these exploratory analyses, we determined that the FAFP data used to estimate any "final" logistic regression model(s) needed to incorporate the following:

1. Only FAFP data in fourth order sub-basins should be used for logistic regression model estimation. The reasons for this decision are explained on page 26.

2. A portion of the FAFP should be excluded from both the LRM building process and the stopping rule optimization procedure[12] and reserved for validating the final model(s). This allows for a quasi-independent assessment of model error.

3. Equal numbers of fish absent (FA) and fish present (FP) points should be in the data set used to estimate the logistic regression model coefficients. One guideline identified in the original work assignment for this project was that model error should be "balanced". For the logistic regression models, we interpreted this as meaning that the probability of the model classifying a true FA point as a FP point had to be similar to the probability of the model classifying a true FP point as a FA point. In our preliminary LRM explorations we demonstrated that when there is a large difference in the proportional contributions of FA and FP to the data set used to estimate LRM coefficients, the classification error tended to be better (smaller) for the group (FA or FP) which had the majority of the points in the data set. For example, if the data set used to estimate the LRM coefficients was composed of 70% fish absent points and 30% fish present points, the classification error for the FA points tended to be smaller than the classification error

[12] The stopping rule optimization procedure will be described in a later section of this report.

Table 10 Numbers of end-of-fish points (EOFPs), fish absent (FA) points, and fish present (FP) points belonging to the unassigned group (not in a fourth order sub-basin) in each WRIA.

WRIA	%[a] of Total	# of EOFPs	Percent of Total	Number of Points FA	FP	Total	Percent of Total
1	7.4%	3	0.3%	1,671	16,609	18,280	2.4%
2	0.0%		0.0%				0.0%
3&4	15.3%	18	1.5%	18,235	26,589	44,824	6.0%
5	2.9%	24	2.0%	5,901	23,653	29,554	3.9%
6	0.9%	2	0.2%	482	320	802	0.1%
7	7.3%	35	3.0%	10,520	38,831	49,351	6.6%
8	2.5%	8	0.7%	1,988	5,068	7,056	0.9%
9	2.1%	37	3.1%	5,632	19,371	25,003	3.3%
10	3.9%	51	4.3%	7,945	31,454	39,399	5.2%
11	3.0%	1	0.1%	2,178	17,677	19,855	2.6%
12	0.0%		0.0%				0.0%
13	1.0%	2	0.2%	341	7,887	8,228	1.1%
14	1.2%		0.0%	0	2,475	2,475	0.3%
15	2.8%	8	0.7%	1,005	8,150	9,155	1.2%
16	1.9%	8	0.7%	548	1,656	2,204	0.3%
17	1.6%	12	1.0%	2,251	5,260	7,511	1.0%
18	2.8%	16	1.4%	25,328	8,296	33,624	4.5%
19	1.3%	11	0.9%	1,060	7,168	8,228	1.1%
20	4.2%	110	9.3%	8,620	30,023	38,643	5.1%
21	4.4%	175	14.9%	10,454	23,475	33,929	4.5%
22	4.7%	148	12.6%	9,931	81,070	91,001	12.1%
23	4.3%	296	25.1%	26,460	76,056	102,516	13.6%
24	3.2%	126	10.7%	9,822	56,039	65,861	8.8%
25	1.5%		0.0%	0	7,863	7,863	1.0%
26	9.7%	57	4.8%	11,265	58,521	69,786	9.3%
27	5.0%	27	2.3%	5,303	25,341	30,644	4.1%
28	1.9%	1	0.1%	295	1,527	1,822	0.2%
29	3.5%	2	0.2%	851	3,604	4,455	0.6%
TOTALS		1,178		168,086	583,983	752,069	

[a] Percentage of total points in the western Washington 10m DEM network with WRIAs 2 and 12 omitted.

for the FP points. When equal numbers of FA and FP points were present in the data set used to estimate the LRM coefficients the classification errors for the two groups (FA points and FP points) tended to be more similar. Therefore, the data sets we used to estimate LRM coefficients all had a 50:50 composition of fish absent and fish present points.

4. <u>The proportional representation of FAFP points in the data set used to estimate the logistic regression model coefficients should be approximately equal to the proportional composition of points by WRIA in the western Washington 10m DEM network.</u> One goal of the project was to develop a single LR model that could be applied to western Washington. Since the model encompasses such a large geographic area it is important that FAFP data representing the entire region are present in the data set used for estimating model coefficients. This helps to insure that the variation in the five GIS-derived physical attributes (BASIZE, ELEV, UPGRD, DNGRD, and PRECIP) across the western Washington landscape is captured. However, because we are concerned with overall model classification accuracy we would like the model to perform well in the WRIAs which have the largest number of points in the 10m DEM network (the points that will eventually be classified by the model). Therefore, we want those WRIA with large numbers of network points to have a greater influence on the estimation of the model coefficients than WRIAs with smaller numbers of network points. We decided that an approach that randomly sampled FAFP points from WRIAs proportionally to their representation in the western Washington 10m DEM network was a good method for addressing these issues.

5. <u>The number of FAFP points used to estimate the logistic regression model coefficients should be less than the maximum number of available points given the previous guidelines.</u> A basic assumption of binary logistic regression analysis is that the observations (the FAFP data points in our application) are statistically independent (Hosmer and Lemeshow 1989). We expect that the FAFP data are spatially autocorrelated because neighboring points on the 10m DEM network are likely to be more similar in the values for the five GIS-derived physical attributes than are points separated by a greater distance on the DEM network. The values of the attributes for FA or FP points are usually similar when they are from the same FOSB, are more different when they are from a different FOSB in the same WRIA, and are even more different when compared to data from a FOSB in a different WRIA. Knapp and Preisler (1999) discuss some of the statistical implications of data that are spatially autocorrelated with respect to a logistic regression problem similar to ours. A major implication of using data that have a high spatial autocorrelation is that the standard errors of the logistic regression coefficients will be under-estimated and the significance of the model coefficients will be over-estimated[13]. Ideally we might prefer to randomly draw one FA point and one FP point from each sub-basin. Unfortunately we do not have a sufficient number of fourth order sub-basins for that approach. Therefore we drew multiple points from the FOSBs. The more points (either FA or FP) we draw from a single FOSB the

[13] An over-estimation of the significance means that the significance of the model or model coefficient(s) is estimated to be more significant (smaller) than it actually should be.

greater the risk of introducing spatial autocorrelation. We investigated a range of FAFP point sample sizes for the LR models we estimated.

6. Only FAFP data generated from EOFPs with end types 1 or 2 should be used to estimate the coefficients of the logistic regression model. A reason for the placement of each end-of-fish point was recorded for each EOFP (see Table 2). The possible reasons for placement of the EOFP were:
    i. Natural end,
    ii. Gradient related,
    iii. Large woody debris,
    iv. Road culvert,
    v. Mass-wasting event,
    vi. Beaver dam or other non-permanent dam,
    vii. Permanent dam,
    viii. Water quality limiter, or
    ix. Other or unknown.

   The logistic regression model predicts fish absence | presence based on the values of the five physical attributes (basin size, elevation, upstream gradient, downstream gradient, and precipitation) derived from the 10m GIS hydro-layer. There is an assumption that, on average, there are differences in the five attributes between FA and FP points. End types 1 and 2 can be considered "natural" end types, i.e., the reason there was a change from fish presence to fish absence on the stream is that the stream became too small or the gradient too steep. The remaining EOFP end types can be considered "unnatural" end types, i.e., the location of these EOFPs is unlikely to be related to the values of the five physical attributes. While beaver dams, large woody debris blockages, and mass-wasting events might be considered "natural" events, their occurrence on the landscape is probably not related to the five physical attributes in the same way as fish absence and presence. Therefore, we decided to use only FAFP data generated from EOFPs with natural or gradient-related end types to estimate the final logistic regression models.

Summary of the FAFP Data Pool:

The data pool from which we drew FAFP points to estimate logistic regression models only included data generated from end-of-fish points with end types 1 (natural) or 2 (gradient related). In addition, only FAFP points that were part of a fourth order sub-basin delineated on the 10m DEM network were included in this FAFP data pool. Some FOSBs contained one or more EOFPs with end types 1 or 2 and other EOFPs with end types other than 1 or 2. In this situation, we could not easily determine which FAFP points in the FOSB were generated from the EOFP(s) with end types 1 or 2 and which were generated from the EOFP(s) with other end types. Therefore, we did not include these FOSBs in the final FAFP data pool but reserved them for the LR model validation process. FAFP data for the data pool were restricted to FOSBs which contained one or more EOFPs with end types 1 or 2 and did not contain EOFPs with other end types.

After screening the 867 original FOSBs (Table 9), there were 573 FOSBs which only contained EOFPs with end types 1 or 2. Similar to earlier summaries (Tables 8, 9, or 10), WRIAs 1 and 3&4 combined were under-represented in the FAFP data pool while WRIAs 22, 23, and 24 were over-represented relative to the proportion of total network points contributed by these WRIA to the western Washington DEM network (Table 11). Generally, the proportional contribution of the other WRIAs to the FAFP data pool was similar to their representation in the 10m DEM network (the proportional contribution of each of the other WRIAs to the data pool was within 3.5% of the WRIAís contribution to the 10m DEM network). There were 400,679 FAFP data points in the FAFP data pool; 270,163 (67%) fish absent points and 130,516 (33%) fish present points.

Most of the EOFPs in the FAFP data pool (78%) were Last Fish Habitat points (Table 12). Field surveys for the EOFPs in the FAFP data pool were conducted between January 1990 and October 2000. The majority (66%) of the EOFPs were collected during field surveys conducted in 1995, 1996, and 1997. Less than 1% of the EOFPs came from surveys conducted prior to 1994. All three boundary types are well represented in the FAFP data pool (Table 12): there were more than 400 EOFPs from each boundary type. More than half (58%) of the EOFPs were determined by electro-shocking during field surveys (Table 13). Visual determination of the EOFP was used for all but three of the remaining EOFPs in the FAFP data pool.

This resulted in the original FAFP data being partitioned into three groups:

1. FAFP data generated from an end-of-fish point with end type 1 or 2 and belonging to a fourth order sub-basin which has only EOFPs of those end types (<u>FAFP data pool</u>),
2. FAFP data generated from EOFPs with all possible end types (including EOFPs with end types 1 or 2 that were in FOSBs which included at least one EOFP of another end type) and belonging to a FOSB (<u>model validation data</u>), and
3. Unassigned FAFP data that are not part of a fourth order sub-basin and have been generated from EOFPs with all possible end types (<u>unassigned data</u>).

Figures 6 and 7 use box-and whiskers plots to compare the distributions of the values for four of the five physical attributes (BASIZE, ELEV, DNGRD, and PRECIP) for fish absent and fish present points belonging to each of the three groups. Each box-and-whiskers plot encompasses the central quartiles of the data (the central 50% of the data values) in the shaded box. The median value is indicated by the heavy line bisecting the box. The box whiskers include all data values not considered outliers or extreme values (all data within 1.5 box lengths of the edge of the central quartile box): outliers or extreme values are not shown on these plots to increase legibility. Log_{10} of basin size is shown to compress the scale and because $\log_{10}$ transformed basin size data are used in the logistic regression analysis. Fish present points generally have much larger basin sizes and are at lower elevations than fish absent points (Figure 6). For the fish present points, the basin sizes for the unassigned group are clearly much larger than for data from fourth order sub-basins. Fish present points generally have lower downstream gradients than fish absent points (Figure 7). There are not large differences among the groups in the distribution of precipitation values (Figure 7). Appendix Tables 2, 3, 4, 5, and 6 summarize the mean, standard error, median, and range for each of the five physical attributes for the three groups defined above.

Table 11 Number of fourth order sub-basins (FOSBs) containing end-of-fish points (EOFPs) with end types 1 or 2 only, by WRIA. Number of EOFPs, fish absent (FA) points, and fish present (FP) points belonging to each WRIA are summarized, also.

WRIA	%[a] of Total	# of FOSBs	% of Total	# of EOFPs	Number of Points in FOSBs FA	FP	Total	Percent of Total
1	7.4%	4	0.7%	4	4,566	315	4,881	1.2%
2	0.0%		0.0%					0.0%
3&4	15.3%	17	3.0%	18	20,767	2,431	23,198	5.8%
5	2.9%	34	5.9%	55	16,929	7,782	24,711	6.2%
6	0.9%	1	0.2%	1	1,341	168	1,509	0.4%
7	7.3%	38	6.6%	54	8,414	7,741	16,155	4.0%
8	2.5%	6	1.0%	8	2,863	1,137	4,000	1.0%
9	2.1%	14	2.4%	19	15,489	1,470	16,959	4.2%
10	3.9%	19	3.3%	31	7,575	2,872	10,447	2.6%
11	3.0%	6	1.0%	9	2,496	955	3,451	0.9%
12	0.0%		0.0%					0.0%
13	1.0%	6	1.0%	7	4,296	864	5,160	1.3%
14	1.2%	1	0.2%	2	117	118	235	0.1%
15	2.8%	9	1.6%	13	2,634	1,734	4,368	1.1%
16	1.9%	6	1.0%	14	2,791	1,235	4,026	1.0%
17	1.6%	6	1.0%	8	5,304	749	6,053	1.5%
18	2.8%	11	1.9%	11	20,333	139	20,472	5.1%
19	1.3%	8	1.4%	17	2,233	1,400	3,633	0.9%
20	4.2%	31	5.4%	112	9,183	7,952	17,135	4.3%
21	4.4%	28	4.9%	88	8,945	7,488	16,433	4.1%
22	4.7%	78	13.6%	185	20,190	23,079	43,269	10.8%
23	4.3%	76	13.3%	383	40,434	19,063	59,497	14.8%
24	3.2%	83	14.5%	191	19,890	25,061	44,951	11.2%
25	1.5%	4	0.7%	9	1,527	914	2,441	0.6%
26	9.7%	55	9.6%	114	24,625	10,195	34,820	8.7%
27	5.0%	22	3.8%	41	15,342	3,142	18,484	4.6%
28	1.9%	6	1.0%	10	10,223	2,148	12,371	3.1%
29	3.5%	4	0.7%	4	1,656	364	2,020	0.5%
TOTALS		573		1,408	270,163	130,516	400,679	

[a] Percentage of total points in the western Washington 10m DEM network with WRIAs 2 and 12 omitted.

Table 12 Summary of the number (#) of end-of-fish points (EOFPs) in the FAFP data pool belonging to each boundary type, by point type (% is the percentage of the total number of EOFPs).

End-of-Fish Point Type		Boundary Type[a]			Total
		A	B	C	
Last Fish	#	69	50	34	153
	%	4.9%	3.6%	2.4%	10.9%
Last Fish Habitat	#	296	442	356	1,094
	%	21.0%	31.4%	25.3%	77.7%
Last Salmonid	#	43	24	24	91
	%	3.1%	1.7%	1.7%	6.5%
Last Salmonid Habitat	#	38	19	13	70
	%	2.7%	1.3%	0.9%	5.0%
TOTALS	#	446	535	427	1,408
	%	31.7%	38.0%	30.3%	100.0%

[a] A = mid-channel boundary; B = confluence point where a nonfish-bearing stream laterally intersects a fish-bearing stream; and C = tributary junction where two or more nonfish-bearing streams join to form a fish-bearing stream.

Table 13 Summary of the number (#) of end-of-fish points (EOFPs) in the FAFP data pool belonging to each determination method, by point type (% is the percentage of the total number of EOFPs).

End-of-Fish Point Type		Determination Method			Total
		Electro-shocking	Visual	Other	
Last Fish	#	123	30		153
	%	8.7%	2.1%		10.9%
Last Fish Habitat	#	573	518	3	1,094
	%	40.7%	36.8%	0.2%	77.7%
Last Salmonid	#	65	26		91
	%	4.6%	1.8%		6.5%
Last Salmonid Habitat	#	51	19		70
	%	3.6%	1.3%		5.0%
TOTALS	#	812	593	3	1,408
	%	57.7%	42.1%	0.2%	100.0%

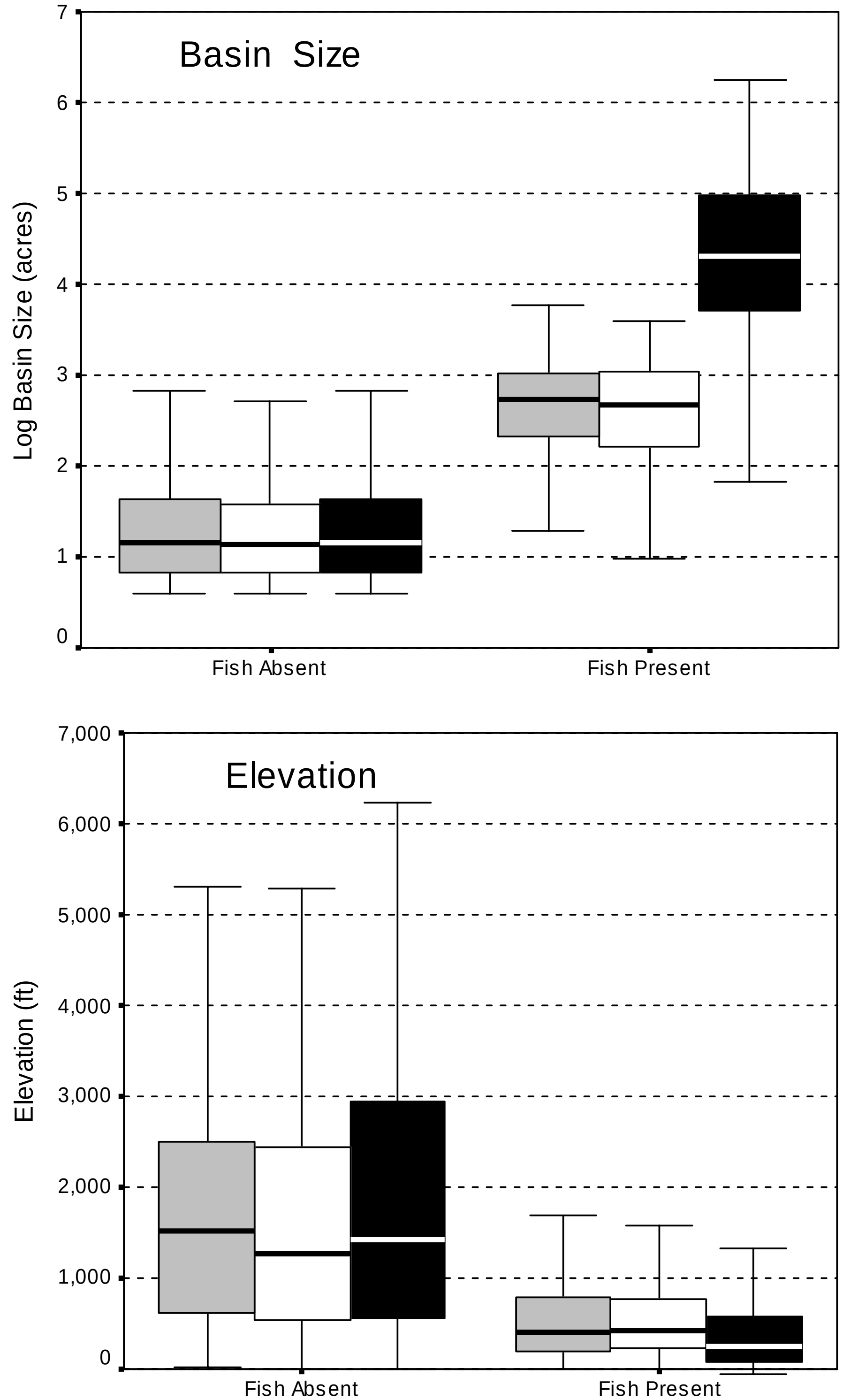


Figure 6 Box-and-whisker plots comparing the distributions of log of basin size and elevation values for fish absent and fish present points, by summary group: = fourth order sub-basin (FOSB) model data, = FOSB validation data, and = unassigned data.

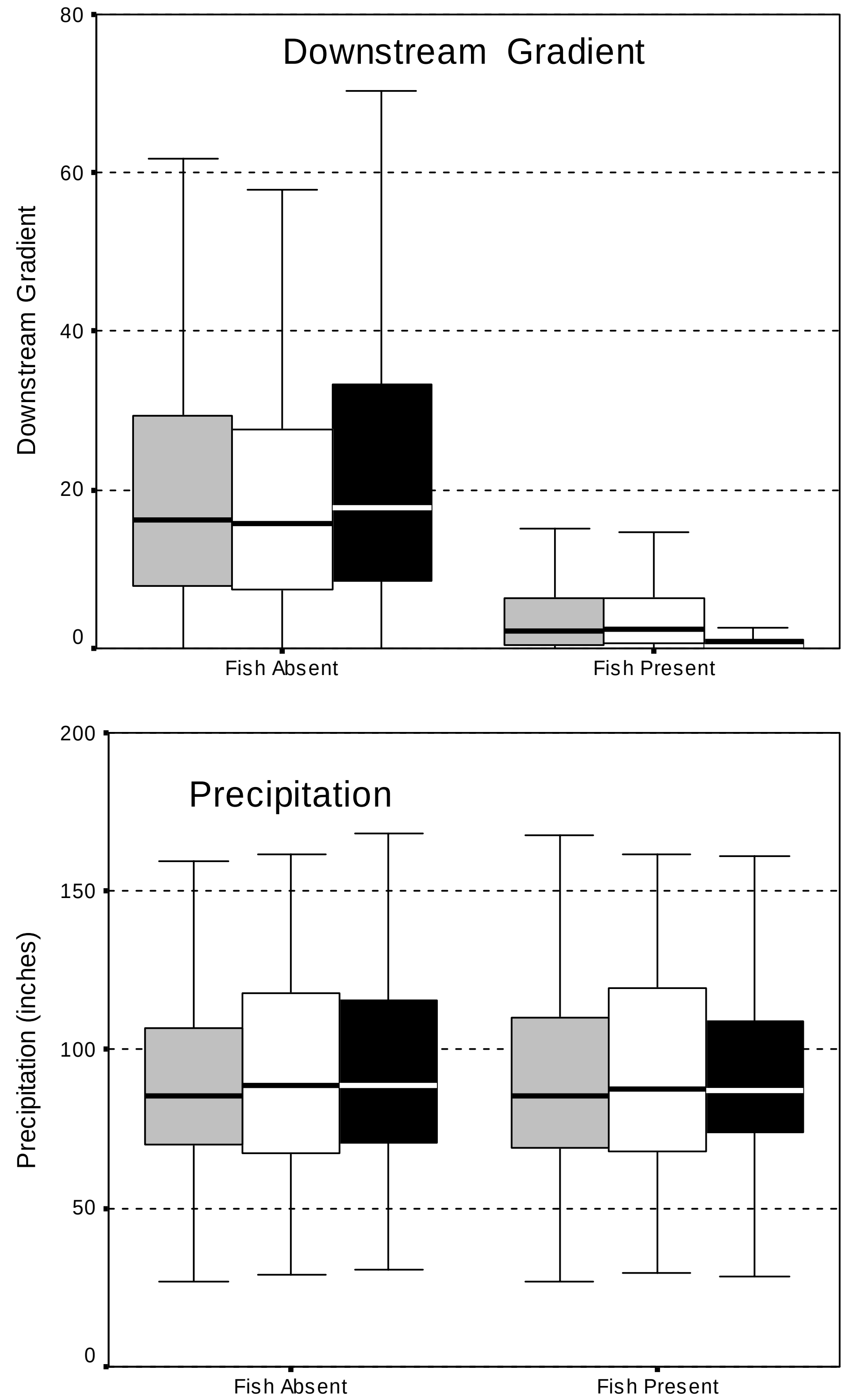


Figure 7 Box-and-whisker plots comparing the distributions of downstream gradient and precipitation values for fish absent and fish present points, by summary group:

▢ = fourth order sub-basin (FOSB) model data, ▢ = FOSB validation data, and ■ = unassigned data.

Construction of Data Sets for Estimating Logistic Regression Models:

We investigated a range of sample sizes for the data sets used to estimate the final logistic regression models. FAFP data selected to estimate a logistic regression model are subsequently referred to as a model estimation data set. For each model estimation data set, equal numbers of FA and FP points were randomly selected from each WRIA proportionally to that WRIAís contribution of points to the western Washington 10m DEM network.

As an example, for a model estimation data set with 6,000 points there should be 3,000 FA points and 3,000 FP points (see item 3 under Selected Approach to Logistic Regression Model Building). Therefore, about 2.9% of the points for each group (FA or FP) should be from WRIA 5 (Table 11). This results in a target sample size of 88 points for WRIA 5 (using percentages calculated out to five decimal points and rounding the target number of points to a whole number). So for this model estimation data set, 88 FA points were randomly selected from the 16,929 FA points available in WRIA 5 (Table 11) and 88 FP points were randomly selected from the 7,782 FP points available in WRIA 5 (Table 11). This procedure was followed for each WRIA.

For the two largest sample sizes examined (10,000 and 12,000 total points), there were three occasions when a WRIA did not have sufficient FP points available to achieve the desired sample size for the WRIA. When this occurred the deficit in points was made up by randomly selecting additional points from a neighboring WRIA. A relatively small number of points were involved in this process (less than 3% of the target sample size for the FP group). The numbers of data points from each WRIA for each model estimation data set sample size are given in Appendix Table 7.

Evaluation of Sample Sizes for Logistic Regression Model Estimation:

Compared to the total number of FAFP points available for estimating the logistic regression models, we wanted to use a relatively small number of points to estimate the models to reduce the possible effects of autocorrelated data (see item 5 under Selected Approach to Logistic Regression Model Building). Based upon our preliminary logistic regression model building exercises, we determined that at least 2,000 FA points and 2,000 FP points were needed to provide a good representation of all WRIAs in the model estimation data set plus provide estimates of model coefficients that were not highly variable for different random draws from the FAFP data pool. Therefore, we examined five possible sample sizes: 2,000 points for each group (FA points and FP points); 3,000 points per group; 4,000 points per group; 5,000 points per group; and 6,000 points per group.

For each sample size, we selected 10 different model estimation data sets following the procedures described in the previous section. Once a model estimation data set had been selected, the procedures described earlier in the report (see Model Estimation and Variable Selection section) were used to select the physical attributes to be included in the logistic regression model, estimate the logistic regression model coefficients, estimate the model fit statistics, and assess model performance. For these analyses we transformed two of the physical attributes prior to analysis. The $\log_{10}$ of BASIZE was used because of the extremely large range of BASIZE values in the data (4 to 5,797 acres). We also transformed ELEV by dividing it by 100 so that the estimated ELEV coefficient was similar in size to the other coefficients in the model.

We used the 290,723 FAFP data points in the model validation data set described earlier to assess model performance in addition to model self-classification accuracy (Table 9 FAFP data ñ Table 11 FAFP data). The results for each model were recorded. Means and coefficients of variation[14] were used to summarize the results for a sample size and to compare the results among the different sample sizes.

[14] The coefficient of variation (CV) is the sample standard deviation divided by the sample mean. It is typically expressed as a percentage.

RESULTS

The results are summarized in two sections:

1. a summary of the results of the sample size evaluation analyses, and
2. a summary of the final logistic regression model estimated.

Evaluation of Sample Sizes for Logistic Regression Model Estimation

Appendix Tables 8a, 8b, 8c, 8d, and 8e present the results for the ten logistic regression models estimated for each of the five sample sizes examined. There is very little difference in the mean model assessment statistics for self-classification accuracy or accuracy estimated from the validation data set among the different sample sizes. For self-classification accuracy, the FP points had a slightly higher mean classification accuracy than FA points. Conversely for the validation data set, FA points had a slightly higher mean classification accuracy than FP points. All classification accuracies were between 87% and 94% regardless of method (self-classification or validation data) or estimation data set sample size. The estimation data sets with 2,000, 3,000, and 4,000 FA and FP points had lower mean values for the Hosmer-Lemeshow (H-L) statistic than the models with 5,000 and 6,000 points. The majority of the models based on the two largest estimation data set sample sizes had significant H-L statistics. Only 2 to 4 of the ten models for each of the smaller sample sizes had significant H-L statistics (non-significance is one indicator of good model fit). In contrast, 8 to 10 of the ten models for each of the two largest sample sizes had significant H-L statistics.

Within the 10 model runs for a specific sample size there is generally little variation in the model constant and the estimated coefficients for the two most important model attributes BASIZE and ELEV. BASIZE was selected by the stepwise variable selection procedure as the first physical attribute entered into the LR model and ELEV as the second attribute for 49 of the 50 models estimated. In one model BASIZE was followed by DNGRD. For 43 of the 50 LR models estimated for the sample size evaluation, the final LR model selected only four of the attributes: BASIZE, ELEV, DNGRD, and PRECIP. There were six models which included UPGRD in addition to the other four attributes. As expected, the coefficients of variation for the means of the model coefficients generally decreased as the sample size for the estimation data set increased.

We decided that a sample size of 4,000 points per group provided an estimation data set that:

- Used a sufficiently small fraction of the FAFP data in each FOSB so that the potential effects of using autocorrelated data to estimate the model coefficients were minimized, and
- Provided estimates of LR model coefficients that were relatively consistent for different random draws from the data.

Four thousand data points represent only 0.64% of the total number of fish absent data points available and only 0.49% of all fish present points available (totals from Tables 9 and 10). When the data are limited to FAFP data generated from EOFPs with end types 1 or 2 that are contained in FOSBs, 4,000 points represents 1.48% of the FA data and 3.06% of the FP data. Table 14 summarizes the percentages of FA and FP points available represented by the sample size target for each WRIA with an estimation data set of 4,000 total points per group (Appendix Table 7). The median percentage of FA points used across the 26 WRIAs was only 2.0% (mean 3.9 %). Because there are fewer FP points available in FOSBs (relative to FA points), a higher fraction of the FP points were used per WRIA (median = 5.9% and mean = 15.1%).

With a sample size of 4,000 FA and 4,000 FP points, the CVs for the model constant and two most important attributes (BASIZE and ELEV) were both less than 5% and the CVs for the other two attributes entered into the model (DNGRD and PRECIP) were both less than 10%. With sample sizes less than 4,000 per group, the CV for ELEV was greater than 5% and the CVs for DNGRD and PRECIP were both greater than 14%. The CVs for the model constant, BASIZE, and ELEV decreased slightly but the CVs for DNGRD and PRECIP were generally greater with sample sizes greater than 4,000 per group. This was probably due to the presence of some models in the set of 10 which included the UPGRD physical attribute. Also, a sample size of 4,000 points did not require "borrowing" from adjacent WRIA to meet the sample size target for any of the WRIA. Each WRIA had a sufficient number of FA and FP points for the target sample size.

Final Logistic Regression Model Coefficients and Model Fit and Assessment Statistics

We drew one final random selection of 4,000 FA and 4,000 FP points from the FAFP data pool with the sample size targets for each WRIA specified in Table 14. These data were then used to estimate a final logistic regression model. This final LR model was similar to the ten models estimated for the sample size evaluation analysis. The final model selected BASIZE, ELEV, DNGRD, and PRECIP. Table 15 summarizes the coefficient values, standard errors, and 95% confidence interval for each model coefficient[15]. The Hosmer-Lemeshow chi-square statistic for this model was 12.832 (P = 0.110) and the -2LL statistic was 3,172.5. Self-classification accuracies for this model were 90.7% for FA points and 93.0% for FP points.

[15] The 95% confidence interval is actually for the exponential of the coefficient since the probability estimated by the logistic regression is expressed as $\frac{1}{1+e^{-(B_0+B_1X)}}$.

Table 14 Percentage of available fish absent and fish present points represented by the final target sample size for each WRIA. Target sample size is based on a desired total of 4,000 data points for each group. Percentages are based on total fish absent or fish present points generated from EOFPs with end types of 1 or 2 and contained in fourth order sub-basins (from Table 11).

WRIA	Target Sample Size	Percent of Points Used	
		Fish Absent	Fish Present
1	295	6.5%	93.7%
3&4	611	2.9%	25.1%
5	116	0.7%	1.5%
6	37	2.8%	22.0%
7	293	3.5%	3.8%
8	99	3.5%	8.7%
9	83	0.5%	5.6%
10	157	2.1%	5.5%
11	118	4.7%	12.4%
13	39	0.9%	4.5%
14	47	40.2%	39.8%
15	110	4.2%	6.3%
16	76	2.7%	6.2%
17	63	1.2%	8.4%
18	110	0.5%	79.1%
19	50	2.2%	3.6%
20	168	1.8%	2.1%
21	177	2.0%	2.4%
22	186	0.9%	0.8%
23	174	0.4%	0.9%
24	126	0.6%	0.5%
25	61	4.0%	6.7%
26	387	1.6%	3.8%
27	198	1.3%	6.3%
28	77	0.8%	3.6%
29	142	8.6%	39.0%
	Mean	3.9%	15.1%
	Median	2.0%	5.9%
	Minimum	0.4%	0.5%
	Maximum	40.2%	93.7%

Table 15 Summary of the final logistic regression model coefficients, standard errors, significance of the coefficients, and 95% confidence intervals for the exponential of the coefficients.

Coefficient	Estimate	Standard Error	Signifi-cance	e^B	95% Confidence Interval
Constant	-7.717073	0.25192	<0.001		
Log_{10} (BASIZE)	3.793994	0.09625	<0.001	44.434	36.794 - 53.657
ELEV/100	-0.110926	0.00687	<0.001	0.895	0.883 - 0.907
DNGRD	-0.062949	0.00656	<0.001	0.939	0.927 - 0.951
PRECIP	0.020166	0.00191	<0.001	1.020	1.017 - 1.024

We examined the classification accuracies for the validation data in more detail. We applied the final LR model to the validation data reserved for model assessment and the unassigned data not used in the model estimation process (FAFP data not in FOSBs). We then calculated the percentage of points correctly classified, by WRIA, for the FA and FP points in each group (Appendix Table 9). Figure 8 summarizes the results for FA points (top graph) and FP points (bottom graph).

For the model validation data, the mean classification accuracy for fish absent points in FOSBs across the 25 WRIAs with data was 92.6% (range: 82.7% to 100.0%). The lowest accuracy (82.7%) was for WRIA 8; the classification accuracy for FA points was greater than 95% for nine of the 25 WRIAs. For fish present points in FOSBs, the mean classification accuracy across the 23 WRIAs with data was 89.8% (range: 48.1% to 100.0%). The lowest accuracy (48.1%) was for WRIA 9; the classification accuracy for FP points was greater than 95% for seven of the 23 WRIAs.

For the model validation data, the mean classification accuracy for unassigned fish absent points across the 24 WRIAs with data was 90.8% (range: 65.6% to 100.0%). The lowest accuracy (65.6%) was for WRIA 6; the classification accuracy for FA points was greater than 95% for six of the 24 WRIAs. For unassigned fish present points, the mean classification accuracy across the 26 WRIAs with data was 98.6% (range: 91.7% to 100.0%). The lowest accuracy (91.7%) was for WRIA 21; the classification accuracy for FP points was greater than 95% for 25 of the 26 WRIAs.

The mean classification accuracy was similar for fish absent points in fourth order sub-basins and unassigned FA points (Appendix Table 9). Fish present points in FOSBs had the lowest mean classification accuracy (89.8%) while FP points in the unassigned group had the highest mean classification accuracy (98.6%).

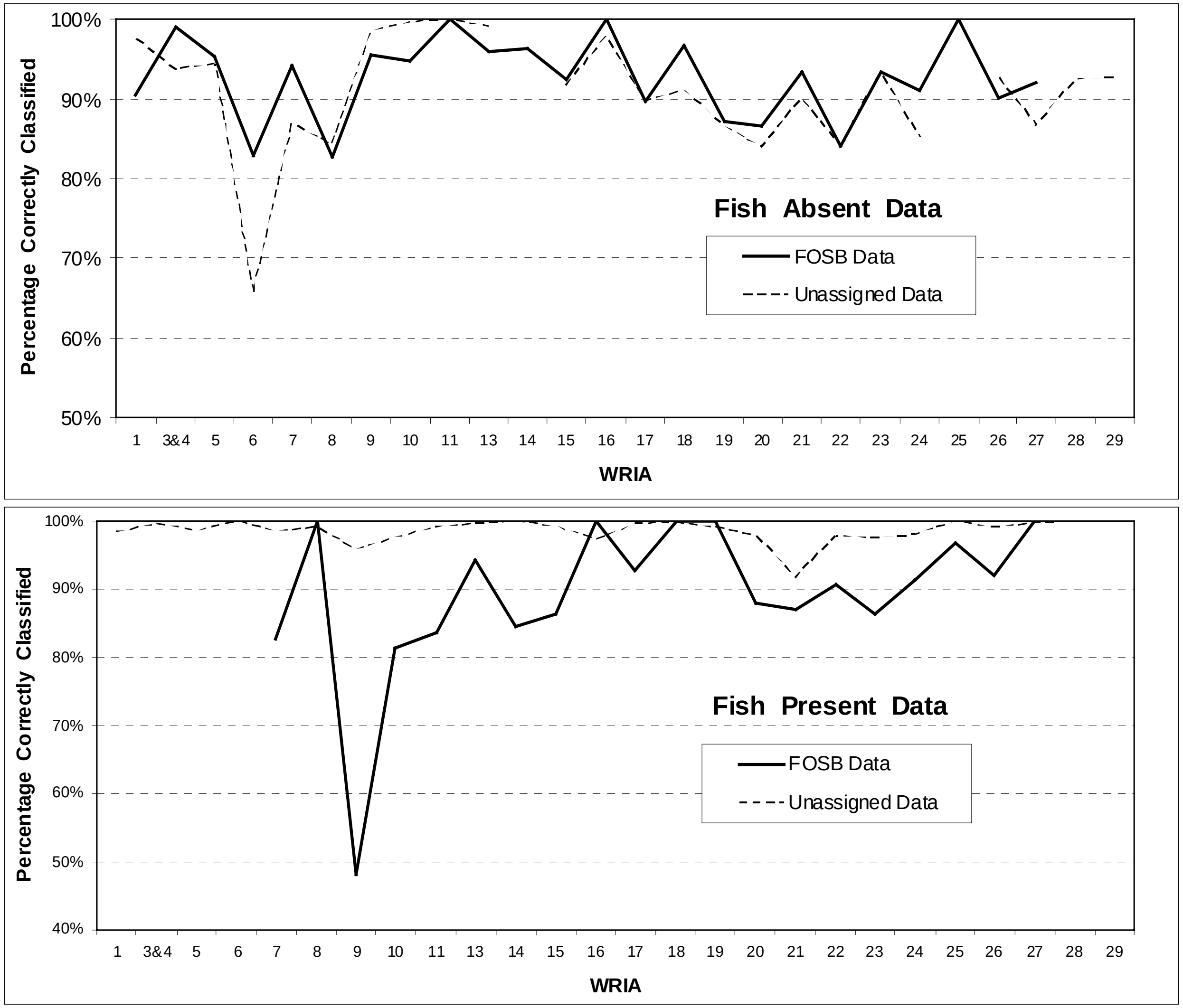


Figure 8 Final logistic regression model classification accuracies for model validation data, by WRIA (summarized by fish absent and fish present points in fourth order sub-basins [FOSB] and for unassigned points).

Examining the results of classifying the end-of-fish points in the model validation data provides additional information on some of the possible influences affecting on the performance of the logistic regression model. The model should classify all EOFPs as fish absent points. We examined the classification accuracies for EOFPs with different end types and different boundary types to determine if there were differences among the categories in these groups.

Table 16 summarizes EOFP classification accuracies for two end type groups, (1) end types 1 and 2 combined and (2) all other end types. This was done separately for EOFPs in fourth order sub-basins and for those in the unassigned group. For EOFPs in FOSBs, 82% of the EOFPs with end types 1 or 2 were correctly classified[16] as fish absent points but only 60% of the EOFPs with other end types were correctly classified. For EOFPs in unassigned group, 73% of the EOFPs with end types 1 or 2 were correctly classified as fish absent points but only 59% of the EOFPs with other end types were correctly classified.

Table 16 Final logistic regression model classification accuracies for end-of-fish points (EOFPs) in the model validation data summarized by end type group for EOFPs in fourth order sub-basins and unassigned EOFPs (all EOFPs should be " Predicted Absent").

Data Set	End Type Group		Predicted Absent	Predicted Present	Total
Fourth Order Sub-basin Data	End type 1 or 2	Number	858	187	1,045
		Percentage	82.1%	17.9%	100.0%
	All other end types	Number	254	167	421
		Percentage	60.3%	39.7%	100.0%
	TOTALS	Number	1,112	354	1,466
		Percentage	75.9%	24.1%	100.0%
Unassigned Data	End type 1 or 2	Number	741	273	1,014
		Percentage	73.1%	26.9%	100.0%
	All other end types	Number	117	82	199
		Percentage	58.8%	41.2%	100.0%
	TOTALS	Number	858	355	1,213
		Percentage	70.7%	29.3%	100.0%

Table 17 summarizes EOFP classification accuracies by boundary type for EOFPs in fourth order sub-basins and for those in the unassigned group. Classification accuracies for the EOFPs in FOSBs are slightly higher than for the unassigned group. Lateral confluence EOFPs have the highest classification accuracies in both data sets. The classification accuracies for the other two boundary types (mid-channel and tributary junction) are similar to each other.

[16] End-of-fish points were all attributed as fish absent points in the database.

Table 17 Final logistic regression model classification accuracies for end-of-fish points (EOFPs) in the model validation data summarized by boundary type for EOFPs in fourth order sub-basins and unassigned EOFPs (all EOFPs should be "Predicted Absent").

Data Set	Boundary Type[a]		Predicted Absent	Predicted Present	Total
Fourth Order Sub-basin Data	Mid-channel	Number	345	149	494
		Percentage	69.8%	30.2%	100.0%
	Lateral confluence	Number	562	105	667
		Percentage	84.3%	15.7%	100.0%
	Tributary junction	Number	205	100	305
		Percentage	67.2%	32.8%	100.0%
	TOTALS	Number	1,112	354	1,466
		Percentage	75.9%	24.1%	100.0%
Unassigned Data	Mid-channel	Number	288	144	432
		Percentage	66.7%	33.3%	100.0%
	Lateral confluence	Number	386	116	502
		Percentage	76.9%	23.1%	100.0%
	Tributary junction	Number	184	85	279
		Percentage	65.9%	34.1%	100.0%
	TOTALS	Number	858	355	1,213
		Percentage	70.7%	29.3%	100.0%

[a] Lateral confluence is a non fish-bearing stream laterally intersecting a fish-bearing stream and a tributary junction is two or more non fish-bearing streams joining to form a fish-bearing stream.

DISCUSSION

For an overall assessment of the logistic regression model and its usefulness in providing a basis for estimating the placement of end-of-fish points on a GIS network, we focus the discussion of the major assumptions on which the model is based and model performance.

Major Model Assumptions

A major assumption for the development of the model is that the FAFP points used in the logistic regression model building process are "representative" of fish absent and fish present locations across all stream systems in western Washington. We selected the data for model building from a pool of FAFP points generated from field surveys conducted in 27 of the 29 western Washington WRIA. The selection process we used to select FAFP data for model building ensured that points were selected from each WRIA approximately proportional to their representation in the western Washington 10m DEM network. However, if the FAFP points in the data pool for a WRIA were not representative of the WRIA this could be a potential source of model error. When additional field survey data is collected the sample design needs to ensure that a random sample of the streams in the WRIA is being collected. These data can be compared to the data used to build the models for this report and used to update any future LR models developed.

It is also assumed that the 10m DEM network accurately defines the location of stream systems on the ground and that the five physical attributes associated with each point in the DEM network accurately represent the conditions at the location. Errors in either of these are another potential source of model error but can only be addressed through the development of more accurate GIS systems.

An assumption of the logistic regression model building process is that the FAFP data used to estimate the logistic regression are independent. The points on the 10m DEM network which define a stream are obviously not independent. The closer the location of the points on the network, the more highly correlated the values for the five physical attributes. We tried to minimize the data dependence by randomly sampling the points in the FAFP data pool for a WRIA and selecting, in most cases, only a small proportion of the fish absent or fish present points available. For some WRIAs, a relatively large proportion of either the FA or FP points were sampled: in six WRIAs more than 25% of the available points were sampled. As more field data are collected in these WRIAs it can be added to their FAFP data pools and used to decrease the proportion of available points used to build future logistic regression models. Violations of the independent data assumption result primarily in an overestimation of the precision of the estimated logistic regression model coefficients.

Model Performance

The assessment of logistic regression model performance should not be limited exclusively to classification accuracy based on known FAFP points in the western Washington 10m DEM network. Caution is needed when basing any conclusions on model performance on a comparison of model classification accuracy to a target standard, whether that standard is 90%, 95%, or some other percentage. Overall classification accuracy is influenced by the location of the FAFP points on the stream network and the type of end-of-fish point being assessed. In general, the classification accuracy results supported our expectations. Mainstem FAFP data[17], which are characterized by large basin sizes, low elevations, and low gradients (see Appendix Tables 2, 3, and 4), are generally " gimmes" and are rarely classified incorrectly (98.6% mean classification accuracy across all 26 western Washington WRIAs for fish present points in the unassigned data set, Appendix Table 9). Approximately 46% of the 31,078,161 points in the western Washington 10m DEM network are outside of FOSBs and the majority of these points are mainstem points, therefore, we expect they will rarely be misclassified. An overall estimate of model classification accuracy will be heavily influenced by the number of mainstem points present in any data set used to assess model performance.

Also, our analyses indicate there are other factors associated with end-of-fish points which affect the performance of the model. The final LR model classifies EOFPs with end types 1 or 2 (naturally occurring end types) more accurately than EOFPs with other end types (Table 16). This supports our decision to restrict the data used to generate the LR model to FAFP data generated from EOFPs with end types 1 or 2. The model also classifies EOFPs associated with lateral confluence points (where a non fish-bearing stream laterally intersects a fish-bearing stream) with higher accuracy than EOFPs with other boundary types (Table 17). This again demonstrates how the overall model classification accuracy can be influenced by the composition of any data set used to assess model performance.

The final LR model assessment results indicate that the model performs well in most WRIAs (Appendix Table 9). However, it is important to remember that the validation FAFP data are not a random sample from the WRIA so we must be cautious in our interpretation of these results. The two WRIAs with very low classification accuracies (< 80%) are associated with relatively small sample sizes (< 500 points) and very few EOFPs (2 to 4). There were a number of other WRIAs where the number of points used to estimate classification accuracy for either FA or FP points was less than 1,000 points (Appendix Table 10). The FAFP data used to assess classification accuracy may not be representative of the WRIA and, therefore, may not be truly indicative of model performance for that WRIA.

[17] Mainstem data are a subset of the unassigned data. The majority of the fish present points in the unassigned data set are mainstem data points.

Limiting our assessment of model performance exclusively to classification accuracies gives no indication of how far in distance the model was off in its classification. For example, if an EOFP was classified as a fish present point, the nearest point classified as fish absent could be the next upstream point on the 10m DEM network (for an error of only one point or about 10 m) or it could be 20 DEM network points upstream. If an EOFP was classified as a fish absent point, the EOFP could be the furthest downstream point classified as fish absent (and therefore correctly placed) or there could be other points classified as fish absent downstream of it (and therefore incorrectly classified). Limiting the assessment of model performance solely to FAFP data correctly classified gives no indication of how far off the placement of the EOFP was on the stream. The error could have been relatively minor, e.g., less than five points on the DEM network, or it could have been very large, e.g., 100 or more points on the DEM network. Also, classification accuracies give no indication of the direction of the error (upstream or downstream). Therefore, a process that allows us to assess error as distance, upstream or downstream, on the stream network is needed.

Finally, classifying DEM network points of unknown fish status with the final logistic regression model does not result in a prediction of an end-of-fish point on the stream. On the DEM network there are strings of points representing a stream and these points can be classified as either fish absent of fish present. Within a string of points representing a stream, there can be points classified as fish absent interspersed in a string classified as primarily fish presents points, or vice-versa. Some decision is needed on the placement of the end-of-fish point on the stream.

The next section of this report addresses two of the important issues identified in this section:

- Development of a procedure to analyze the results of applying the final logistic regression model to points of unknown fish status on the 10m DEM network and using those results to place predicted end-of-fish points on the stream network, and
- Estimating and evaluating the error of this prediction process as a distance on the stream network.

PAGE LEFT BLANK INTENTIONALLY

PROCESSING FISH PRESENT PROBABILITIES TO DEFINE AN END-OF-FISH POINT ON THE STREAM NETWORK

The logistic regression model described in the preceding section estimates the probability of fish presence for every point on the stream network. After the string of points representing a stream has been classified with the logistic regression model, a procedure for defining the placement of the end-of-fish point in the string of data is needed. Because fish absent and fish present points are often interspersed with one another in a string, a procedure is needed to ensure that EOFPs are consistently placed. This procedure needs to be repeatable and have a method for assessing its accuracy so that different rules used within the procedure can be compared and evaluated. This section of the report defines the methods developed to address these issues.

INTRODUCTION

Definition of a Cut Point

The cut point is the value which is used to convert the estimated probability of fish presence from the logistic regression into a prediction of fish absence or presence for any point on the 10m DEM network. If a pointís probability is less than the cut point, the point is designated as fish absent. If a pointís probability is equal to or greater than the cut point, the point is designated as fish present. A cut point of 0.5 was used in the preceding section as an unbiased way of assessing the performance of the logistic regression models in predicting fish absence | presence. However, the choice of a cut point of 0.5 may not be optimal when developing a procedure to predict the placement of an end-of-fish point on the DEM network. For this process, there is no *a priori* reason to choose one cut point over another without additional analyses.

Once a cut point is selected, the logistic regression model can be used to predict fish absence or presence at each point on the network. On some streams, however, the model fails to give a clear prediction of an EOFP. Figure 9 shows an example of this behavior. There is no question about the predictions for points far upstream or far downstream, but there is an ambiguous zone in the center where the points move back and forth across the cut point. The problem that must be resolved is where to place the predicted EOFP within this ambiguous zone.

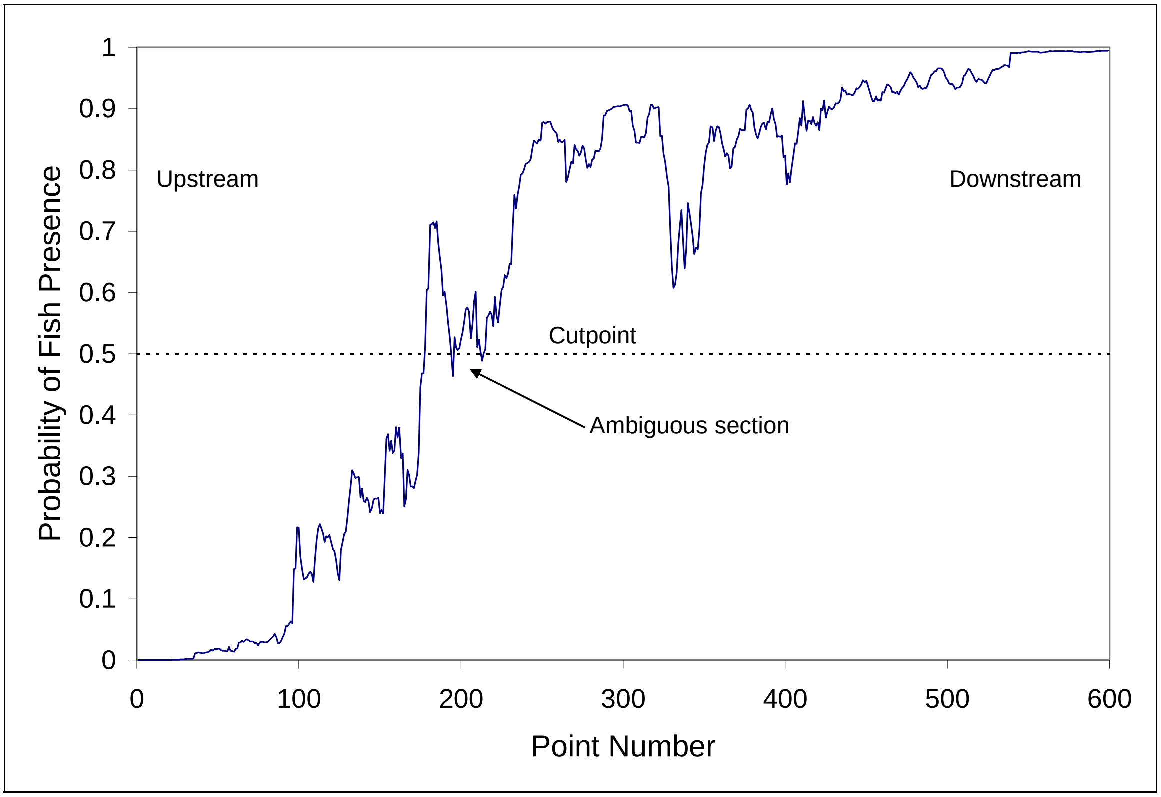


Figure 9 Logistic regression probability values for an example stream profile plotted against GIS point number. Using a cut point of 0.5, the probability values suggest that fish will be present far downstream and absent far upstream. The exact location of the predicted EOFP is unclear, however, so a stopping rule is required to resolve the ambiguity.

Following Fransen et al. (2003), we fit a stopping rule to the data. The stopping rule has three parameters which are described in more detail in the Methods section: a cut point, a trigger size, and an upstream habitat block size. The stopping rule mimics the process a person might use walking up a stream looking for the EOFP and having the logistic regression model as a guide. Starting downstream where we know there are fish, we move upstream until reaching a trigger, which is some number of consecutive fish absent points. A temporary EOFP is defined at this location. We then continue walking upstream and if there is a large enough block of fish habitat upstream, we move past the trigger and continue upstream. If there is no such upstream block, we stop and predict the EOFP just before the trigger. This section describes the process developed to define an optimal stopping rule.

METHODS

Stopping Rule Components

A stopping rule consists of three parameters: a cut point, trigger size, and upstream habitat block size.

Cut Point:

The cut point is a value between 0 and 1 which divides the LRM probabilities into predictions of fish absence and fish presence. We considered values of 0.1, 0.2, Ö , 0.9. Although the cut point could be treated as a continuous variable, we used these discrete values to simplify the optimization process.

Trigger Size:

The trigger is a block of consecutive predicted fish absent points which indicates a possible EOFP. We considered trigger sizes of 1, 3, 5, Ö , 13.

Upstream Habitat Block Size:

A large block of fish habitat upstream from a trigger suggests that the trigger should be ignored. Upstream habitat block size could be measured in several ways. We followed the method used by Fransen et al. (2003) and measured it as the sum of LRM probability values for contiguous points above the cut point. This method counts smaller blocks of high quality habitat the same as larger blocks of more marginal habitat. For example, suppose the cut point is 0.4. A block of 50 consecutive points, each having an estimated probability of fish presence of 0.7 would have a block size of 35, as would a block of 70 points, each having an estimated probability of 0.5. In our optimization, we considered values for the upstream block size of 10, 15, 20, Ö , 500.

Modeling Data and Strings

Strings:

There is an important difference in the form of the data for fitting the stopping rule. The LRM was fit with a set of scattered individual points from the stream network. The LRM uses the physical characteristics at a point and whether fish are absent|present, but it analyzes each point separately. By contrast, the stopping rule works on linear sections of the stream network which we call " strings" . A string is a linear trace through the stream network starting downstream at the bottom of the network, passing through a survey point, and ending at the top of the network. The string follows the branch with the larger basin area when the network branches. Figure 10 shows a section of stream network and a string within it.

There is one unique string for each EOFP, so in this section the terms "string" and "EOFP" are used interchangeably.

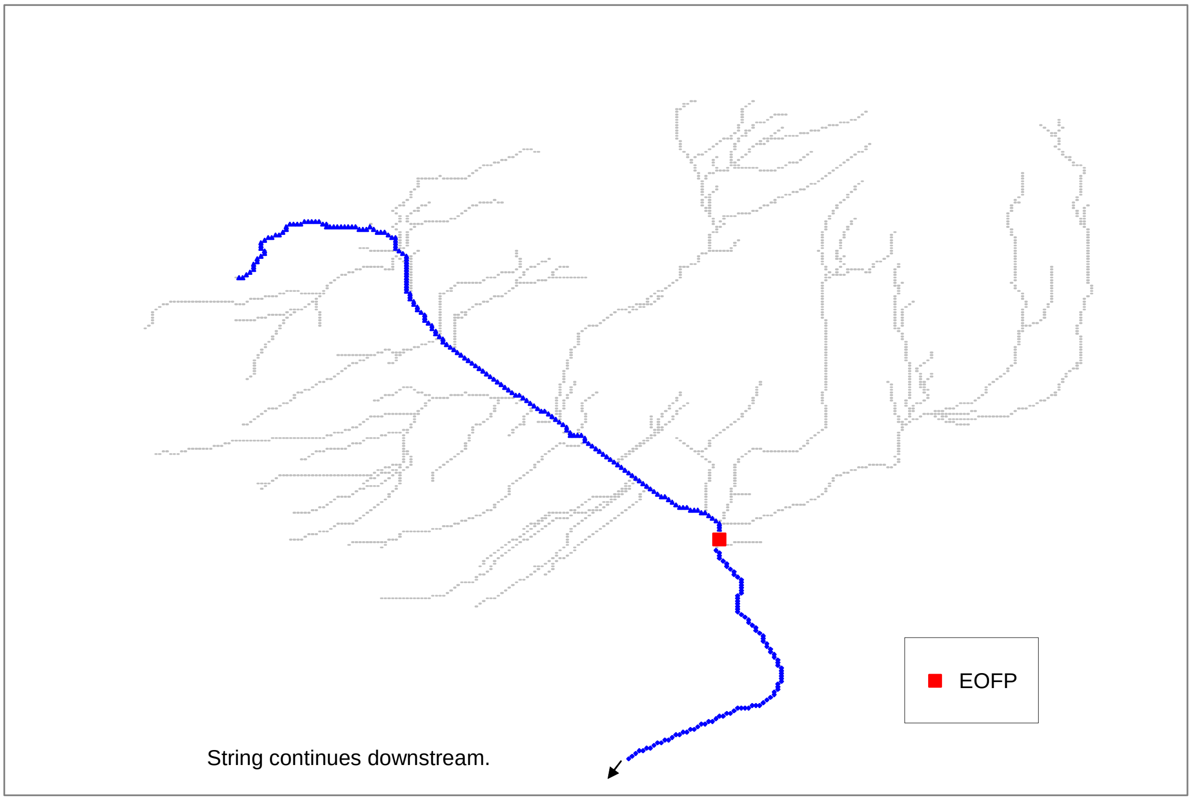


Figure 10 Illustration of a string which is a linear trace through the stream network that passes through an EOFP. There is only one EOFP and, consequently, one string in the section of stream network shown. Where EOFPs are more dense strings can overlap.

Construction of the Data Set for Stopping Rule Development:

As described in the LRM section, the available data were divided into modeling data and validation data. For the stopping rule fitting process, we kept the same division but the data sets were larger. In our initial investigations of stopping rules, we discovered that the stopping rule process required around 2,000 or more strings in the modeling data to produce a stable result. Since we had a total of 4,052 EOFPs, we wanted about half the data in each set (rule development data set and the rule validation data set) and we required the validation data to be independent of both the LRM and stopping rule fitting processes.

The LRM modeling data started with a pool of points based on EOFPs from fourth order sub-basins and having end types 1 or 2. To these 1,408 EOFPs (Table 11), we added 600 randomly selected EOFPs from the unassigned data, resulting in the final total of 2,008 modeling strings. The validation data consisted of EOFPs from FOSBs having at least one EOFP with an end type other than 1 or 2 and the remaining EOFPs from the unassigned data. There were 2,044 EOFPs in the validation data.

Error and Balance

Measuring Error:

For the logistic regression models, the measures of error were the percentages of FA and FP points correctly classified. With the stopping rules and strings, however, it was possible to define error as the distance on the DEM network between the modelís predicted EOFP and the "true[18]" EOFP. Distance is a better metric for assessing model error as it can be directly compared to the results from field validation surveys and recognizes that not all errors are equal. With classification accuracy a point is correctly classified or it is not: when using distance to define error it is clear that an error of +10 ft is not as serious as an error of +1,000 ft. We considered several possible measures of error, including mean error distance, mean absolute error distance, median error distance, median absolute error distance, and 10% trimmed mean[19] error distance.

The distribution of error distances had characteristics that helped define the choice of a measure. See Figure 12 for a typical example. For most candidate stopping rules there were many strings with zero error and there tended to be large outliers in both tails. The mean error distance and the mean absolute error distance are both heavily influenced by these outliers. The median and 10% trimmed mean are both resistant to these outliers. The numerous zero errors, however, made the median error distance uninformative as an objective function because the median error distance was 0 for most rules. The 10% trimmed mean had the same problem to a lesser degree and seemed better suited as a measure of balance. We focused on the median absolute error, which is simple and resistant to both problems for the distribution of errors.

Error Terminology:

There are two kinds of errors that result from predicting EOFPs. When the prediction was upstream from the true EOFP, we called the error an "over-prediction." Our convention was to label this error negative.

---

[18] "True" meaning the location of the end-of-fish point was identified by a field survey.

[19] A trimmed mean removes a specified percentage (commonly 10%) of the observations from each end of the distribution, and then averages the remaining observations. It is commonly used where outliers are a problem (Hoaglin et al. 1983).

When the prediction was downstream from the true EOFP, we called the error an " under-prediction. " Our convention was to label this error positive.

Balance:

An ideal stopping rule should be balanced. That is, its errors should be equal in the upstream and downstream directions. One way to view this is that there would be equal numbers of under-predictions and over-predictions. This approach weights errors equally regardless of their length. Instead, we defined balance based on the 10% trimmed mean error distance. This measure ignores the large outlier errors, but weights the remaining errors according to their severity (length of stream in error). Although this measure does not work well as an objective function, it does a good job of measuring balance. Values close to 0 indicate that, excluding outliers, there are approximately equal lengths of error distance in both directions (upstream and downstream).

Benchmark Rules

Following Fransen et al. (2003), we compared the rules we developed to two simple benchmarks. Both benchmarks use a cut point of 0.5, which is the default cut point used in many logistic regression analysis situations.

Benchmark 1: Stop at the first predicted fish absent point using a cut point of 0.5.

The first benchmark is the stopping rule one would use to stop at the beginning of the ambiguous stream section. The parameter values used for Benchmark 1 are cut point=0.5, trigger size=1, upstream habitat block size=100,000. The value for upstream habitat block can be arbitrarily large, since the intent is never to move past the first trigger. For this rule, the EOFP is placed where the first fish absent point (a point with a probability of fish presence < 0.5) is encountered as you move upstream.

Benchmark 2: Stop after the last predicted fish using a cut point of 0.5.

The second benchmark is the stopping rule one would use to continue upstream until there is no more predicted fish habitat (any point with a probability of fish presence 7 0.5). The parameter values used for Benchmark 2 are cut point=0.5, trigger size=1, upstream habitat block size=0. With these parameters, even a single fish present point in a long-string of fish absent points will cause the EOFP to be moved to the point above the last fish present point.

Optimization Method

We fit the stopping rule to minimize error while achieving a degree of balance on the strings. The method we used for finding an optimum stopping rule was a simple grid search with a constraint for balance. In other words, we examined every combination of stopping rule parameters for the ranges listed earlier. Each stopping rule was applied to the modeling strings, and the median absolute error was computed. For the set of rules that met our balance constraint (-50 ft > 10% trimmed mean error > 50 ft), the rule with the minimum median absolute error was chosen. Because we used discrete values of the stopping rule parameters, it was possible to find the minimum value by simply sorting the results.

RESULTS

The final stopping rule based on minimizing the median absolute error distance subject to the constraint that -50 ft > 10% trimmed mean error > 50 ft was:

Cut point	0.4
Trigger size	1
Upstream habitat block size	85

As Figure 11 shows, the response surface is fairly flat in the " Upstream habitat block size" dimension. There are many stopping rules having the same cut point and trigger size across a range of values for upstream habitat block size that give the same median absolute error. In cases like this our convention was to choose the stopping rule with the smallest upstream habitat block size in the region of the response surface where the results stabilize. In this case, there were also two trigger sizes that gave the same minimum, and again we chose the smaller of the two. Favoring smaller parameter values where ties exist should improve the ruleís sensitivity, although there is no direct evidence of this.

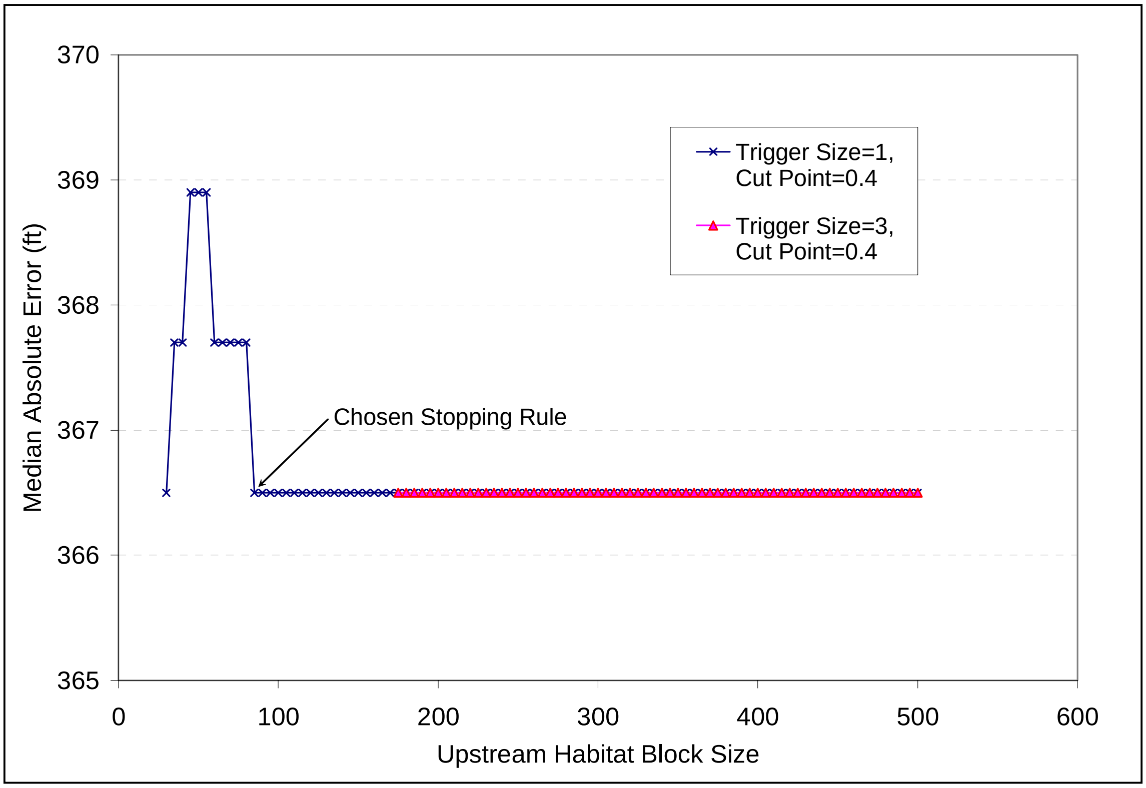


Figure 11 Candidate stopping rules that satisfy the balance constraint.

The minimum median absolute error and 10% trimmed mean attained by the final stopping rule are shown in Table 18. The same results for the unconstrained optimal rule (cut point=0.4, trigger size=11, upstream habitat block size=30) are also shown. There is a small penalty (6 ft) in median absolute error for constraining the rule to achieve better balance on the strings.

Table 18 Final results for modeling data with the optimal constrained rule and for the optimal rule without a balance constraint.

Error Measurement	Optimal Rule with Constrained Balance	Optimal Rule with No Constraints on Balance
Median absolute error (ft)	367	361
10% trimmed mean error (ft)	-40	-95

Stopping Rules Applied to Withheld Validation Strings

Figure 12 is a histogram of the errors that resulted from applying the final rule stopping rule to the 2,044 withheld validation strings. The structure of the distribution is typical of every error distribution of this type we examined:

1. The distribution is centered at zero and there is a very large proportion of zero errors (32%).
2. The distribution is not symmetrical but is skewed to the left. This means there is a tendency to make over-predictions (place the predicted EOFP upstream of the true EOFP).
3. There are a number of outliers, some of which are extremely large. Like the skewing, these tend to be over-predictions.

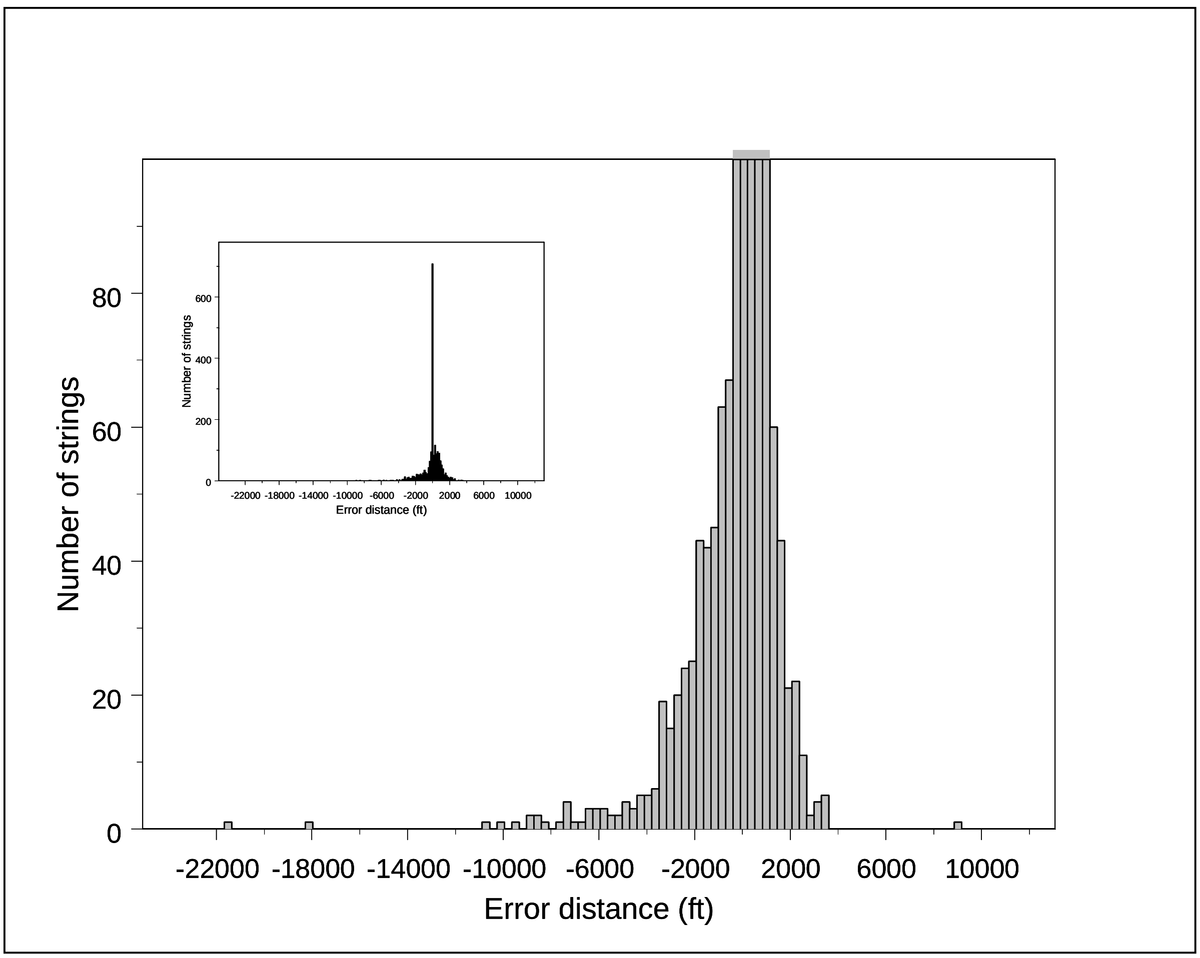


Figure 12 Histogram of error distances when the final stopping rule is applied to the validation strings. The large histogram is truncated to show outliers and other fine structure that is hidden when the large bar at zero is fully displayed. The inset shows the shape of the entire histogram.

Table 19 compares four stopping rules using the independent validation data. In addition to the final rule, there are the unconstrained optimal rule and the two benchmark rules. The two optimal rules are better than either benchmark, both in median absolute error and in balance. Compared to each other, however, there is little if any apparent difference between the constrained and unconstrained optimal rules.

Table 19 Final results on independent validation data for the optimal rule with constrained balance, the unconstrained optimal rule, and two benchmark rules.

Assessment Statistic	Optimal Rule with Constrained Balance	Unconstrained Optimal Rule	Benchmark 1: Stop at First Predicted Non-Fish	Benchmark 2: Stop After Last Predicted Fish
Median absolute error (ft)	327	327	400	395
10% trimmed mean error (ft)	37	2	196	132
Mean error (ft) [a]	-140	-196	66	-33
% Under [b]	39%	37%	45%	43%
% Exact [c]	32%	33%	33%	33%
% Over [d]	29%	30%	23%	24%
% Within 100 ft [e]	36%	36%	36%	37%

[a] The mean error includes all outliers.

[b] Percentage of strings with under-predictions (i.e., positive errors).

[c] Percentage of strings with zero errors.

[d] Percentage of strings with over-predictions (i.e., negative errors).

[e] Percentage of strings with errors between -100 ft and 100 ft.

DISCUSSION

The string form of the stream network data was convenient for implementing the optimization process but it has one major drawback; the assessment of model performance based on strings may not correctly characterize performance when the model is applied to the 10m DEM network. Specifically, the estimates of model precision (total error) and balance based on strings may not represent what occurs on the stream network. This problem is described in more detail in part 2 of the next section. The following section describes the methods we developed for estimating precision and balance for entire sections of the 10m DEM stream network.

MODEL VALIDATION AND ASSESSMENT OF ERROR

This section has three main parts. Part 1 discusses the model validation approach and presents validation results based on strings. Part 2 focuses on estimates of model precision and balance based on an assessment of an entire basin or sub-basin. Part 3 concludes with some observations on the different sources of model error.

PART 1: VALIDATION RESULTS BASED ON STRINGS

INTRODUCTION

There are several possible methods for assessing model performance. Cross-validation has the advantages of being easily described and applied, readily interpretable, and commonly used (Efron and Tibshirani 1993). The basic idea behind cross-validation methods is to fit the model with one portion of the data (the training set) and then assess the modelís performance by applying it to the remaining data (the validation set). The "cross" in cross-validation means that this procedure is repeated many times with each set acting as training and validation data (for different iterations). The most common form of cross-validation, k-fold cross-validation, consists of breaking the data into *k* sets, fitting the model using *k*-1 of the sets, and validating the model using the remaining set. This procedure is repeated *k* times so that each set acts as a validation set once. The measures of fit from the *k* iterations are then summarized to describe the prediction error. Leave-one-out cross-validation is the special case where each validation set is just a single observation.

A closely related procedure, called hold out or split sample validation, only performs the fit/validation step once. The data is fit to a training set and the prediction error is calculated using a single validation set. This procedure is much less computationally intensive but can perform poorly when used with small data sets (Goutte 1997). Because the data set we were analyzing was very large, we chose this approach (split sample validation) to avoid the prohibitive amount of computer time that would be necessary for a complete k-fold cross-validation.

The details of how the training and validation sets were chosen, along with the motivation for this approach, were presented in the earlier section "Selected Approach to Logistic Regression Model Building" (page 29).

Types of Errors

The discrepancy between the predicted fish presence map and the observed (field survey) data can be characterized by the error in the classification of points on the stream network, or by the distance between the predicted and observed end-of-fish points.

Point-based Metrics:

The approach described in the Forest and Fish Agreement is based on the number or percentage of points that are misclassified. Table 20 illustrates how each point in the stream network can be correctly or incorrectly classified. Comparing the percentage of incorrectly classified fish present points ***N***(FP,FA)/***T***(FP) to the percentage of incorrectly classified fish absent points ***N***(FA,FP)/***T***(FA) is problematic since ***T***(FP) and ***T***(FA) are determined by decisions with no obvious correct choice. For example, the GIS program that generates the stream network depends on a parameter that determines the number of square meters that must drain into a point before it becomes part of the stream network. Decreasing this value will increase the number of fish absent points while the number of fish present points stays the same. A more robust comparison is between the number of misclassified fish present points ***N***(FP,FA) and the number of misclassified fish absent points ***N***(FA,FP).

Table 20 Summary of the possible outcomes from classifying points of known fish presence status with a logistic regression.

Observed (True)	Predicted fish present (FP) points	Predicted fish absent (FA) points	Total
Fish present points	**N** (FP,FP)	**N** (FP,FA)	**T**(FP)
Fish absent points	**N** (FA,FP)	**N** (FA,FA)	**T**(FA)

Attempts to characterize the overall misclassification error are again made non-trivial by the problem above. The total number of points in the stream network can be arbitrarily adjusted by changing a parameter with no obvious " correct " value. So for example, if you wanted to make the percentage of misclassified points small you could just decrease the drainage area required to start a stream because DEM network points with very small basin drainage areas are rarely misclassified.

Error Distances:

An alternative measure of prediction error is the distance between the predicted EOFP and the observed EOFP. This is only possible when stopping rules have been applied to the results of the logistic regression so that there is a clear demarcation between fish present and fish absent habitat. Error distances can be summarized in a number of ways. Subtracting the stream meter location of the predicted EOFP from the observed EOFP will yield a negative number if the predicted point is above the observed point and a positive value if the predicted point is below the observed. By calculating the mean or median of these values we can compare the lengths of stream that were misclassified in the two possible ways. This should yield results similar to subtracting ***N***(FP,FA) from ***N***(FA,FP). By first taking the absolute value of the error distance, we can examine how far the predicted EOFP is from the observed EOFP on average (or 50% of the time in the case of the median). Because the error distance is measured at the EOFP scale instead of the DEM network point scale (like the metrics above), it is more appropriate for discussing error as it relates to an individual stream.

METHODS

We used two types of error distance in our analyses:

1. **Absolute error distance**: The distance between the observed and predicted end-of-fish points.
2. **Error distance**: The same as the absolute error distance except the error distance is negative when the observed EOFP is downstream of the predicted point (i.e., the error distance is negative when the length of fish bearing stream is over-estimated).

The different measures of error described above were summarized for the entire validation data set.

Error Summary by Important Covariates

For every end-of-fish point there was a list of associated variables (covariates) derived from the field survey records or from the GIS network. These included: WRIA, survey sponsor, survey date, survey protocol, boundary type, end type, flow accumulation, elevation, precipitation, upstream and downstream gradient, and stream order. There are other covariates that were not available for this analysis that would have been useful such as stream discharge at the time of data collection relative to average and maximum discharge, relative stream condition (e.g., wood supply, any upstream sediment events, etc.), geology, and water temperature. In addition, some covariates could be highly correlated but further investigations are needed to understand their importance. For example, different survey seasons and survey protocols may be more common in some WRIAs compared to others.

For the different discrete covariates, the error distances were summarized by the major categories of the covariate. A table with the mean, standard deviation, and coefficient of variation of the error distances, by category, was produced for each covariate examined. In addition, a modified box-and-whiskers plot was produced so that the error distributions could be visually compared. These modified box-and-whiskers plots were over-laid with a strip chart showing the values for the individual observations (which were randomly jittered [Chambers et al. 1983] to reduce overlap of points with identical values).

For the continuous covariates, a scatter plot was produced showing the value for the continuous covariate on the x-axis and the error distance for each point on the y-axis. Because of the density of points, Friedmanís (1984) super-smoother was used to produce a line indicating the trend of the data.

RESULTS

Overall, the distribution of all error distances was skewed. Very large negative errors were more common than very large positive errors (Figure 13, Table 21). This is expected since there are many fish blockages on the network that cannot be predicted based on the physical variables available (e.g., blockages due to culverts, beaver dams, and waterfalls).

Table 21 Summary statistics for the two measurements of error distance (in ft) by quantiles.

Error Measurement	Mean	Stand. Dev.	0%	1%	25%	50%	75%	99%	100%
Number of points			0	20	510	1,021	1,532	2,022	2,042
Absolute Error Distance	762	1,306	0	0	0	327	994	2,817	21,666
Error Distance	-140	1,505	-21,666	-6,214	-183	0	468	2,423	9,163

<u>Error Summary by Important Covariates</u>

End Type:

The probable cause for the placement of the end-of-fish point was assigned to each EOFP in the field. The codes assigned were:

1: Natural end (stream size related or boundary types B and C)
2: Gradient related (e.g., water falls)
3: Large woody debris
4: Road culvert
5: Mass wasting event (landslide)
6: Beaver dam or other non-permanent dam
7: Other dam (permanent)
8: Water quality limiter
9: None
10: Unknown

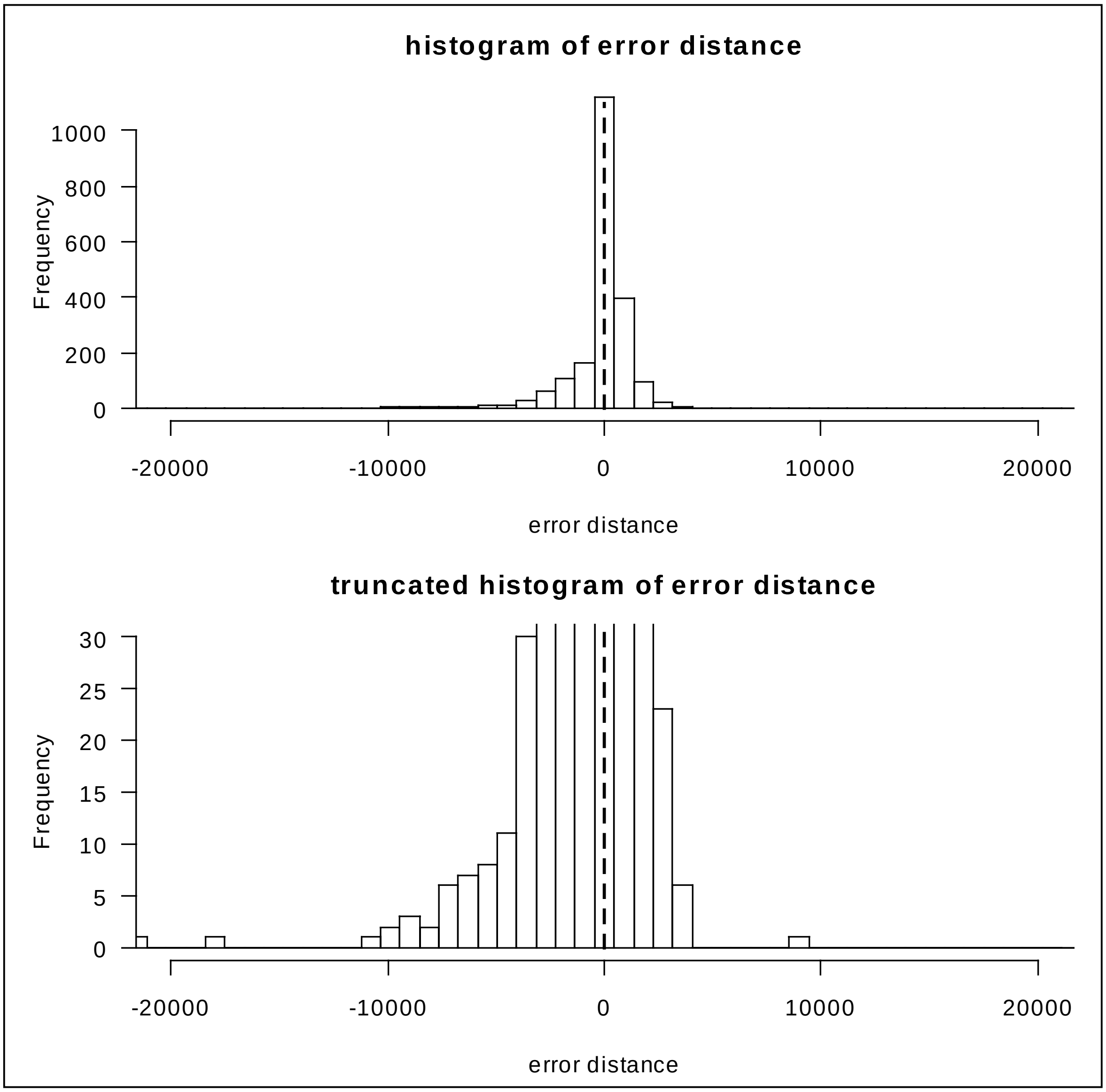


Figure 13 Histograms for all of the error distances (in ft) resulting from the application of the final logistic regression model and stopping rules to the validation data set.

These categories can be broken into end types we would expect to be able to predict with the physical information we have (end types 1 and 2) and end types that would be difficult to predict (3, 4, 5, and 6). The error distances in the predictable group were smaller than those in the less predictable group (Figure 14, Table 22). In addition, the values were more likely to be negative in the less predictable group. This makes sense since the model should tend to over-predict fish habitat when the observed last fish point is caused by a blockage (e.g., a culvert, mass wasting, or beaver dam). Some of the CVs were extremely large because of the outliers present in some end types.

Table 22 Summary statistics for the two measurements of error distance (in feet) by end-of-fish point end type (only end types with a sample size 7 20 points are shown).

End Type	SS[a]	Absolute Error Distance			Error Distance		
		Mean	St. Dev.	CV[b]	Mean	St. Dev.	CV
1	933	448	884	197%	12	991	8,258%
2	498	702	940	134%	89	1,170	1,315%
3	129	1,181	1,910	162%	-416	2,209	531%
4	126	1,185	1,312	111%	-416	1,721	414%
5	37	851	1,065	125%	-494	1,276	258%
6	31	1,439	1,594	111%	-1,360	1,665	122%
9	29	1,276	1,522	119%	-841	1,808	215%
10	254	1,447	2,143	148%	-574	2,522	439%

[a] SS = sample size. [b] CV = coefficient of variation.

Sponsor:

Most of the observed EOFPs in the validation set were collected by two sponsors (Quinault Natural Resources and Hoh Fisheries). The distribution of errors for most sponsors with a large number of points tended to be centered near 0 with relatively low variability (Figure 15, Table 23). Two notable exceptions were Washington Trout and Plum Creek. In the case of Washington Trout, the surveyed points tended to be lower in the stream network than the predictions while the opposite occurred for the Plum Creek data.

Protocol:

The Last Fish Habitat (LFH) protocol was used for most of the survey points in the validation set. As expected, the protocols based on habitat (LFH and Last Salmonid Habitat) tended to have more positive error distances than the fish-based surveys (Last Salmonid and Last Fish), since they tend to be placed higher in the watershed (Figure 16, Table 24).

Table 23 Summary statistics for the two measurements of error distance (in feet) by survey sponsor for the end-of-fish point (only sponsors with a sample size 7 20 points are shown).

Survey Sponsor	SS[a]	Absolute Error Distance			Error Distance		
		Mean	St. Dev.	CV[b]	Mean	St. Dev.	CV
Aquatic Tech	43	1,226	2,172	177%	-782	2,373	303%
Champion	40	358	636	178%	283	674	238%
WA Dept. Nat. Res.	59	811	1,192	147%	-302	1,413	468%
Hoh Fisheries	249	732	1,158	158%	-183	1,359	743%
Olympic Resource	20	540	727	135%	-86	910	1,058%
Pt. Blakely Tree Farm	63	709	1,150	162%	-35	1,353	3,866%
Plum Creek	23	1,161	1,145	99%	1,057	1,245	118%
Quinault DNR	1,333	631	929	147%	-37	1,122	3,032%
Tulalip DNR.	30	1,401	1,543	110%	153	2,095	1,369%
WA Trout	67	1,999	3,487	174%	-1,493	3,736	250%
Weyerhaeuser Co.	92	1,147	1,925	168%	-560	2172	388%

[a] SS = sample size. [b] CV = coefficient of variation.

Table 24 Summary statistics for the two measurements of error distance (in feet) by survey protocol for the end-of-fish point.

Survey Protocol	SS[a]	Absolute Error Distance			Error Distance		
		Mean	St. Dev.	CV[b]	Mean	St. Dev.	CV
Last Fish	102	1,376	2,191	159%	-834	2,451	294%
Last Fish Habitat	1,793	677	1,199	177%	-90	1,374	1,527%
Last Salmonid	85	1,502	1,538	102%	-795	2,002	252%
Last Salmonid Habitat	62	1,199	1,341	112%	446	1,748	392%

[a] SS = sample size. [b] CV = coefficient of variation.

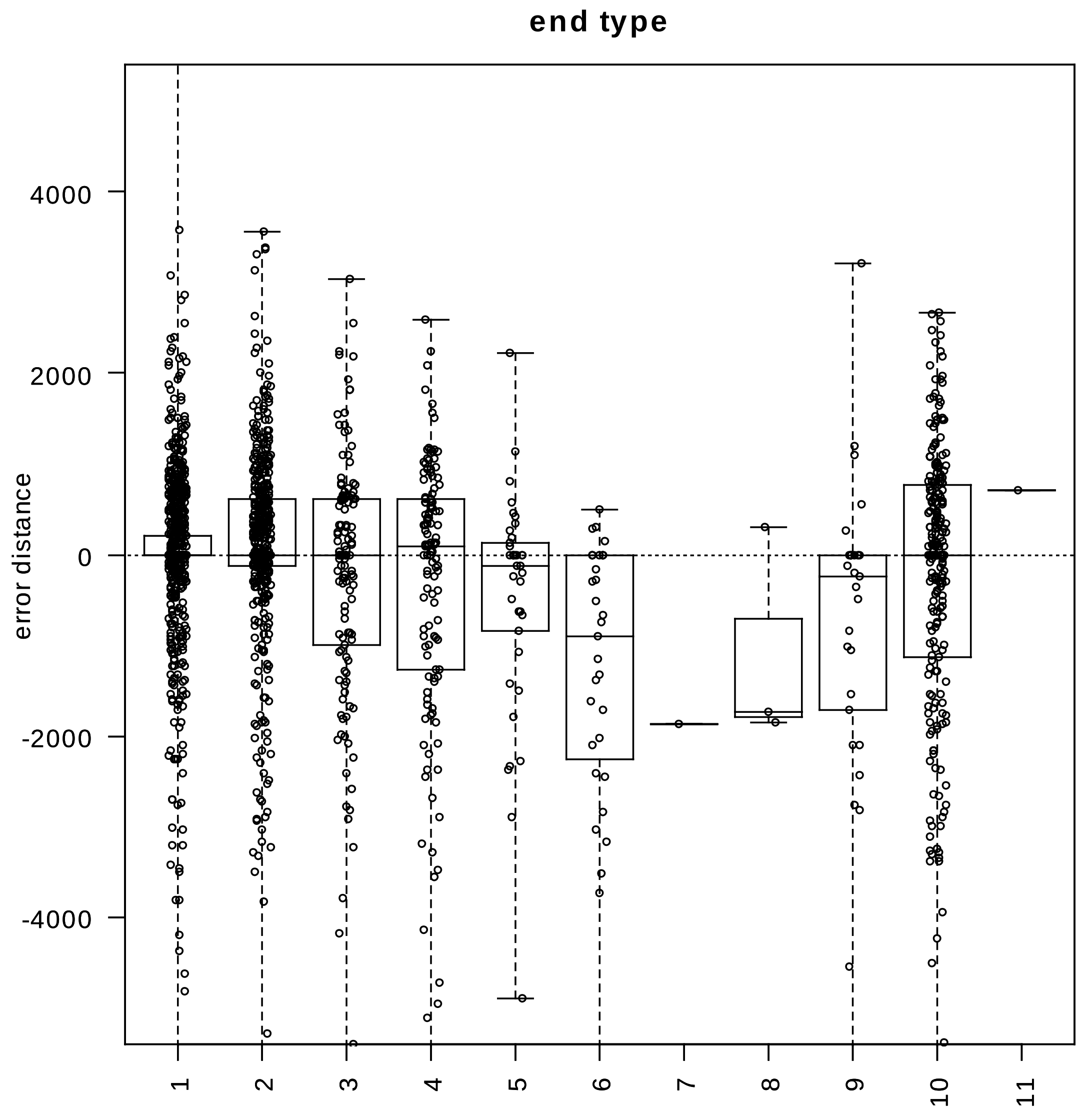


Figure 14 Box-and-whiskers plots summarizing distributions of error distances (in ft) by end type of the surveyed end-of-fish point.

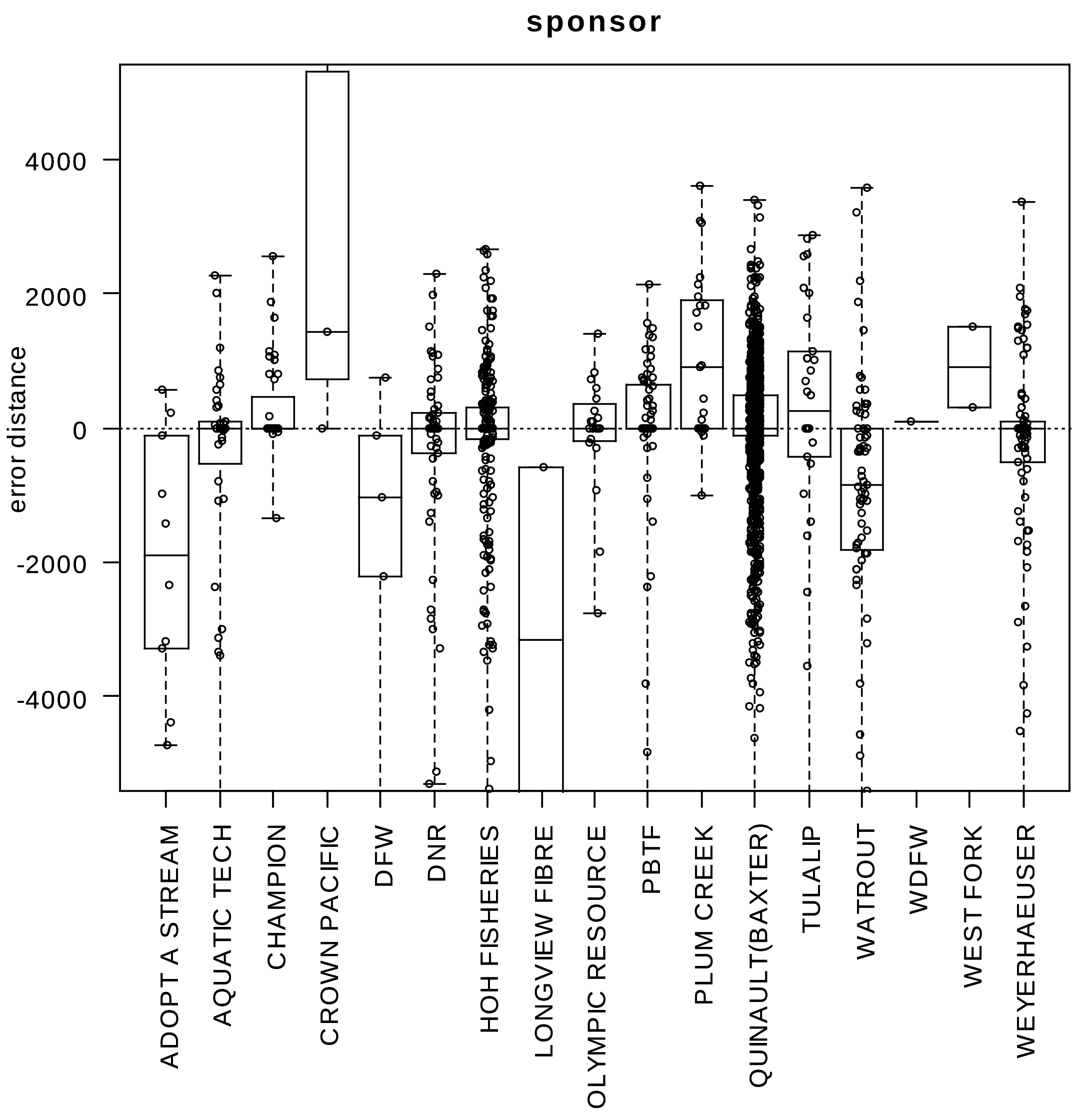


Figure 15 Box-and-whiskers plots summarizing distributions of error distances (in ft) by sponsor of the end-of-fish point survey.

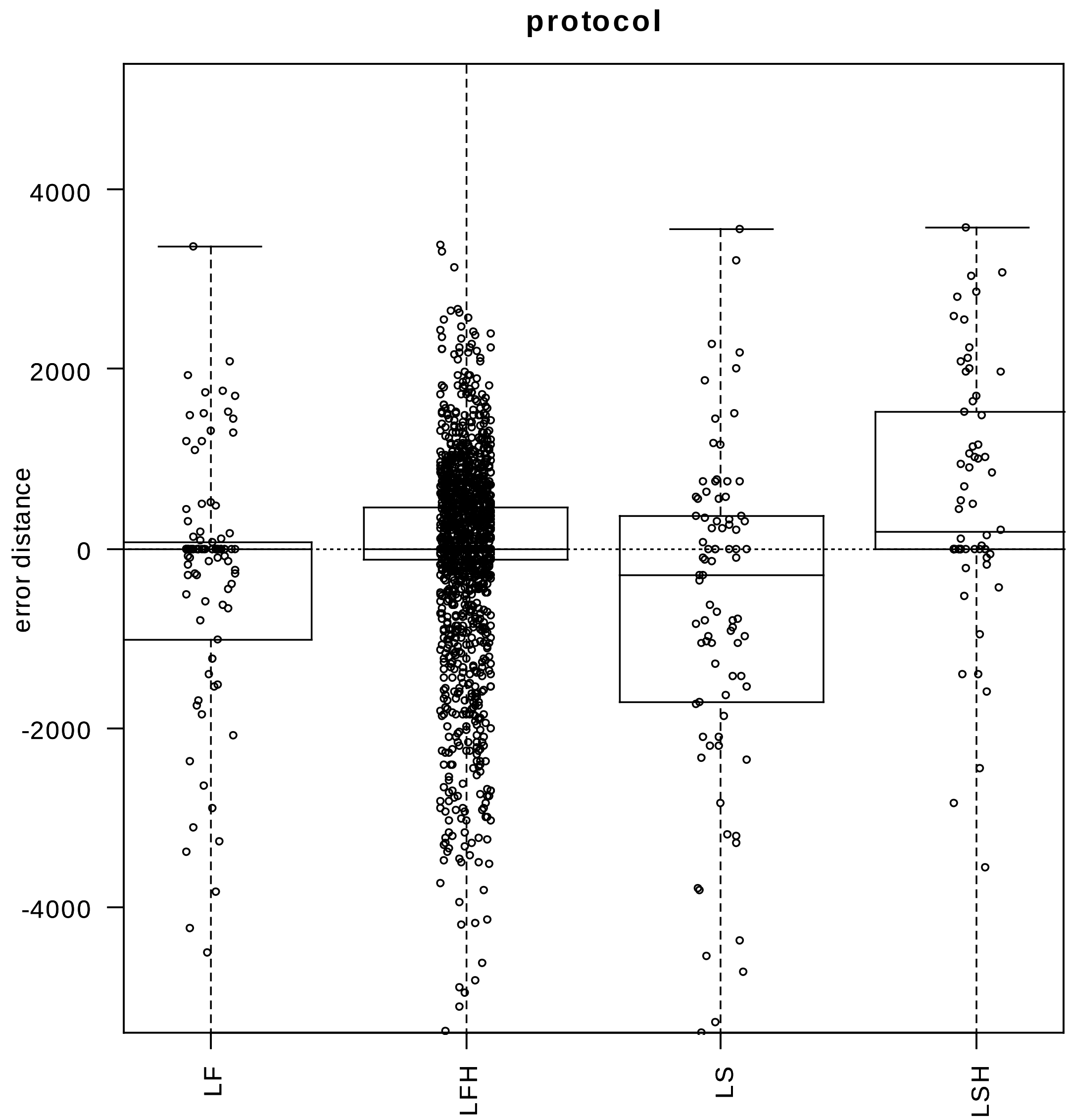


Figure 16 Box-and-whiskers plots summarizing distributions of error distances (in ft) by survey protocol for the end-of-fish point.

Boundary Type:

Boundary type describes the relationship of the end-of-fish point to confluences with other streams (see Figure 2). Because basin size usually changes drastically when moving into a small tributary, and basin size is an important predictor of fish presence, the lateral confluence EOFPs (type B) tended to be predicted with much higher accuracy than the other two boundary types. Of the 882 type B points in the validation set, 600 (68%) were predicted exactly. In contrast, only 54 of the other 1,160 EOFPs of types A and C (less than 5%) had zero error distances (Figure 17, Table 25).

Because lateral confluence EOFPs occur at the confluence with a larger river, incorrect predictions are much more likely to be upstream than downstream (200 of the 282 non-zero error distances were negative). Type A and C boundary EOFPs, however, tend to have more positive error distances (707 out of 1,160 or 65%). This is probably due to compensation for the propensity of negative type B points. The optimization procedure used to develop the stopping rules constrained the average error distance to be close to zero.

Table 25 Summary statistics for the two measurements of error distance (in feet) by the boundary type of the end-of-fish point.

Boundary Type	SS[a]	Absolute Error Distance			Error Distance		
		Mean	St. Dev.	CV[b]	Mean	St. Dev.	CV
Mid-channel (A)	726	1,202	1,637	136%	-132	2,028	1,536%
Lateral Confluence (B)	882	337	881	261%	-222	917	413%
Tributary Junction (C)	434	888	1,118	126%	12	1,428	11,900%

[a] SS = sample size. [b] CV = coefficient of variation.

WRIA:

There was substantial variability in error distributions among the different WRIAs (Figure 18). However, most of the WRIAs with distributions above or below zero or with large variance had few EOFPs in the validation set (Table 26). Some effort was put into stratifying the training set across WRIAs so that the importance of fitting points within a WRIA was not just decided by the number of survey EOFPs in the WRIA. If this were the case, minor improvements in fit for the WRIAs with many EOFPs would come at the expense of large decreases in the overall fit for the WRIAs with fewer EOFPs. However, WRIAs with few EOFPs have poorer fits due to their paucity of points. It is therefore not unexpected that many of the WRIAs with few points did not appear to perform as well. Only the collection of more EOFPs from the under-represented WRIAs will remedy this problem.

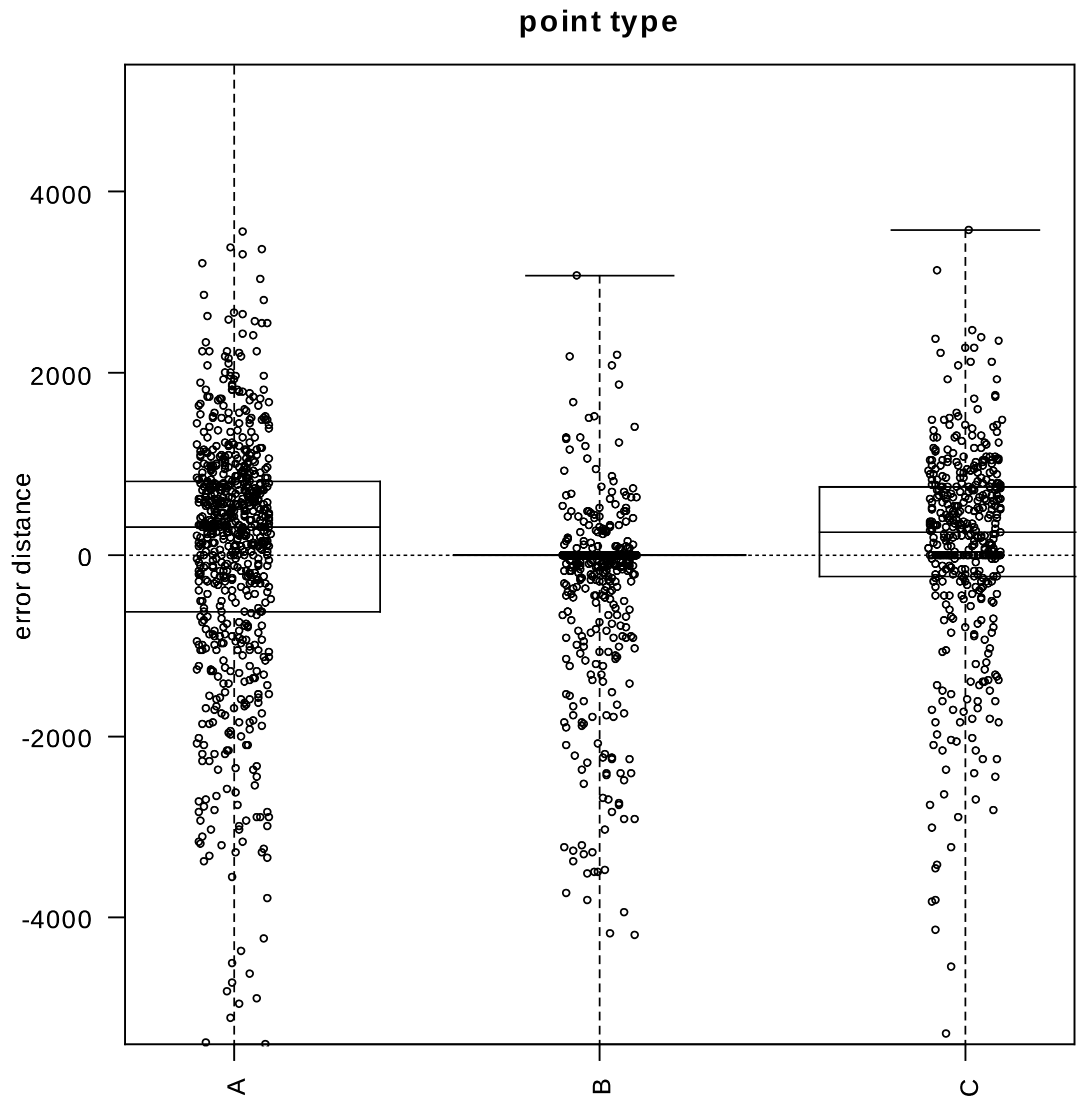


Figure 17 Box-and-whiskers plots summarizing distributions of error distances (in ft) by the boundary type of the end-of-fish point.

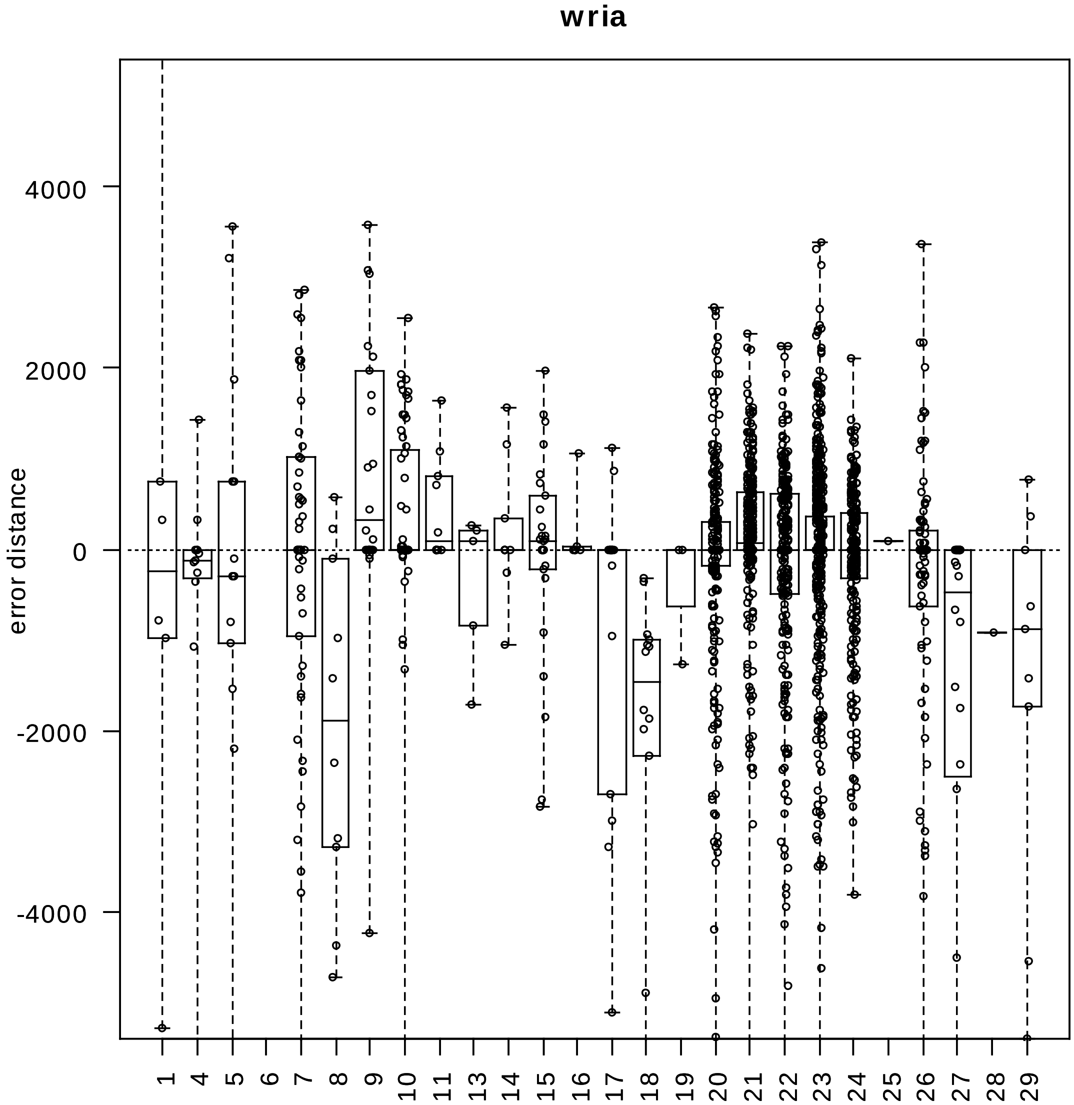


Figure 18 Box-and-whiskers plots summarizing distributions of error distances (in ft) by WRIA.

Table 26 Summary statistics for the two measurements of error distance (in feet) by WRIA (only WRIAs with a sample size > 10 points are shown).

WRIA	SS[a]	Absolute Error Distance Mean	Absolute Error Distance St. Dev.	Absolute Error Distance CV[b]	Error Distance Mean	Error Distance St. Dev.	Error Distance CV
3 and 4	11	914	1,835	201%	-596	1,973	331%
5	13	1,704	1,621	95%	-140	2,398	1,713%
7	50	1,355	1,396	103%	-120	1,951	1,626%
9	22	1,196	1,359	114%	793	1,639	207%
10	59	720	1,372	191%	258	1,531	593%
15	26	766	852	111%	-42	1,155	2,750%
17	14	1,232	1,644	133%	-949	1,835	193%
18	14	4,174	6,776	162%	-4,174	6,776	162%
20	265	707	1,128	159%	-170	1,321	777%
21	294	542	686	127%	164	859	524%
22	246	894	1,096	122%	-214	1,399	654%
23	599	575	988	172%	-18	1,143	6,350%
24	255	594	681	115%	-147	892	607%
26	93	1,197	1,915	160%	-633	2,170	343%
27	20	2,039	3,073	151%	-2,039	3,073	151%

[a] SS = sample size. [b] CV = coefficient of variation.

Table 27 Summary statistics for the two measurements of error distance (in feet) by month of the end-of-fish point survey.

Month	SS[a]	Absolute Error Distance Mean	Absolute Error Distance St. Dev.	Absolute Error Distance CV[b]	Error Distance Mean	Error Distance St. Dev.	Error Distance CV
Undefined	143	675	1,012	150%	-86	1,215	1,413%
1	163	792	949	120%	-70	1,236	1,766%
2	227	739	1,080	146%	-144	1,302	904%
3	301	720	1,164	162%	-63	1,368	2,171%
4	316	734	1,217	166%	-141	1,415	1,004%
5	204	882	1,413	160%	-297	1,640	552%
6	103	1,047	1,654	158%	10	1,960	19,600%
7	108	830	1,256	151%	-370	1,461	395%
8	101	920	2,064	224%	-549	2,193	399%
9	129	921	2,134	232%	-169	2,319	1,372%
10	109	402	620	154%	104	733	705%
11	136	590	888	151%	-8	1,068	13,350%

[a] SS = sample size. [b] CV = coefficient of variation.

Month and Year:

Seasonal changes in fish distribution might suggest that error distribution for the model would change with survey month since time is not considered in the model. There does not appear to be a pattern in the errors (Figure 19, Table 27). However, month is potentially confounded with many other variables (e.g., year, sponsor, and geographic location).

Differences in yearly weather regimes (snow pack, summer low flows, etc.) may also affect fish distribution. There is no strong evidence of this in the error distribution summaries by year but like month, year is potentially confounded with many other variables (Figure 20, Table 28).

Current survey work which includes repeated visits to the same sites across seasons and years will provide much more useful information on seasonal and annual variability in fish extent.

Table 28 Summary statistics for the two measurements of error distance (in feet) by year of the end-of-fish point survey.

		Absolute Error Distance			Error Distance		
Year	SS[a]	Mean	St. Dev.	CV[b]	Mean	St. Dev.	CV
1994	176	714	1,127	158%	-311	1,298	417%
1995	461	845	1,051	124%	-151	1,341	888%
1996	720	538	877	163%	61	1,027	1,684%
1997	375	839	1,356	162%	-229	1,579	690%
1998	95	1,005	1,461	145%	-522	1,697	325%
1999	94	757	957	126%	33	1,222	3,703%
2000	117	1,467	3,096	211%	-629	3,370	536%

[a] SS = sample size. [b] CV = coefficient of variation.

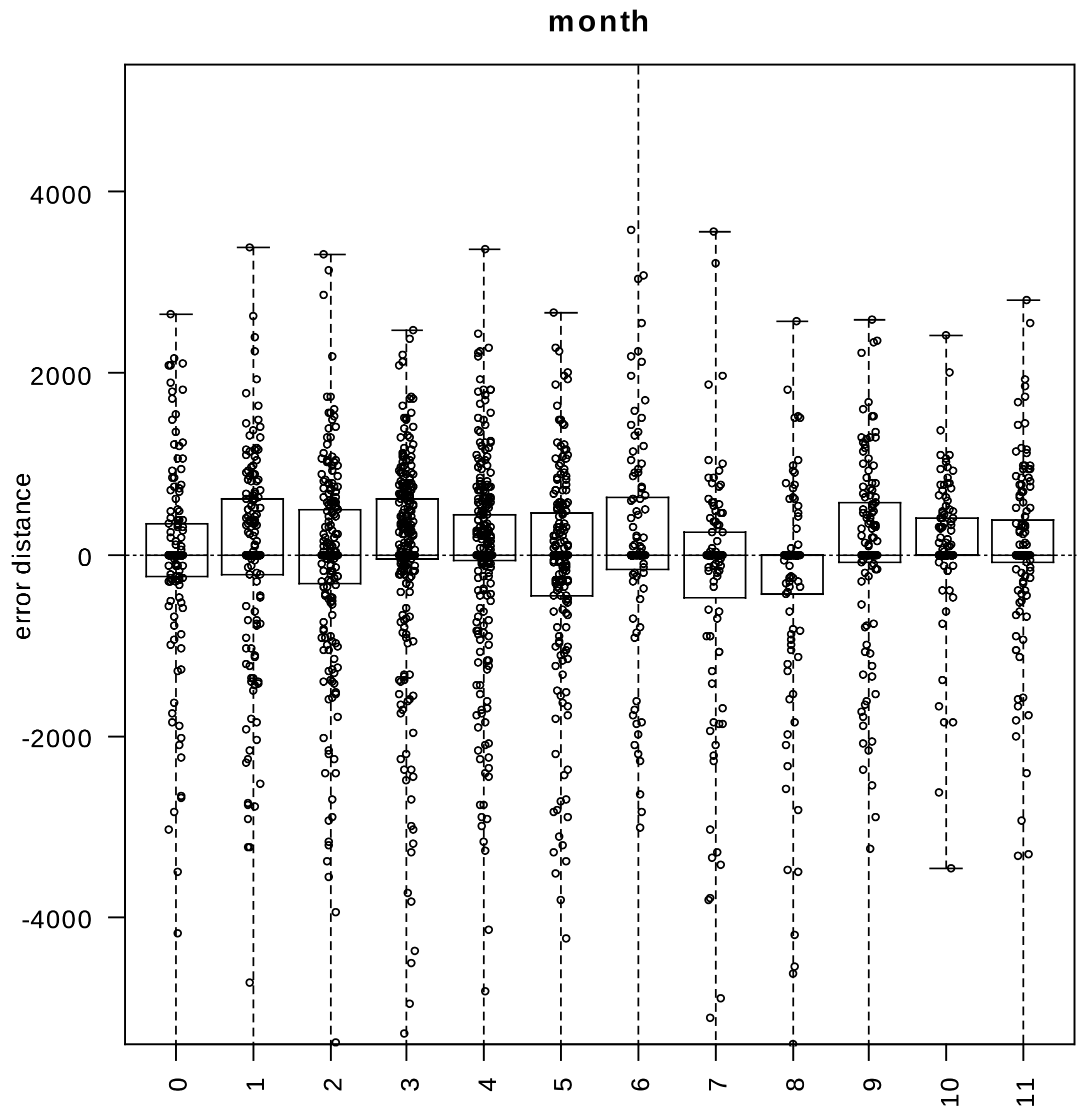


Figure 19 Box-and-whiskers plots summarizing distributions of error distances (in ft) by month of survey.

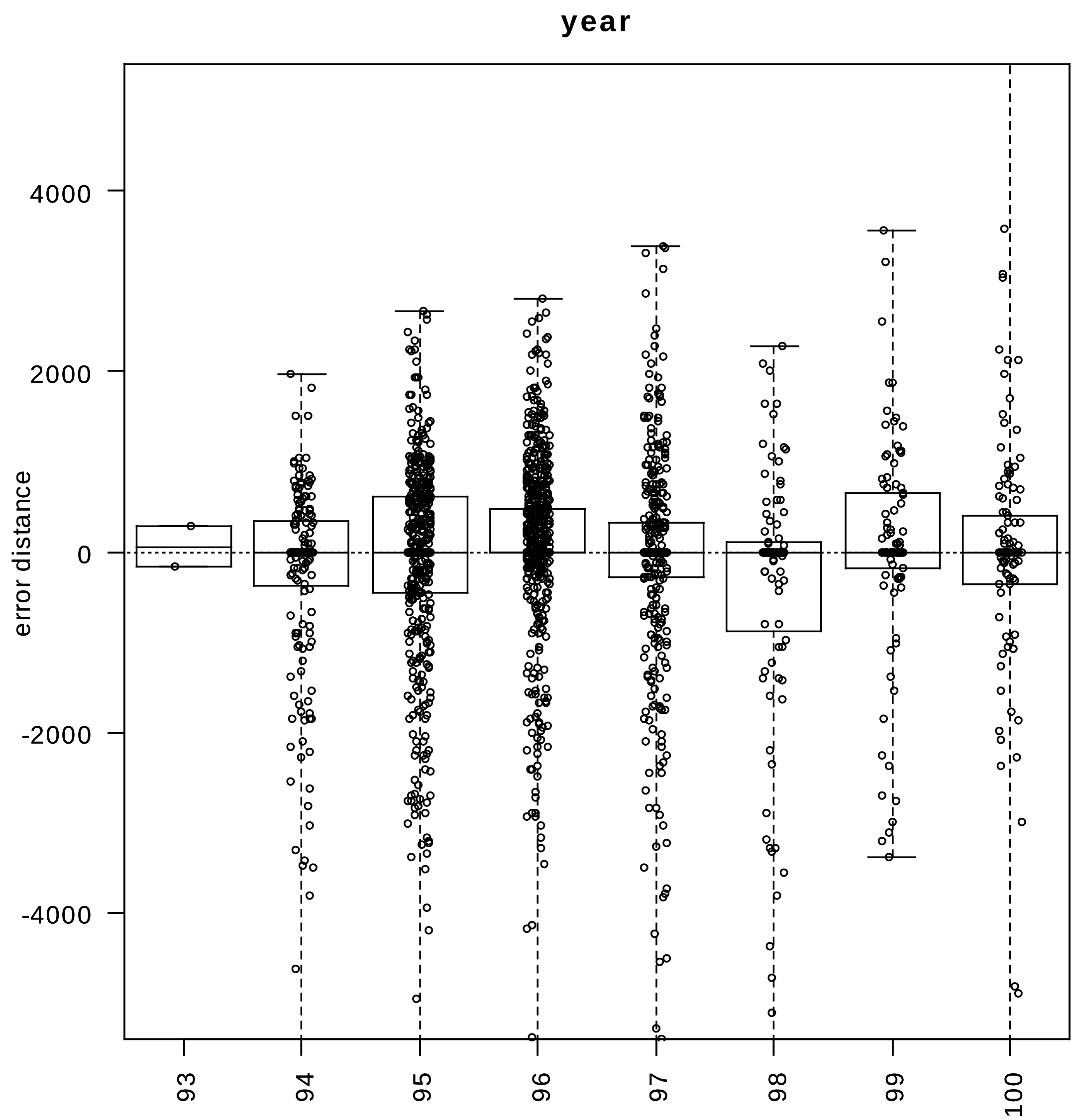


Figure 20 Box-and-whiskers plots summarizing distributions of error distances (in ft) by year of survey.

GIS Generated Variables:

All of the physical variables have the potential of being correlated with one another. Smaller basins tend to be at higher elevation and have higher gradient streams. Precipitation can also be related to elevation. This implies that we need to be conservative in our interpretations of these univariate relationships.

Elevation - There appears to be a general increasing trend in error distance with increasing elevation above 2,000 ft (Figure 21). This means that the model tends to predict downstream of the observed EOFP at higher elevations and the size of the error increases as elevation increases. This could be caused by elevation effects such as lower temperatures, or it could be caused by correlated attributes such as flow accumulation or gradient. Conversely, at elevations less than 500 ft the model tends to predict upstream of the observed EOFP. This relationship may be influenced by the frequency of very large negative outliers present in this range of the data.

Flow Accumulation - There is a strong downward trend in error distance with increasing flow accumulation (Figure 22). Because the flow accumulation (basin size) physical attribute was the most important parameter of the model, and the probability of fish presence increases rapidly as basin size increases, this is not surprising. Points with flow accumulations 7 100 acres are generally expected to be fish present points and the model typically places the predicted EOFP above them. When these points are truly fish absent points the predicted EOFP is then too far upstream which results in a negative error distance.

Downstream Gradient and Precipitation - Compared to elevation and flow accumulation, the trends exhibited by these two physical attributes are not nearly as strong (Figure 23 and Figure 24).

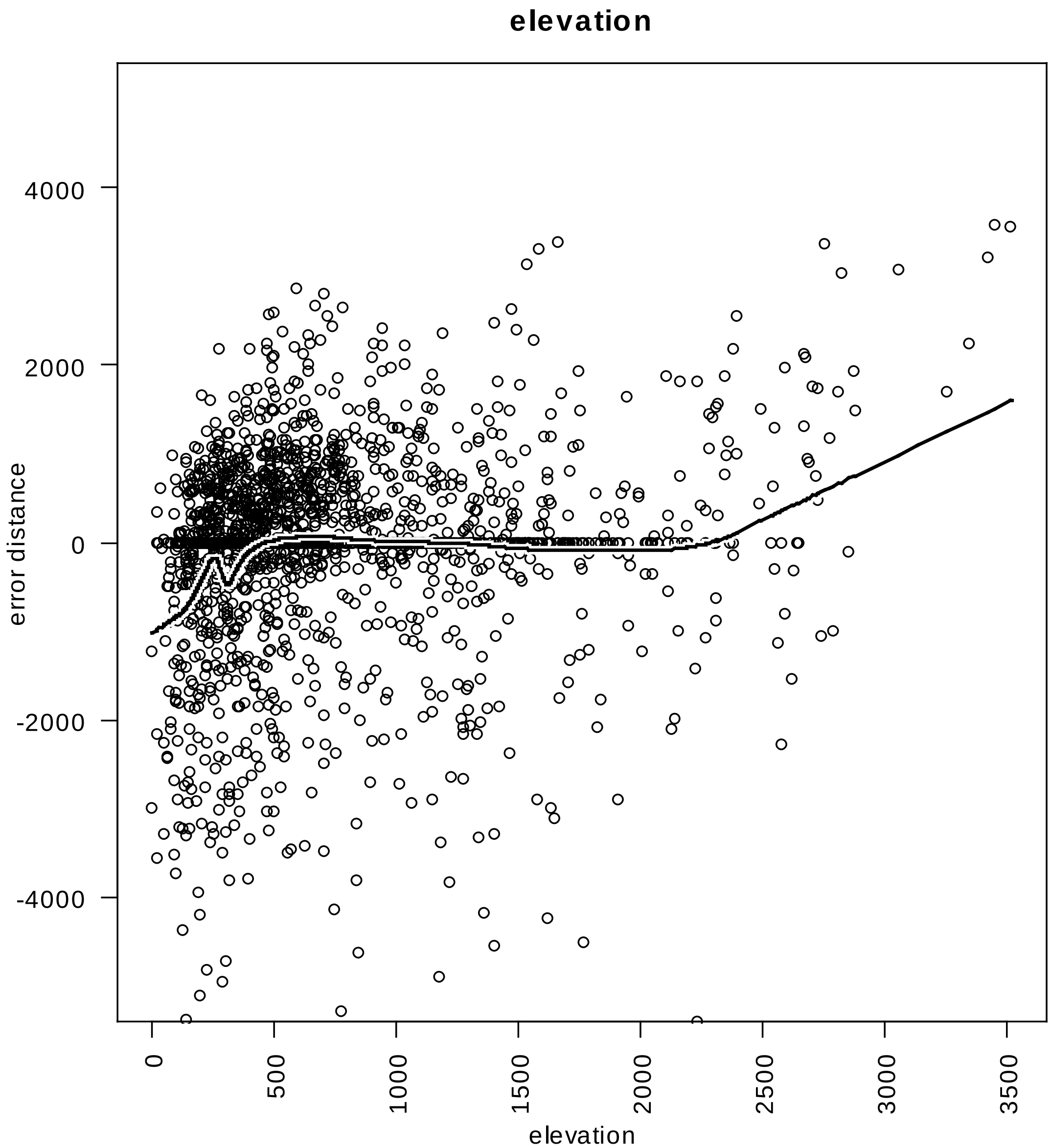


Figure 21 Scatter plot showing the relationship between error distance (in ft) and the elevation (ft) of the end-of-fish point.

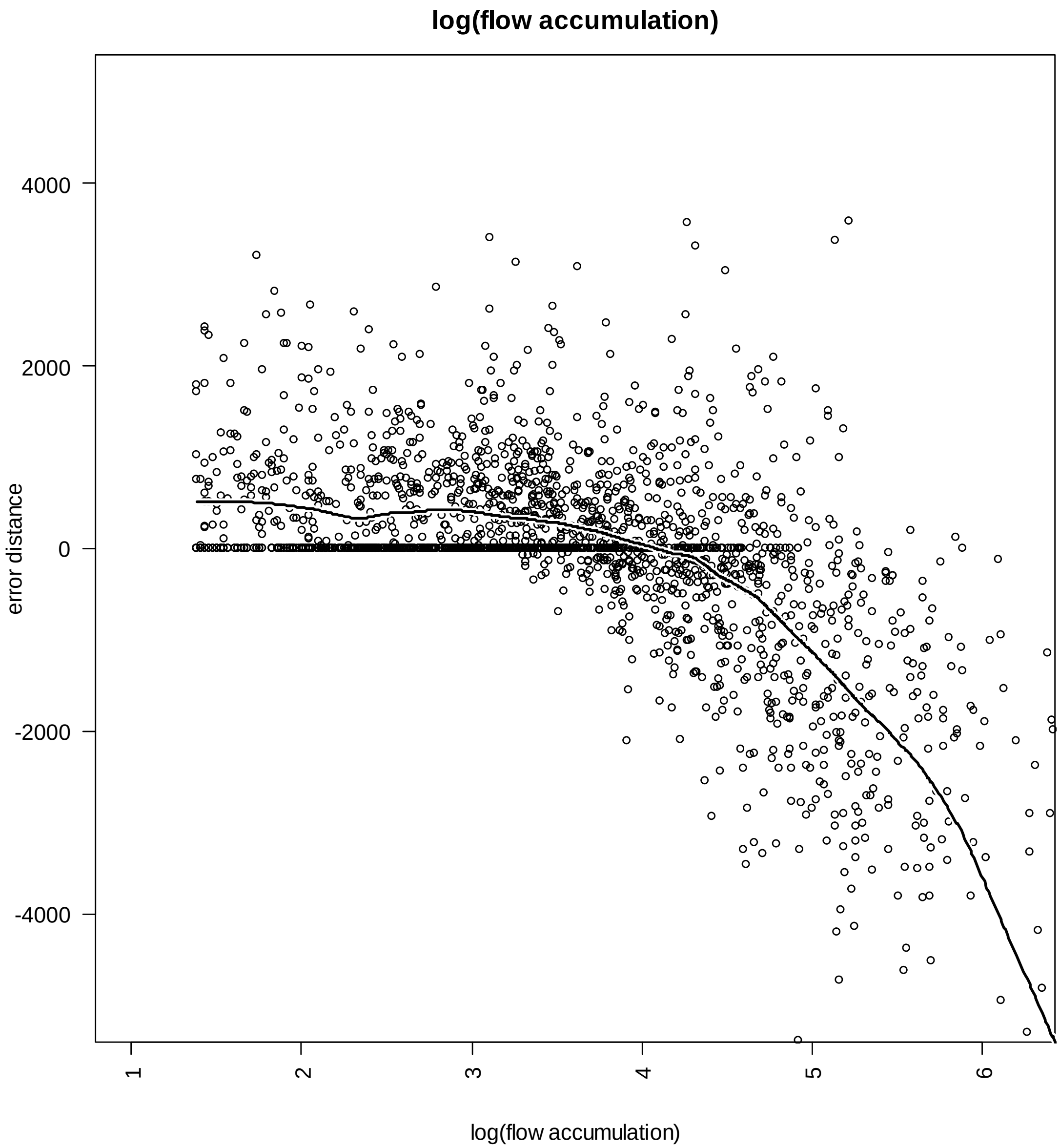


Figure 22 Scatter plot showing the relationship between error distance (in ft) and the $\log_{10}$ of flow accumulation (basin size in acres) of the end-of-fish point.

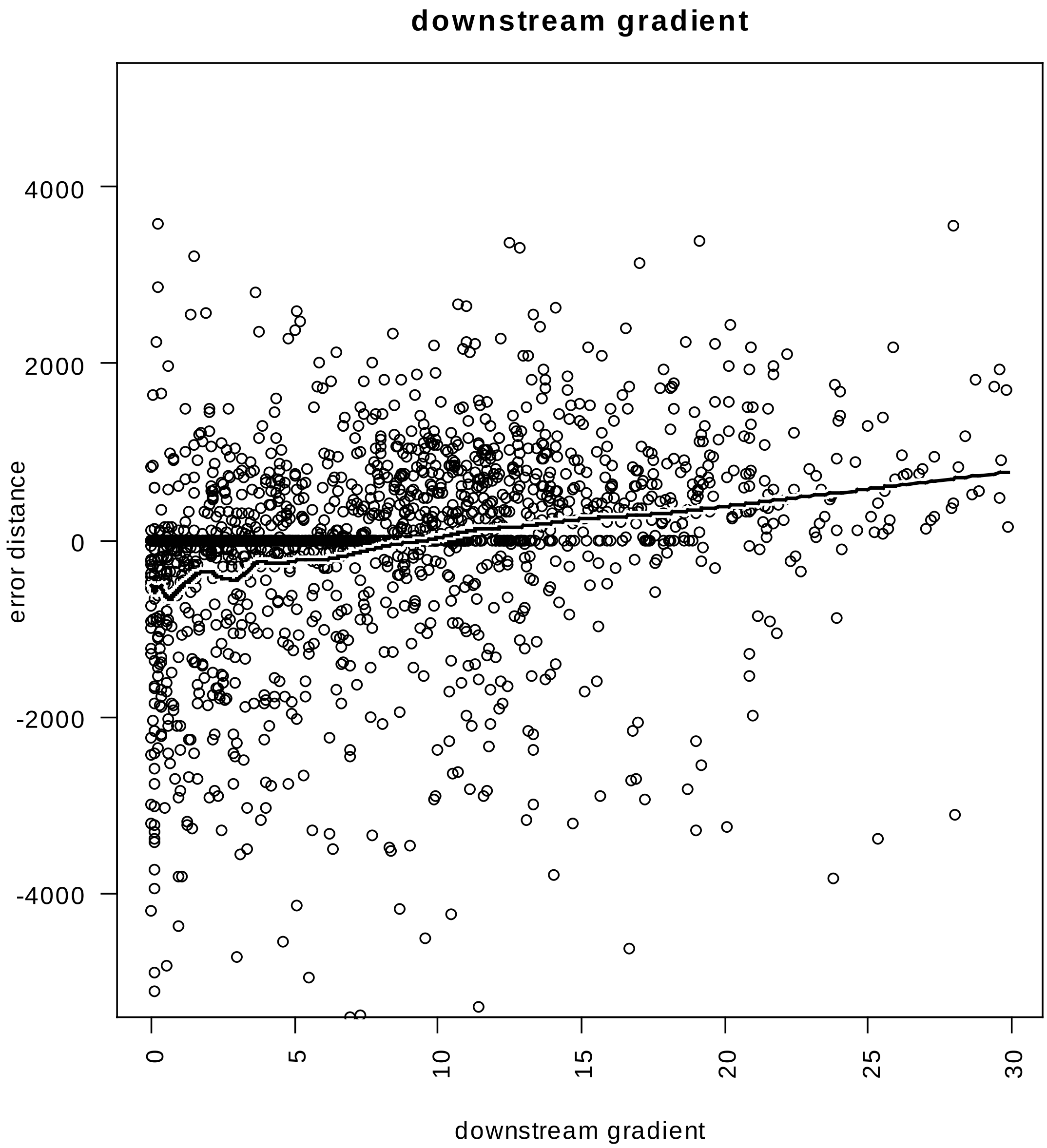


Figure 23 Scatter plot showing the relationship between error distance (in ft) and the downstream gradient of the end-of-fish point.

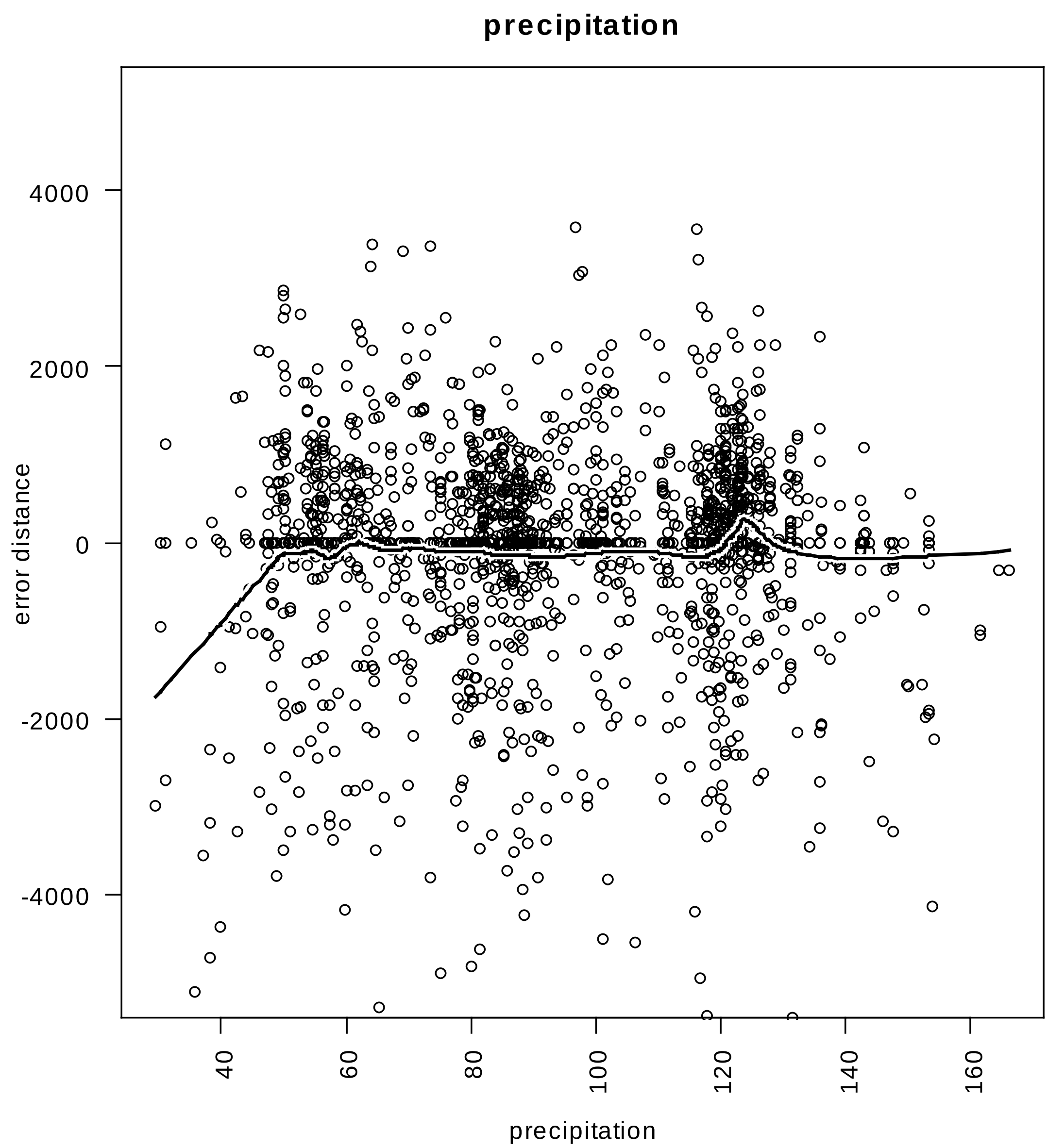


Figure 24 Scatter plot showing the relationship between error distance (in ft) and the precipitation (in inches) of the end-of-fish point.

DISCUSSION

The distribution of error distances was skewed, with more large error distances associated with over-prediction. This is likely due to blockages in fish passage that are not predictable with the available GIS coverages. The median error distance was 327 feet, with over 25% of the points predicted perfectly (zero error distance). Examining the error distribution in relation to different variables revealed some patterns. Points for which the probable cause for EOFP placement (end type) was due to gradient or stream size were more accurately predicted than those points associated with more difficult to predict end types such as those due to road culverts, mass wasting events, and poor water quality. Points placed using the LFH (last fish habitat) protocol tended to be higher in the stream network than the LF (last fish) points. Also, EOFPs at the mouth of tributaries (type B) tended to have much smaller errors than the other two boundary types. Because type B points were under represented in the survey data, measures of average absolute error and average error derived from the string-based errors will tend to be biased. This is addressed in the next section.

Because many of the variables are correlated it is difficult to speculate why error distributions vary across these variables. For example, sponsors tended to collect their data in one location, making it difficult to distinguish between patterns in error due to geography from error patterns due to sponsor. Similarly physical characteristics derived from the 10m DEM network (elevation, gradient and flow accumulation) all tend to be correlated, frustrating attempts to attribute variability in error to a specific variable.

PART 2: PRECISION AND BALANCE ESTIMATES BASED ON FOURTH ORDER SUB-BASINS

INTRODUCTION

In Part 1, we used the validation strings to explore the effects of various covariates on model error. Error distances were used, but the values should be considered relative rather than absolute. Computing realistic estimates of model precision and balance requires that the structure of the stream network be taken into account. In Part 2 we describe the approach we used to make estimates of model precision and balance which account for the structure of the stream network using the available data.

Problems with String-based Estimates of Precision and Balance

The error results presented in the previous section are based upon error measured in feet per string. Using these numbers directly to compute total error for a basin or any other section of the stream network would result in a serious overestimate of error. There are two reasons for this. The first is that strings can overlap so that sections of a string that are in error can overlap with error sections in other strings and be counted two or more times. This multiple counting overestimates the total error. Figure 25 illustrates this problem. The section of stream below the confluence of the two streams with surveyed EOFPs (labeled duplicated error) would be included in the error estimate for both strings (and hence counted twice).

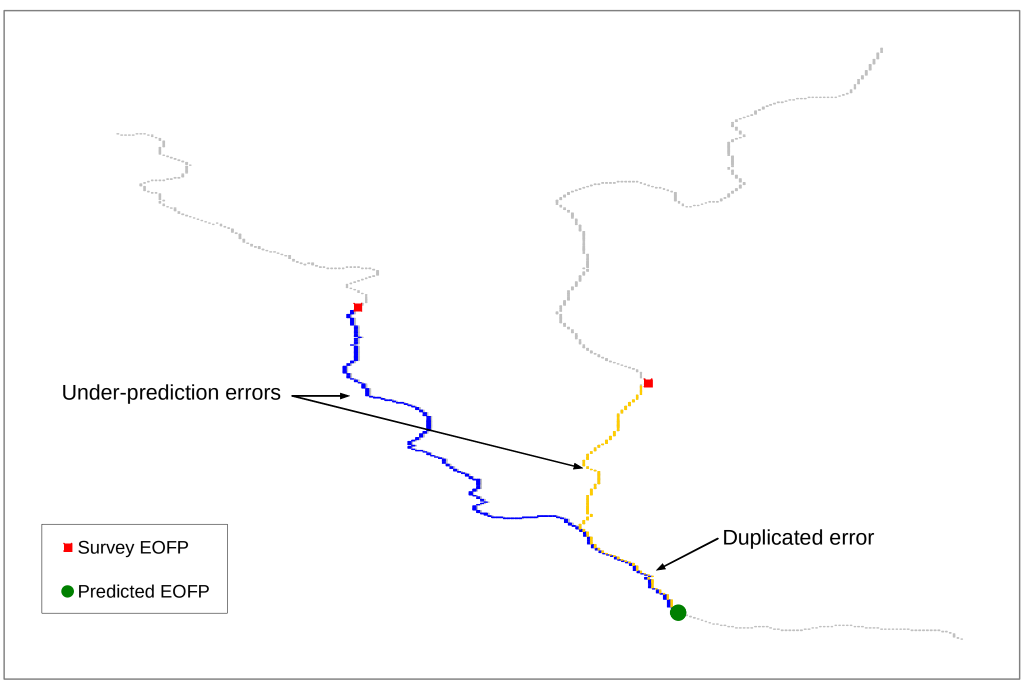


Figure 25 Schematic illustrating how strings can overlap and result in some error being double-counted.

The second reason strings tend to overestimate total error is that the string data contain a smaller proportion of lateral confluence EOFPs than are found in the stream network. Lateral confluences (Figure 2) are side tributaries entering a main channel and are frequently associated with a gradient break and/or a large change in basin area. The string validation data has 43% laterals, while stream network estimates of laterals present in a basin range from 70% to 78%. As Table 29 shows, the model and stopping rule tend to have less error on laterals so having them weighted correctly, as they would be if error were assessed for an entire stream network, would tend to reduce the total error.

Table 29 Validation string results from the final model and stopping rule showing the greater accuracy in predicting lateral confluence EOFPs.

Type of EOFP	10% Trimmed Mean Error (ft)	Strings with Zero Error
Mid-channel (A)	118	1%
Lateral Confluence (B)	-48	68%
Tributary Junction (C)	196	10%

The final stopping rule is reasonably balanced on the strings in that the 10% trimmed mean error is close to zero. This indicates that the total distance in under-prediction error is close to the total distance in over-prediction error *on the strings*. Given the problems with measuring error on the strings, however, balance on the stream network may be different.

Complete Basin or Sub-basin Error Assessment

Earlier sections of this report have described the current dataset and some of its shortcomings. In addition to such problems as varying protocols, however, there is a structural problem that makes the dataset as it stands unsuitable for assessing model performance on the stream network. Although there are over 4,000 surveyed EOFPs in the data, they tend to be widely scattered on the western Washington stream network. This is not inherently a problem for building the logistic regression model or the stopping rule, neither of which make use of the stream networkís structure. Making realistic estimates of precision and balance, however, does require accounting for the network structure of the GIS data. Multiple counting of errors must be avoided and boundary types, especially lateral confluences, must be present in the correct proportion. Data from basins or sub-basins that have been completely surveyed are needed for a more realistic error assessment. Unfortunately, there is very little of this type of data currently available. The Stillman Creek basin data were the only complete basin data set available to us. Therefore, we developed a method for approximating the results that would be obtained from entire sub-basin surveys using data from sub-basins that had many survey points.

Stillman Creek Basin:

There were data available from one basin which had been completely surveyed. This was the Stillman Creek basin in WRIA 23 in southwestern Washington. Stillman Creek is a tributary of the East Fork of the Chehalis River. Only a portion of the Stillman Creek data were included in the logistic regression model building, stopping rule development, and previous model error assessment analyses. The basin contains approximately 30,000 acres and has approximately 1,810,000 feet (343 miles) of stream length. The basin was completely surveyed by crews from the Quinault Indian Nation and Weyerhaeuser Company. We applied the final logistic regression model and stopping rule throughout the Stillman basin and assessed the precision and balance of the errors associated with the model predictions. Although it is only one basin, we use these results as a point of reference, comparing them to the results we obtained using the sub-basin error assessment method described below.

Overview of the Fourth Order Sub-basin Error Assessment Method:

We tried to make a realistic estimate of the precision and balance of model error using the data available from a number of fourth order sub-basins which had relatively many survey points. We applied the final logistic regression model and stopping rule throughout each FOSB to obtain predicted EOFPs. For each prediction that did not have a corresponding survey point in the sub-basin, we assigned an error at random from a pool of realistic model errors. After doing this for each prediction, we could compute error totals for the entire FOSB. We did this for each of our chosen FOSBs and then calculated mean estimates of precision and balance.

By using only FOSBs, we ignore those parts of the stream network that occur in higher order streams where the model tends to perform very well. The net result is that our FOSB-based estimates will tend to overestimate the true total error.

Measuring Precision and Balance:

We used the average amount of absolute error in a prediction as a measure of precision. For the purposes of precision, over-prediction errors and under-prediction errors do not offset each other. We measured precision in two ways, feet of absolute error per mile of stream and feet of absolute error per prediction.

Feet/mile has the advantage of corresponding directly to the popular measure " % correct." Figure 26 illustrates the relationship between these two measures. It is useful to know that a ±1% change in % correct corresponds to a change of ±53 ft/mile on the stream network.

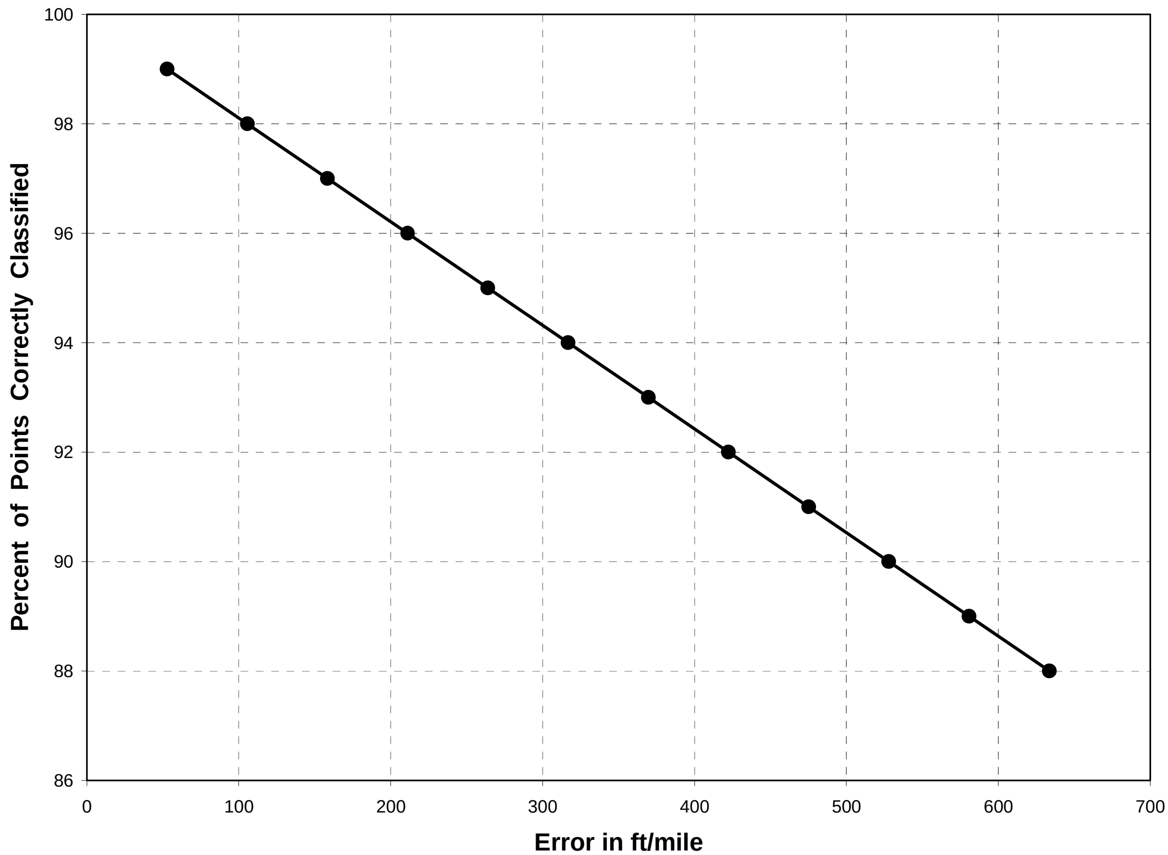


Figure 26 Schematic showing the relationship between two common measures of precision.

Both feet/mile and % correct require conventions on the upper and lower limits of streams in a basin in order to get the total stream length. Our convention for the upper limit is a GIS-based measure that amounts to 3.7 acres of basin area. Our convention for the lower limit is the lower boundary of the basin or FOSB.

The other measure of precision, feet/prediction, is comparable to the error measured on strings. It is mainly useful for comparing string results to stream network results.

Balance refers to the tendency of the distribution of errors to be centered upstream or downstream of the true EOFP. We measure balance as the net error per mile or per prediction. The ideal balance value, 0 ft/mile, would indicate that there was no bias in either the upstream or downstream directions. As with precision, the balance measure of feet/prediction is mainly useful for comparing network results to string results.

METHODS

Selecting Fourth Order Sub-basins for Error Assessment

We wanted FOSBs with relatively many survey points but we also wanted as broad a representation of WRIAs as possible. Starting with a complete list of FOSBs, we decided to limit our attention to those with at least eight survey points. This gave us a reasonable compromise between density of survey points and number of basins; 86 FOSBs were selected. Table 30 summarizes the distribution of the selected FOSBs across WRIAs. Figure 27 shows a map of one of the FOSBs as an example.

Table 30 Summary of FOSBs used for precision and balance estimates. Each FOSB had eight or more survey points.

WRIA	Number of FOSBs
7	1
20	13
21	15
22	8
23	40
24	7
26	1
27	1
Total	86

Sub-basin Error Assessment Process

The following steps were involved in making an estimate of precision and balance for each of the 86 selected FOSBs.

1. Use the final LR model and stopping rule to predict EOFPs throughout the FOSB.
2. Match up the predictions with the survey points, where they exist, and compute error distances.
3. Draw random errors with replacement for the predictions that do not have corresponding survey points (see following sub-section).
4. Reconcile any inconsistencies and eliminate duplicated error. In some cases, for example, additional predictions must be made on laterals that are downstream from an imputed EOFP that results from a randomly assigned error. The detailed list of rules for this reconciliation process is given in Appendix C.
5. Add up over-prediction and under-prediction errors.

Final estimates were then be made by averaging the individual FOSB results weighted by total stream length.

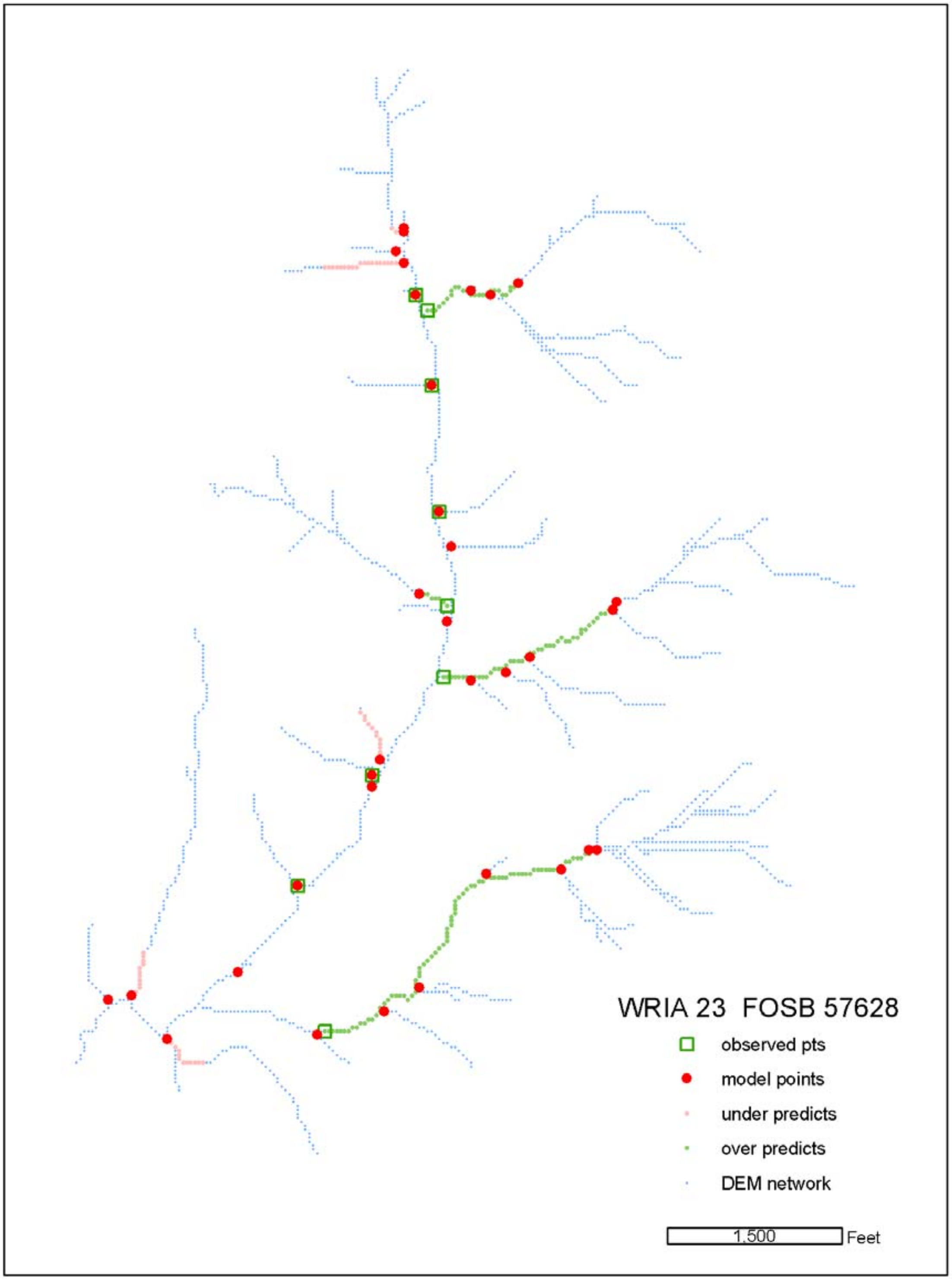


Figure 27 Schematic showing an example FOSB with 9 observed points (survey points) and 33 modeled points (predictions).

Errors for Random Assignment to Predictions Without Corresponding Surveys:

A critical aspect of this method was the assignment of random errors to predictions that had no corresponding survey point. We started with a pool of errors generated by applying the final model and stopping rule to the 2,042 withheld validation strings.

As discussed earlier, the string errors contain a number of outliers. We did not want to include unusual errors in our estimates so we removed outliers from the pool of candidate errors. We used an outlier detection method based on the definition of outliers in the standard boxplot (Hoaglin et al. 1983). Treating each boundary type separately, we deleted as outliers errors smaller than the 1st quartile ñ (2.5 x IQR)[20] or larger than the 3rd quartile + (2.5 x IQR). This resulted in deleting 81 type A errors (mid-channel), 21 type B errors (lateral confluence), and 32 type C errors (tributary junction).

In addition to removing error distance outliers, we carefully considered the question of stratifying the errors. Although it would be possible to assign the errors completely at random, that would only be appropriate if there were no strong relationships between the error distances and other characteristics of the EOFP. It was clear after examining the other available variables that the errors should first be stratified by boundary type. We found no additional stratification that would be useful within the mid-channel (A) and tributary junction (C) boundary types, but further grouping the lateral confluence (B) errors by flow accumulation (i.e., basin area) proved useful. After looking at several alternatives, we grouped the lateral errors in 10-acre classes. The final list of errors for use in the random error assignment process is summarized in Table 31. For the randomization process, the errors were drawn randomly with replacement.

Table 31 Summary of the pool of errors used for random assignment summarized by the strata used for the randomization process.

Predicted Point Type	Total Number of Errors	Number of Zero Errors	Number of Under-predictions	Number of Over-predictions	Percent Zero Errors	Average Under-prediction (ft)	Average Over-prediction (ft)
A	788	28	405	355	4%	862.8	-1,035.6
B: 0-10 ac	115	101	14	0	88%	328.1	
B: 11-20 ac	156	130	26	0	83%	353.3	
B: 21-30 ac	95	71	24	0	75%	363.7	
B: 31-40 ac	63	46	17	0	73%	475.9	
B: 41-50 ac	29	13	16	0	45%	447.9	
B: >50 ac	64	30	34	0	47%	437.1	
C	598	235	197	166	39%	708.9	-901.7

[20] IQR = Inner Quartile Range = 3rd Quartile-1st Quartile. The IQR measures the width of the central 50% of the data distribution.

Sensitivity Analysis:

We wanted to assess the impact of the random error assignment on our final estimates of precision and balance. To do this we conducted the entire process 10 times to obtain a distribution of results.

RESULTS

Stillman Creek Basin Results:

Applying the model to the Stillman Creek basin, which was fully surveyed, resulted in a precision estimate of 297 feet/mile or about 94% correct. The balance estimate for the Stillman Creek data is -226 feet/mile, which is a net over-prediction.

Four of the 86 FOSBs from the FOSB error process are located in the Stillman Creek basin. Table 32 summarizes the results from these FOSBs compared to the basin as a whole. The precision estimates are in reasonable agreement but the balance estimates are not. The balance estimate from the Stillman FOSBs of -1 foot/mile is quite different from the whole-basin value of -226 feet/mile.

Table 32 Estimates of model error precision and balance for the four FOSBs located in Stillman Creek basin and for the Stillman Creek basin as a whole.

FOSB	Under (ft)	Over (ft)	Network Length (ft)	% Correct	Abs. error (ft per mile)	Balance (ft per mile)
11317	4,012	-5,432	152,768	94	326	-49
19353	1,643	-1,659	71,788	95	243	-1
58130	1,643	-2,220	128,010	97	159	-24
60668	2,093	-197	81,407	97	149	123
Weighted mean				95.6	230	-1
Total Basin	12,226	-89,703	1,810,462	94	297	-226

Results for the FOSB Error Assessment Method

The FOSB error assessment process estimates an average absolute error for the model of 445 feet/mile or about 92% correct. We expect this to be a conservative estimate, because the FOSBs do not include a significant number of mainstem and mainstem tributary miles where the model predicts very well. The balance estimate is 170 feet/mile of net error, which is a net under-prediction. As described later, we have conflicting estimates of balance and are not confident in which balance estimate best indicates model performance.

Table 33 summarizes the 10 runs of the FOSB error assessment process. Each run involved evaluating each of the 86 FOSBs for under-prediction and over-prediction errors. Variability in the results of the 10 runs comes from differences in the random assignment of errors to predictions without corresponding survey points.

Table 33 Results from running the FOSB error assignment process 10 times.

Run Number	Precision		Balance	
	Absolute Error (ft/mile)	Absolute Error (ft/model prediction)	Net Error (ft/mile)	Net Error (ft/model prediction)
1	440	159	173	62
2	449	162	169	61
3	455	165	176	64
4	447	162	169	61
5	450	163	173	63
6	443	160	176	64
7	440	159	171	62
8	433	157	163	59
9	459	166	171	62
10	435	157	161	58
Mean	445	161	170	62
Std Dev	8.6	3.1	5.0	1.8

Final Estimates of Precision and Balance

Table 34 compares the three available estimates of precision and balance for the final model. These estimates come from the validation strings, the FOSB error assessment process, and the Stillman Creek entire basin assessment. The differences in the measures are noted in the table. For example, the string measure of precision is based on the median while the FOSB measure is a mean. Nevertheless, these measures are fairly comparable and it is instructive to compare them.

The string measures must be compared to the FOSB measures on the basis of feet/prediction. The smaller absolute error estimate from the FOSB error process (161 ft/prediction) appears to reflect the improvement we obtained by eliminating the duplicate counting that is inherent in the string process and having a realistic proportion of lateral confluence points.

At first glance, the absolute error estimates from the FOSBs (445 ft/mi) and Stillman Creek basin assessment (297 ft/mi) appear quite different. The % correct values of 92% for the FOSBs and 94% for Stillman help put the ft/mile values in perspective, however.

Table 34 Summary of the available estimates of precision and balance for the final model.

Dataset	Precision			Balance	
	Absolute Error (ft/ model point)	Mean Absolute Error (ft/mile)	% Correct	Error (ft/ model point)	Mean Error (ft/mile)
Validation strings	327[a]	-	-	37[c]	-
FOSBs	161[b]	445	92	62[b]	170
Stillman total basin	-	297	94	-	-226

[a] Median [b] Mean [c] 10% Trimmed Mean

In terms of balance, the validation strings and the FOSBs are in good agreement. The FOSBs and Stillman basin assessment, however, are in sharp disagreement. The FOSB error assessment process results in an estimated mean net under-prediction of 170 ft/mile. For the Stillman Creek basin, however, there was a mean net over-prediction of -226 ft/mile.

DISCUSSION

We have made several estimates of precision and balance (Table 34), but none of them is as reliable or precise as we would like. We have used the available data to the fullest extent possible.

The string estimates of error are reliable, but cannot be used to directly estimate stream network results. The FOSB error assessment process represents our best attempt to overcome the limitations in the string dataset and make realistic estimates of precision and balance for the stream network. As noted earlier, however, the estimates of error from the FOSB error assessment process are very likely to be greater than the true value. Stillman Creek is a fully surveyed basin but represents only a single location.

The overall conclusion we draw from comparing the estimates is that the preliminary model is likely to be reasonably close to the target 95% correct in terms of precision, where percent correct is based on the stream distance correctly classified as fish or no fish. In terms of balance, our measures are too much in conflict to draw any firm conclusion.

PART 3: FURTHER CONSIDERATIONS

The discussion for this section focuses on sources of errors that are not explicitly accounted for in the previous analyses.

Potential Sources of Error Not Explicitly Accounted for in the Analyses

Errors Due to GIS Data and Manipulation:

There are significant errors introduced to the modeling process by the GIS coverages used to place end-of-fish points geographically and to generate physical attributes for points in the stream network. Errors occur in:

1. The DNR stream network: This affects placement of observed EOFPs and the location of streams in the final map delineating fish habitat. Rosenfeld et al. (2001) estimated that in lower gradient topographies, 1:20,000 scale topographic maps underestimated fish habitat by 34.2% to 100%. Fish use of ephemeral streams (e.g. Hartman and Brown 1986) exacerbates this problem.
2. The Digital Elevation Map (DEM): The 10m DEMs (each point on a 10-meter grid has an elevation) used for this project are derived from human interpretation of aerial photographs. Fine scale topography associated with small streams is difficult to see on aerial photographs, especially with the dense forest cover typical of Western Washington.
3. Generation of the stream network and physical attributes: The DEMs were used to generate the stream network (used for modeling) and the attached physical attributes (gradient and flow accumulation). The algorithms used to generate the stream networks are sensitive to errors in the DEMs and perform poorly in low gradient areas.
4. Translation between maps: The EOFPs were recorded on the DNR stream layer and then translated to the 10m DEM-generated stream points network used for modeling. When the modeling was completed the predicted EOFPs were then translated back to the DNR coverage. Because the two maps were constructed using completely different methods, there were many streams that did not align. This led to potential errors associated with the translation process.
5. Precipitation data: The precipitation data is only available on a very course resolution. This creates sudden shifts in predicted precipitation when moving from one square in the grid to another. Local topographic effects on precipitation are not included at this resolution.

Data Collected in the Field:

The stream surveys used to identify EOFPs also introduced error into the modeling process. There are a number of potential sources of this error.

1. Different organizations and individuals collected the data.
2. Different protocols were used (e.g. Last Fish, Last Fish Habitat, Last Salmonid, Ö).
3. Data were collected in different seasons and years.
4. Failure to see a fish that is present during a survey. If a fish is seen at a location, the habitat is clearly fish habitat. However, if a fish is not seen, the location may still contain fish. This error leads to an underestimation of fish habitat.
5. Placement of an EOFP on the map. Finding your location on a map in the field can be non-trivial when working in small streams.

Potential Fish Habitat Versus Fish Observations:

This model predicts the point above which you are unlikely to observe a fish if you follow the field methodology used to collect the data. Potential fish habitat, as defined in the forest fish agreement, is " habitat which is used by any fish at any life stage at any time of the year, including potential habitat likely to be used by fish which could be recovered by restoration or management and includes off-channel habitat" .

CONCLUSIONS AND RECOMMENDATIONS

The Fish Habitat model requested in the Forest and Fish Report is complete. The preliminary model described in this report (final logistic regression equation and stopping rule) and resulting assessments represent the best possible science given the limitations of the available data and guidelines specified in the Forest and Fish Report.

We recommend caution when interpreting the model results due to:

1. The unknown variability in the field survey protocols (both between surveys with different protocols and among surveys with the same protocol) which were used to establish the "end-of-fish" points for this study;
2. The non-random nature in which the data that were the basis for the analyses were collected; and
3. The methods used to estimate the precision and balance of the model predictions were not ideal because entire basins were not covered by field surveys.

The "found data" that were used for this project presented problems in model development and assessment that cannot be addressed at this time. Specifically:

- Differences in the survey protocols among the end-of-fish points used to generate data for the model prevent assessment of the model against a single target.
- The proportion of field surveyed end-of-fish points representing various situations (e.g., mid-channel, lateral confluence, and tributary junction boundary types) is not representative of the real world in both the model building and model validation data sets.
- Field survey data collection was not intended for this purpose. There was no deliberate sampling design which resulted in a non-random geographic distribution of the end-of-fish points and the Fish Absent | Fish Present data available for both model building and assessment of the final model.

There are some situations where the model will not work as well as it does in others:

- Predictions where fish stop along continuous stream segments (mid-channel boundary types);
- Streams with barriers preventing access by fish;
- Streams containing headwater lakes (e.g., stocked lakes);
- Geographies and situations not represented in the available survey data;
- Areas where modification of the stream network is prevalent (e.g., agricultural, residential, or urban areas);
- Areas of very low topographic relief (e.g., flood plains); and
- Areas with atypical hydrology (e.g., springs or unusually large dry channels).

We are reasonably confident that, on average, the distance between a predicted EOFP and the "true" EOFP will be less than 500 ft. This is largely based upon the results presented in Table 34 for the fourth order sub-basin error assessment and their comparability to the results from the Stillman basin assessment:

- On average, 92% of the habitat in fourth order sub-basins will be correctly classified (assuming a four acre channel initiation point) based on the FOSB assessment compared to 94% for the assessment based on the Stillman basin.
- On average, there is an absolute error of ±445 ft per mile of stream classified based on the FOSB assessment compared to ±297 ft/mile for the Stillman basin assessment.
- The assessment of precision for fourth order sub-basins (±445 ft per mile of stream classified) is likely an over-estimate of the error. The overall error would be smaller (<445 ft per mile) if larger order streams (e.g., mainstem channels) were included in the data used to make the estimate.

We are <u>not</u> confident in the assessment of the balance of model errors because of the conflicting results of the FOSB and Stillman basin error assessments:

- +170 ft per mile of stream classified in fourth order sub-basins (excluding mainstem streams).
- -226 ft per mile of stream classified in the Stillman Creek total basin survey.

<u>Recommendations</u>

We recommend that the preliminary model be implemented with consideration of the following:

- ➢ Where site-specific field survey information exists, it should be considered in addition to the model predictions. Natural variability in the last fish location, unexplained by the model, is likely to be much larger than observational error. This suggests that where survey points exist, they should be substantially more accurate than the predicted point.

- ➢ More field surveys need to be conducted to provide a more reliable assessment of model performance and to update and improve the model. These data should be collected following a statistically-based sampling design which includes a consistent and easily reproducible field protocol. This will require a more precise description of both the quantitative goals and the definition of "potential fish habitat". The current model results should be used when constructing the sampling design in order to focus some of the effort on specific points or types of points where the errors were largest.

- ➢ Alternative data sources and statistical modeling approaches should be explored to improve and streamline future stream-typing models. For example, including data from a GIS lithology layer may provide useful information for modeling. Statistical methods other than logistic regression analysis may provide more efficient techniques for predicting end-of-fish points.

- Where model predictions are in areas of greatest uncertainty, such as mid-channel or above known barriers, additional surveys could be used to improve stream typing precision.

- Situations that might not be appropriate for inclusion in the next generation of modeling data should be reviewed and a protocol for exclusion of data developed (e.g., stocked headwater lakes with high gradient reaches below them or unusually large streams with barriers preventing fish use).

- The assumption that all habitat downstream from a known fish is suitable fish habitat for the purposes of model development should be re-examined. Including all stream points below known fish could adversely impact model and stopping rule performance.

- The process for transferring survey points to the DEM stream network needs to be reviewed and standardized. Points should be placed where DEM gradient and stream network cues closely match the conditions observed in the field (e.g., gradient breaks or tributary junctions).

REFERENCES

Bahls, P., and M. Ereth. 1994. Stream typing error in Washington water type maps for watersheds of Hood Canal and the southwest Olympic Peninsula. Point No Point Treaty Council Technical Report. TR 94-02.

Baxter, B., and M. W. Mobbs. 1992. Quinault Indian Nation stream type verification project annual report. *Unpublished manuscript*. Quinault Indian Nation Technical Report, Taholah, WA.

Baxter, B. 1993. Johns River stream type verification project. *Unpublished manuscript*. Quinault Indian Nation Technical Report, Taholah, WA.

Chambers, J. M., W. S. Cleveland, B. Kleiner, and P. A. Tukey. 1983. *Graphical Methods for Data Analysis*. Duxbury Press, Boston.

Cole, M. B., and J. L Lemke. 2003. Eastern Washington last fish variability characterization resurvey: final report. CMER-ISAG-02-__ [to be numbered], Washington Departments of Natural Resources and Fish and Wildlife, Olympia, WA.

Daly, C., C. Taylor, and G. Taylor. 1998. 1961-1990 Mean Monthly Precipitation Maps for the Conterminous United States. Oregon State University Spatial Climate Analysis Service.

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

Forest Practices Board (FAB). 1996a. Washington Administrative Code (WAC) 222-16-030. Emergency rule adopted November, 14, 1996.

Forest Practices Board (FAB). 1996b. Washington Administrative Code (WAC) 222-12-090. Forest Practices Board Manual "Guidelines for determining fish use for the purpose of typing waters under WAC 222-16-030".

Fransen, B. R., S. D. Duke, G. McWethy, J. K. Walter, and R. E. Bilby. 2003. A landscape scale model for predicting the upper extent of fish-bearing streams. *Unpublished manuscript*. The Weyerhaeuser Co., Federal Way, WA.

Fransen, B. R., S. Needham, G. McWethy, and V. Kim. 1997. Development of a process to delineate potential fish habitat based on physical characteristics measured at the upper extent of known fish distribution. *Unprocessed report* to the Timber, Fish, and Wildlife Water Typing Committee.

REFERENCES (continued)

Freidman, J. H. 1984. A variable span scatterplot smoother. Laboratory for Computational Statistics Technical Report No. 5. Stanford University, Palo Alto, CA.

Goutte, C. 1997. Note on free lunches and cross-validation. Neural Computation 9:1211-1215.

Hartman, G. F., and Brown, T. G. 1987. Use of small, temporary, floodplain tributaries by juvenile salmonids in a west coast rain-forest drainage basin, Carnation Creek, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 44:262-270.

Hoaglin, David C., Frederick Mosteller, and John W. Tukey, eds. 1983. Understanding Robust and Exploratory Data Analysis. John Wiley and Sons, New York.

Hosmer, D. W. and S. Lemeshow. 1989. Applied Logistic Regression. John Wiley and Sons, New York.

Knapp, R. A. and H. K. Preisler. 1999. Is it possible to predict habitat use by spawning salmonids? A test using California golden trout (Oncorhynchus mykiss aquabonita). Canadian Journal of Fisheries and Aquatic Sciences 56:1576-1584.

Kruse, C. G., W. Hubert, and F. J. Rahel. 1997. Geomorphic influences on the distribution of Yellowstone cutthroat trout in the Absaroka Mountains, Wyoming. Transactions of the American Fisheries Society 126:418-427.

Mobbs, M. W., J. James, and B. Baxter. 1995. Newaukum River stream type verification project. Unpublished manuscript. Quinault Indian Nation Technical Report, Taholah, WA.

Needham, S. 2001. Fish Habitat Reclassification Process for Western Washington. Unprocessed report to the Washington Dept. of Natural Resources under contract number PSC 01-056. March 30, 2001.

Norusis, M. J. 1999. SPSS Regression Models. SPSS Inc. Chicago, IL.

Paul, A. J., and J. R. Post. 2001. Spatial distribution of native and nonnative salmonids in streams of the eastern slopes of the Canadian Rocky Mountains. Transactions of the American Fisheries Society 130:417-430.

Porter, M., S. J. Rosenfeld, and E. A. Parkinson. 2000. Predictive models of fish species distribution in the Blackwater Drainage, British Columbia. North American Journal of Fisheries Management: 20:349-359.

REFERENCES (continued)

Rieman, B. E., and J. D. McIntyre. 1995. Occurrence of bull trout in naturally fragmented habitat patches of varied size. Transactions of the American Fisheries Society 124:285-296.

Rosenfeld, J., S. MacDonald, D. Foster, S. Amrhein, B. Bales, T. Williams, F. Race, and T. Livingstone. 2002. Importance of small streams as rearing habitat for coastal cutthroat trout. North American Journal of Fisheries Management 22:177-187.

Swartzman, G. L., and S. P. Kaluzny. 1983. Ecological Simulation Primer. MacMillan Publishing Co., New York.

Watson, G.,and T. W. Hillman. 1997. Factors affecting the distribution and abundance of bull trout: an investigation at hierarchical scales. North American Journal of Fisheries Management: 17:237-252.

APPENDIX TABLES

Appendix Table 1 Summary of the number of end-of-fish points (EOFPs) collected using the pre-emergency rules survey protocol by Washington Trout and excluded from the logistic regression model building process, by end-of-fish point type and WRIA.

EOFP	WRIA	End-of-Fish Point Type[a]				Total
		LF	LFH	LS	LSH	
Fourth Order Sub-basin (FOSB)	**1**		2	3		5
	5		3			3
	7		23	7		30
	10		7	6		13
	11		9	3		12
	14		2			2
	23		1			1
	Total	0	47	19		66[b]
Unassigned	**1**		2	1		3
	5			1		1
	7		13	3		16
	10		5	1		6
	11	2	6		1	9
	Total	2	26	6	1	35

[a] LF = Last Fish; LFH = Last Fish Habitat; LS = Last Salmonid; and LSH = Last Salmonid Habitat.

[b] Because our analysis methods were based on using FAFP data associated with fourth order sub-basins (FOSB), there were an additional five EOFPs collected by other agencies that were excluded because they were in the same FOSB as one of the excluded Washington Trout EOFPs.

Appendix Table 2 Summary of the mean, standard error, median, and range of basin size for fish absent and fish present points by summary group; fish absent | fish present (FAFP) data pool, model validation data, and unassigned data.

Fish Status	Group	Mean	Number of Points	Standard Error of Mean	Median	Minimum	Maximum
	FAFP Data Pool	62.9	270,163	0.36	14.0	4.0	3,018.4
Fish Absent	Validation Data	44.5	187,237	0.22	13.4	4.0	1,401.4
	Unassigned Data	81.4	168,086	0.88	14.2	4.0	7,102.7
	Combined	62.4	625,486	0.29	13.9	4.0	7,102.7
	FAFP Data Pool	733.9	130,516	1.96	528.6	4.0	5,797.4
Fish Present	Validation Data	723.0	103,486	2.25	473.4	4.1	4,010.8
	Unassigned Data	128,363.1	583,983	377.12	20,181.9	4.0	1,769,485.9
	Combined	91,850.7	817,985	276.69	6,759.3	4.0	1,769,485.9
	FAFP Data Pool	281.5	400,679	0.84	37.2	4.0	5,797.4
Total	Validation Data	286.0	290,723	1.01	37.4	4.0	4,010.8
	Unassigned Data	99,692.4	752,069	299.25	9,200.6	4.0	1,769,485.9
	Combined	52,076.9	1,443,471	161.30	296.1	4.0	1,769,485.9

Appendix Table 3 Summary of the mean, standard error, median, and range of elevation for fish absent and fish present points by summary group; fish absent | fish present (FAFP) data pool, model validation data, and unassigned data.

Fish Status	Group	Mean	Number of Points	Standard Error of Mean	Median	Minimum	Maximum
	FAFP Data Pool	1,792.8	270,163	2.62	1,520.5	5.2	7,153.5
Fish Absent	Validation Data	1,713.5	187,237	3.38	1,263.5	3.0	6,495.5
	Unassigned Data	1,875.2	168,086	3.64	1,443.0	2.0	6,504.3
	Combined	1,791.2	625,486	1.81	1,426.1	2.0	7,153.5
	FAFP Data Pool	614.0	130,516	1.64	421.5	1.0	3,981.6
Fish Present	Validation Data	586.3	103,486	1.62	425.2	0.9	3,500.0
	Unassigned Data	435.6	583,983	0.66	255.0	-53.2	3,615.3
	Combined	483.1	817,985	0.58	303.0	-53.2	3,981.6
	FAFP Data Pool	1,408.8	400,679	2.04	977.0	1.0	7,153.5
Total	Validation Data	1,312.3	290,723	2.47	748.2	0.9	6,495.5
	Unassigned Data	757.4	752,069	1.19	366.0	-53.2	6,504.3
	Combined	1,050.0	1,443,471	1.01	548.0	-53.2	7,153.5

Appendix Table 4 Summary of the mean, standard error, median, and range of downstream gradient for fish absent and fish present points by summary group; fish absent | fish present (FAFP) data pool, model validation data, and unassigned data.

Fish Status	Group	Mean	Number of Points	Standard Error of Mean	Median	Minimum	Maximum
Fish Absent	FAFP Data Pool	20.46	270,163	0.033	16.39	0.00	134.04
	Validation Data	19.92	187,237	0.039	15.62	0.00	156.48
	Unassigned Data	23.02	168,086	0.046	17.64	0.00	142.81
	Combined	20.99	625,486	0.022	16.48	0.00	156.48
Fish Present	FAFP Data Pool	4.24	130,516	0.014	2.31	0.00	80.09
	Validation Data	4.16	103,486	0.014	2.46	0.00	43.74
	Unassigned Data	1.27	583,983	0.004	0.20	0.00	91.94
	Combined	2.11	817,985	0.004	0.45	0.00	91.94
Total	FAFP Data Pool	15.18	400,679	0.025	10.09	0.00	134.04
	Validation Data	14.31	290,723	0.029	9.18	0.00	156.48
	Unassigned Data	6.14	752,069	0.015	0.54	0.00	142.81
	Combined	10.29	1,443,471	0.013	3.03	0.00	156.48

Appendix Table 5 Summary of the mean, standard error, median, and range of upstream gradient for fish absent and fish present points by summary group; fish absent | fish present (FAFP) data pool, model validation data, and unassigned data.

Fish Status	Group	Mean	Number of Points	Standard Error of Mean	Median	Minimum	Maximum
Fish Absent	FAFP Data Pool	21.77	270,163	0.036	17.53	0.00	227.50
	Validation Data	21.31	187,237	0.042	17.03	-0.01	171.40
	Unassigned Data	24.39	168,086	0.050	19.15	-0.06	199.50
	Combined	22.34	625,486	0.024	17.79	-0.06	227.50
Fish Present	FAFP Data Pool	4.73	130,516	0.016	2.61	-0.10	80.09
	Validation Data	4.69	103,486	0.016	2.74	-0.07	63.38
	Unassigned Data	1.38	583,983	0.004	0.20	-0.10	91.94
	Combined	2.33	817,985	0.005	0.49	-0.10	91.94
Total	FAFP Data Pool	16.22	400,679	0.028	10.47	-0.10	227.50
	Validation Data	15.39	290,723	0.032	9.73	-0.07	171.40
	Unassigned Data	6.52	752,069	0.016	0.53	-0.10	199.50
	Combined	11.00	1,443,471	0.014	2.96	-0.10	227.50

Appendix Table 6 Summary of the mean, standard error, median, and range of precipitation for fish absent and fish present points by summary group; fish absent | fish present (FAFP) data pool, model validation data, and unassigned data.

Fish Status	Group	Mean	Number of Points	Standard Error of Mean	Median	Minimum	Maximum
	FAFP Data Pool	88.58	270,163	0.057	85.40	26.77	218.11
Fish Absent	Validation	91.19	187,237	0.069	88.98	29.13	161.42
	Unassigned	92.38	168,086	0.069	89.25	30.71	218.11
	Combined	90.38	625,486	0.037	87.40	26.77	218.11
	FAFP Data Pool	88.05	130,516	0.075	85.49	26.79	213.32
Fish Present	Validation	92.29	103,486	0.087	88.19	29.65	161.42
	Unassigned	91.01	583,983	0.035	87.14	28.73	205.14
	Combined	90.70	817,985	0.030	86.97	26.79	213.32
	FAFP Data Pool	88.41	400,679	0.045	85.43	26.77	218.11
Total	Validation	91.58	290,723	0.054	88.58	29.13	161.42
	Unassigned	88.58	270,163	0.057	85.40	26.77	218.11
	Combined	91.19	187,237	0.069	88.98	29.13	161.42

Appendix Table 7 Summary of the number of fish absent (FA) and fish present (FP) points available, by WRIA, in the FAFP data pool and the target samples sizes for model estimation data sets of different sizes.

	Target	Number of Points		Desired Number of FA or FP Points				
WRIA	**Percentage**	**FA**	**FP**	2,000	3,000	4,000	5,000	6,000
1	7.38%	4,566	315	147	221	295	369[a]	443[b]
3&4	15.26%	20,767	2,431	304	458	611	761	916
5	2.90%	16,929	7,782	59	87	116	145	174
6	0.92%	1,341	168	18	28	37	46	55
7	7.32%	8,414	7,741	146	220	293	366	439
8	2.48%	2,863	1,137	50	74	99	124	149
9	2.07%	15,489	1,470	42	62	83	104	124
10	3.93%	7,575	2,872	79	118	157	197	236
11	2.95%	2,496	955	59	89	118	148	177
13	0.96%	4,296	864	19	29	39	48	58
14	1.19%	117	118	24	36	47	59	71
15	2.75%	2,634	1,734	55	83	110	138	165
16	1.91%	2,791	1,235	38	57	76	95	115
17	1.58%	5,304	749	32	47	63	79	95
18	2.76%	20,333	139	56	83	110	138	166[c]
19	1.25%	2,233	1,400	25	38	50	63	75
20	4.20%	9,183	7,952	85	126	168	210	252
21	4.43%	8,945	7,488	89	133	177	222	266
22	4.66%	20,190	23,079	94	140	186	233	279
23	4.34%	40,434	19,063	88	130	174	217	260
24	3.16%	19,890	25,061	63	95	126	158	189
25	1.51%	1,527	914	30	45	61	76	91
26	9.67%	24,625	10,195	191	288	387	483	580
27	4.96%	15,342	3,142	98	149	198	248	297
28	1.92%	10,223	2,148	38	58	77	96	115
29	3.54%	1,656	364	71	106	142	177	213
Totals	100.00%	270,163	130,516	2,000	3,000	4,000	5,000	6,000

[a] An additional 54 FP points were randomly selected from WRIA 3&4 combined to make up the deficit.

[b] An additional 128 FP points were randomly selected from WRIA 3&4 combined to make up the deficit.

[c] An additional 14 FP points were randomly selected from WRIA 17 and an additional 13 points were randomly selected from WRIA 19 to make up the deficit.

Appendix Table 8a Summary of the results of estimating a logistic regression using 10 different model estimation data sets with 2,000 fish absent points and 2,000 fish present points.

A. Model fit and assessment statistics:

Model	Hosmer-Lemeshow		-2 x log	Self-Classification		Validation Data	
Number	Chi sq.	Signif.[a]	likelihood	Absent	Present	Absent	Present
1	8.248	0.410	1,722.3	89.6%	92.4%	91.7%	89.0%
2	14.644	0.066	1,579.0	90.9%	93.4%	91.9%	88.9%
3	10.412	0.237	1,617.7	90.4%	92.6%	91.5%	89.1%
4	8.646	0.373	1,521.7	90.8%	93.5%	92.0%	88.5%
5	18.635	**0.017**	1,576.0	90.2%	92.6%	92.1%	88.0%
6	13.861	0.085	1,498.0	91.4%	93.3%	91.9%	88.5%
7	11.384	0.181	1,588.9	90.8%	93.0%	92.2%	88.2%
8	2.308	0.970	1,560.9	91.4%	93.2%	91.3%	89.5%
9	17.474	**0.026**	1,621.0	90.8%	92.5%	91.5%	89.2%
10	15.249	0.054	1,646.7	90.1%	92.4%	91.8%	88.6%
Means	12.086			90.6%	92.9%	91.8%	88.8%

[a] Bolded significance levels are significant at $\alpha > 0.05$.

B. Estimated logistic regression model coefficients:

Number	CONSTANT	BASIZE	ELEV	UPGRD	DNGRD	PRECIP
1	-6.54940	3.45914	-0.10558		-0.07250	0.01526
2	-7.55885	3.73571	-0.11360		-0.04729	0.02122
3	-6.99252	3.58562	-0.09003		-0.09023	0.01779
4	-7.29470	3.76833	-0.10124		-0.08511	0.01694
5	-7.19707	3.74708	-0.08642		-0.07057	0.01310
6	-7.74062	3.92758	-0.10397		-0.06096	0.01670
7	-7.22992	3.70976	-0.10512		-0.07153	0.01627
8	-7.37983	3.74978	-0.11344		-0.06447	0.01922
9	-7.34048	3.60904	-0.09902		-0.07704	0.02105
10	-7.22604	3.65562	-0.09297		-0.06823	0.01667
Means	-7.25094	3.69477	-0.10114		-0.07079	0.01742
St. Dev.	0.32042	0.12631	0.00920		0.01211	0.00252
Coef. Var.	4.4%	3.4%	9.1%		17.1%	14.4%

Appendix Table 8b Summary of the results of estimating a logistic regression using 10 different model estimation data sets with 3,000 fish absent points and 3,000 fish present points.

A. Model fit and assessment statistics:

Model Number	Hosmer-Lemeshow Chi sq.	Hosmer-Lemeshow Signif.[a]	-2 x log likelihood	Self-Classification Absent	Self-Classification Present	Validation Data Absent	Validation Data Present
1	40.274	**0.000**	2,398.4	90.7%	92.8%	92.1%	88.3%
2	17.689	**0.024**	2,513.6	90.0%	92.6%	91.8%	88.9%
3	24.535	**0.002**	2,469.2	89.7%	92.2%	92.4%	87.6%
4	18.124	**0.020**	2,535.6	89.8%	92.5%	91.8%	88.9%
5	11.555	0.172	2,467.6	90.4%	92.9%	92.0%	88.4%
6	14.126	0.079	2,415.8	90.2%	92.8%	92.1%	88.1%
7	6.677	0.572	2,420.3	90.7%	92.6%	91.9%	88.6%
8	8.223	0.412	2,557.7	89.9%	92.7%	92.1%	88.3%
9	8.810	0.359	2,377.6	90.9%	93.0%	91.5%	89.2%
10	11.258	0.188	2,450.9	90.3%	92.9%	92.1%	88.3%
Means	16.127			90.3%	92.7%	92.0%	88.5%

[a] Bolded significance levels are significant at $\alpha > 0.05$.

B. Estimated logistic regression model coefficients:

Number	CONSTANT	BASIZE	ELEV	UPGRD	DNGRD	PRECIP
1	-6.99211	3.64139	-0.09860		-0.07944	0.01521
2	-6.57328	3.57237	-0.11067		-0.06435	0.01273
3	-7.12508	3.69173	-0.09312		-0.08726	0.01500
4	-6.56115	3.47407	-0.09865		-0.07976	0.01481
5	-7.32778	3.66630	-0.10184		-0.06037	0.01741
6	-6.94972	3.73589	-0.09965		-0.06863	0.01175
7	-7.51620	3.74782	-0.11253		-0.05558	0.01862
8	-7.02668	3.53269	-0.09618		-0.08541	0.01831
9	-7.12627	3.65933	-0.09895		-0.07537	0.01746
10	-6.84081	3.58343	-0.10148		-0.06567	0.01393
Means	-7.00391	3.63050	-0.10117		-0.07218	0.01552
St. Dev.	0.29975	0.08850	0.00605		0.01082	0.00236
Coef. Var.	4.3%	2.4%	6.0%		15.0%	15.2%

Appendix Table 8c Summary of the results of estimating a logistic regression using 10 different model estimation data sets with 4,000 fish absent points and 4,000 fish present points.

A. Model fit and assessment statistics:

Model Number	Hosmer-Lemeshow Chi sq.	Hosmer-Lemeshow Signif.[a]	-2 x log likelihood	Self-Classification Absent	Self-Classification Present	Validation Data Absent	Validation Data Present
1	10.691	0.220	3,102.9	90.9%	93.0%	91.8%	88.8%
2	17.700	**0.024**	3,111.7	90.7%	93.2%	91.9%	88.8%
3	16.658	**0.034**	3,268.3	90.3%	92.7%	91.8%	88.8%
4	13.153	0.107	3,264.4	90.4%	92.4%	91.8%	88.9%
5	9.152	0.330	3,169.9	90.8%	93.0%	91.7%	89.0%
6	9.025	0.340	3,219.0	90.6%	92.7%	92.1%	88.4%
7	21.691	**0.006**	3,200.3	90.3%	93.1%	92.6%	87.5%
8	12.800	0.119	3,194.0	90.3%	92.8%	91.8%	88.9%
9	14.734	0.065	3,231.5	90.5%	93.0%	92.0%	88.6%
10	14.466	0.070	3,162.9	90.8%	93.1%	91.7%	88.8%
Means	14.007			90.6%	92.9%	91.9%	88.7%

[a] Bolded significance levels are significant at $\alpha > 0.05$.

B. Estimated logistic regression model coefficients:

Number	CONSTANT	BASIZE	ELEV	UPGRD	DNGRD	PRECIP
1	-7.76135	3.88991	-0.11098		-0.06912	0.01968
2	-7.42960	3.77924	-0.11097		-0.07105	0.01828
3	-7.17478	3.63829	-0.10231		-0.07056	0.01770
4	-7.28136	3.72171	-0.10604		-0.07715	0.01809
5	-7.33115	3.68023	-0.10638		-0.08002	0.02011
6	-7.37201	3.77738	-0.10510		-0.07002	0.01644
7	-7.44779	3.76651	-0.09595		-0.07926	0.01633
8	-7.16892	3.68494	-0.10278		-0.07334	0.01703
9	-7.05989	3.60788	-0.10324		-0.08188	0.01776
10	-7.30005	3.70142	-0.09652		-0.08114	0.01823
Means	-7.33269	3.72475	-0.10403		-0.07535	0.01797
St. Dev.	0.19374	0.08166	0.00510		0.00505	0.00123
Coef. Var.	2.6%	2.2%	4.9%		6.7%	6.9%

Appendix Table 8d Summary of the results of estimating a logistic regression using 10 different model estimation data sets with 5,000 fish absent points and 5,000 fish present points.

A. Model fit and assessment statistics:

Model Number	Hosmer-Lemeshow Chi sq.	Hosmer-Lemeshow Signif.[a]	-2 x log likelihood	Self-Classification Absent	Self-Classification Present	Validation Data Absent	Validation Data Present
1	20.566	**0.008**	4,151.2	90.0%	92.8%	92.1%	88.5%
2	22.327	**0.004**	3,981.0	90.9%	92.9%	91.9%	88.7%
3	17.805	**0.023**	4,028.9	90.5%	92.9%	92.0%	88.5%
4	19.600	**0.012**	4,018.5	90.4%	92.7%	92.1%	88.4%
5	23.104	**0.003**	4,056.8	90.3%	92.5%	91.9%	88.5%
6	8.970	0.345	4,093.8	90.4%	93.0%	91.9%	88.7%
7	19.643	**0.012**	3,963.3	90.7%	92.9%	91.9%	88.6%
8	35.392	**0.000**	3,846.2	90.9%	93.2%	92.1%	88.3%
9	23.340	**0.003**	4,103.4	90.2%	92.5%	91.9%	88.5%
10	12.198	0.143	3,957.6	90.7%	92.8%	92.1%	88.3%
Means	20.295			90.5%	92.8%	92.0%	88.5%

[a] Bolded significance levels are significant at $\alpha > 0.05$.

B. Estimated logistic regression model coefficients:

Number	CONSTANT	BASIZE	ELEV	UPGRD	DNGRD	PRECIP
1	-6.91996	3.52477	-0.10053	-0.02261	-0.05009	0.01726
2	-7.29955	3.73119	-0.10789		-0.07233	0.01768
3	-7.27174	3.68196	-0.10068		-0.07669	0.01785
4	-7.31098	3.73650	-0.10327		-0.07372	0.01680
5	-7.26114	3.72764	-0.09630		-0.07531	0.01619
6	-7.00478	3.58570	-0.10236		-0.07349	0.01699
7	-7.17530	3.69346	-0.10350		-0.06565	0.01590
8	-7.42720	3.83523	-0.10287	-0.01458	-0.06261	0.01639
9	-7.01478	3.63623	-0.09415		-0.07629	0.01535
10	-7.56838	3.79122	-0.10542		-0.06449	0.01793
Means	-7.22538	3.69439	-0.10170	-0.01860	-0.06907	0.01683
St. Dev.	0.20095	0.09284	0.00406	0.00568	0.00840	0.00087
Coef. Var.	2.8%	2.5%	4.0%	30.5%	12.2%	5.2%

Appendix Table 8e Summary of the results of estimating a logistic regression using 10 different model estimation data sets with 6,000 fish absent points and 6,000 fish present points.

A. Model fit and assessment statistics:

Model	Hosmer-Lemeshow		-2 x log	Self-Classification		Validation Data	
Number	Chi sq.	Signif.[a]	likelihood	Absent	Present	Absent	Present
1	26.678	**0.001**	4,951.9	90.2%	92.5%	92.1%	88.4%
2	20.667	**0.008**	4,818.2	90.4%	92.8%	91.7%	89.0%
3	21.252	**0.007**	5,022.8	90.1%	92.9%	92.0%	88.6%
4	16.009	**0.042**	4,930.7	90.3%	92.5%	92.0%	88.5%
5	16.144	**0.040**	4,798.6	90.5%	93.0%	91.9%	88.7%
6	16.894	**0.031**	4,638.2	90.9%	93.1%	92.0%	88.5%
7	16.378	**0.037**	4,940.8	90.4%	92.8%	92.0%	88.5%
8	37.511	**0.000**	5,011.6	89.8%	92.5%	92.0%	88.5%
9	16.415	**0.037**	4,848.3	90.9%	92.7%	91.9%	88.7%
10	23.004	**0.003**	4,826.9	90.2%	92.7%	92.1%	88.4%
Means	21.095			90.4%	92.8%	92.0%	88.6%

[a] Bolded significance levels are significant at $\alpha > 0.05$.

B. Estimated logistic regression model coefficients:

Number	CONSTANT	BASIZE	ELEV	UPGRD	DNGRD	PRECIP
1	-7.20264	3.63689	-0.10161		-0.07278	0.01763
2	-7.09268	3.69374	-0.10754		-0.06803	0.01616
3	-6.85523	3.56254	-0.09770	-0.01527	-0.06117	0.01572
4	-7.35315	3.69262	-0.10032		-0.07105	0.01797
5	-7.12475	3.70270	-0.10465		-0.07238	0.01599
6	-7.41567	3.81290	-0.10309	-0.01773	-0.05859	0.01698
7	-7.49091	3.71374	-0.10712		-0.06421	0.01907
8	-7.15015	3.58797	-0.09786	-0.01715	-0.05882	0.01842
9	-7.00607	3.63333	-0.09993		-0.07525	0.01591
10	-7.27509	3.75642	-0.10372		-0.07008	0.01572
Means	-7.19663	3.67929	-0.10235	-0.01672	-0.06724	0.01696
St. Dev.	0.19330	0.07587	0.00349	0.00129	0.00611	0.00124
Coef. Var.	2.7%	2.1%	3.4%	7.7%	9.1%	7.3%

Appendix Table 9 Final logistic regression model classification accuracies for model validation data, by WRIA, (summarized for fish absent [FA] and fish present [FP] points in fourth order sub-basins [FOSB] and for unassigned points).

	FOSB Data		Unassigned Data	
WRIA	**FA**	**FP**	**FA**	**FP**
1	90.4%	96.7%	97.5%	98.3%
3&4	99.0%		93.7%	99.7%
5	95.3%	88.4%	94.4%	98.5%
6	83.0%		65.6%	100.0%
7	94.2%	82.6%	87.0%	98.5%
8	82.7%	100.0%	84.2%	99.1%
9	95.7%	48.1%	98.5%	95.8%
10	94.7%	81.3%	99.7%	97.6%
12	100.0%	83.8%	100.0%	99.1%
13	96.0%	94.4%	99.1%	99.7%
14	96.5%	84.5%		100.0%
15	92.4%	86.3%	91.6%	99.3%
16	100.0%	100.0%	97.8%	97.2%
17	89.7%	92.7%	89.8%	99.6%
18	96.7%	100.0%	91.0%	99.9%
19	87.2%	100.0%	86.5%	99.1%
20	86.8%	88.1%	84.0%	97.9%
21	93.3%	87.0%	90.2%	91.7%
22	84.1%	90.7%	84.0%	97.9%
23	93.4%	86.5%	93.4%	97.4%
24	91.0%	91.4%	85.4%	98.0%
25	100.0%	97.0%		100.0%
26	90.2%	92.1%	92.7%	99.1%
27	92.1%	100.0%	86.8%	99.9%
28			92.5%	100.0%
29	91.0%	94.9%	92.8%	100.0%
Mean	92.6%	89.8%	90.8%	98.6%
Median	93.3%	91.4%	92.1%	99.1%
Minimum	82.7%	48.1%	65.6%	91.7%
Maximum	100.0%	100.0%	100.0%	100.0%

Appendix Table 10 Number of points used to estimate classification accuracies, by WRIA, for the final model validation.

	FOSB Data		Unassigned Data	
WRIA	FA	FP	FA	FP
1	1,890	513	1,671	16,609
3&4	9,585	0	18,235	26,589
5	1,535	896	5,901	23,653
6	2,199	0	482	320
7	4,718	4,072	10,520	38,831
8	965	223	1,988	5,068
9	2,199	237	5,632	19,371
10	5,563	2,511	7,945	31,454
12	843	647	2,178	17,677
13	1,169	284	341	7,887
14	650	737	0	2,475
15	1,703	1,761	1,005	8,150
16	112	387	548	1,656
17	3,761	521	2,251	5,260
18	17,853	139	25,328	8,296
19	156	172	1,060	7,168
20	16,099	13,035	8,620	30,023
21	12,017	13,765	10,454	23,475
22	13,882	16,751	9,931	81,070
23	46,391	23,943	26,460	76,056
24	16,261	15,082	9,822	56,039
25	262	134	0	7,863
26	17,665	6,479	11,265	58,521
27	5,831	608	5,303	25,341
28	0	0	295	1,527
29	3,428	589	851	3,604

PAGE LEFT BLANK INTENTIONALLY

APPENDIX A

Unresolved and Conflicting End-of-Fish Points

This appendix discusses unresolved and conflicting end-of-fish points (EOFPs) and summarizes the analyses that were conducted related to these points.

Definition of Unresolved and Conflicting EOFPs

Unresolved end-of-fish points were single EOFPs whose recorded location could not be associated with a stream on the 10m DEM network during processing by EarthRes.i. These EOFPs were not included in the process used to generate Fish Absent | Fish Present (FAFP) points or included in the logistic regression analysis. There were 48 unresolved, single EOFPs. Appendix Table A1 summarizes the number of unresolved, single EOFPs present in each WRIA.

In the data set compiled by EarthRes.i there were occasionally end-of-fish points that had been submitted which conflicted with each other, i.e., there were two or more EOFPs specified on the same stream. We wanted to resolve as many of these conflicts as possible so that data from the affected streams could be included in the analyses, if possible, and additional upstream and downstream FAFP data generated. The two most common reasons for the conflicting EOFPs were:

- a survey agency had identified two types of EOFPs for the stream during a survey (e.g., a Last Fish EOFP and a Last Fish Habitat EOFP), or
- two or more distinct surveys had been conducted on the stream.

The majority of the conflicting EOFPs involved a conflicting pair of points. We developed an initial protocol (described below) to examine these pairs of conflicting points and allow some of them to be used in the process to generate FAFP points for use in the logistic regression model building process. In cases when there were more than two EOFPs in conflict, we felt it was not possible to develop an easily applied protocol to resolve the conflict. Therefore, when there were more than two EOFPs in conflict these potential EOFPs were excluded from all analyses. A total of 42 potential EOFPs were omitted from consideration due to multiple (>2) conflicting EOFPs on a stream (Appendix Table A1).

An additional benefit of the two-conflicting points analysis was that it allowed us to gain some understanding of the differences between EOFP types (for example, LF and LFH) when they are located on the same stream and the variability associated with identifying an EOFP on a stream during independent surveys. There were 268 conflicting pairs of EOFPs in the final data set submitted to EarthRes.i (Appendix Table A1).

Appendix Table A1 Summary of the number of single, unresolved and multiple, conflicting end-of-fish points (EOFPs) present in the original data set, by WRIA.

WRIA	Unresolved Single Points	Two-point Conflicts BSD[a] > 50 acres	Two-point Conflicts BSD > 50 acres	Multiple (>2) Point Conflicts[b]
1	0	7	0	0
2	0	0	0	0
3 and 4	0	18	2	2
5	0	7	1	1
6	0	0	0	0
7	18	38	5	4
8	0	0	0	0
9	0	2	1	0
10	1	17	2	4
11	6	13	3	3
12	0	0	0	0
13	0	4	3	4
14	0	0	3	0
15	3	8	2	1
16	0	2	1	0
17	0	3	0	1
18	0	15	3	2
19	1	2	0	0
20	2	10	0	4
21	2	14	1	3
22	10	15	4	0
23	1	16	5	5
24	3	16	1	5
25	0	1	0	0
26	0	2	2	1
27	0	4	1	0
28	0	2	1	0
29	1	8	3	2
Totals	48	224	44	42

[a] EOFP pairs where the basin size difference (BSD) between the two conflicting EOFPs was > 50 acres.

[b] Number of potential EOFPs that could be defined by the conflicting points (anywhere from three to eight points were in conflict in each case).

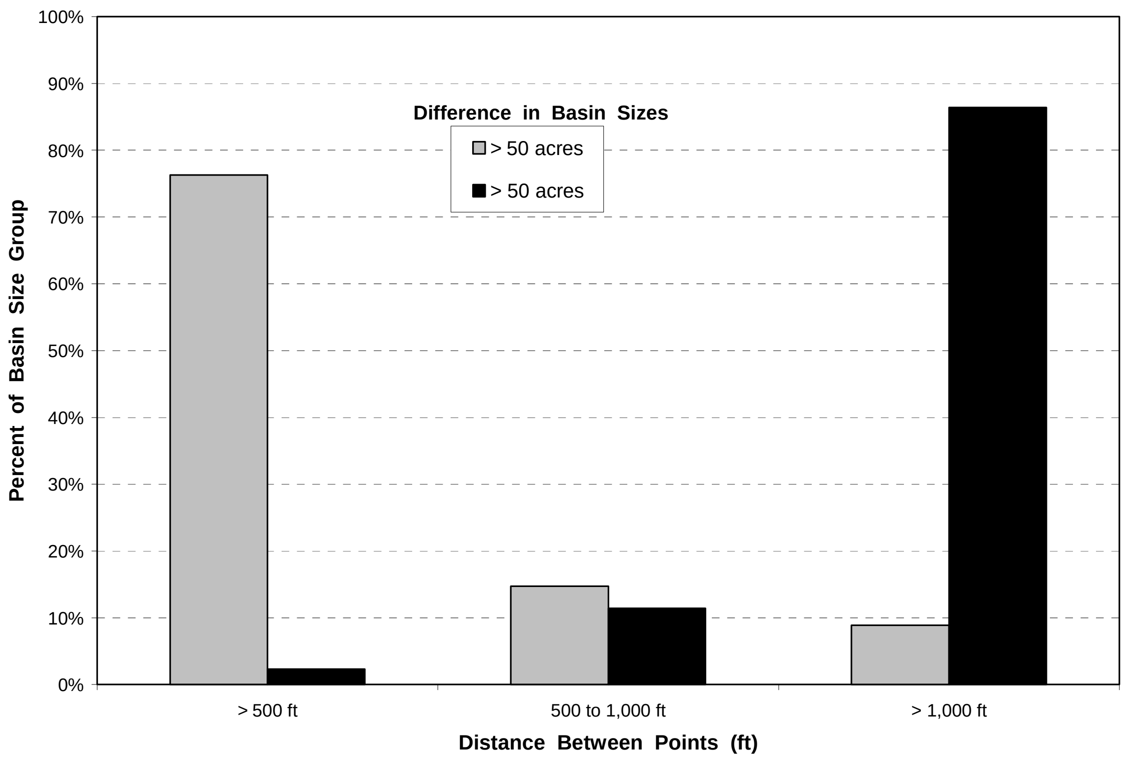
Initial Screening Protocol for Conflicting Pairs:

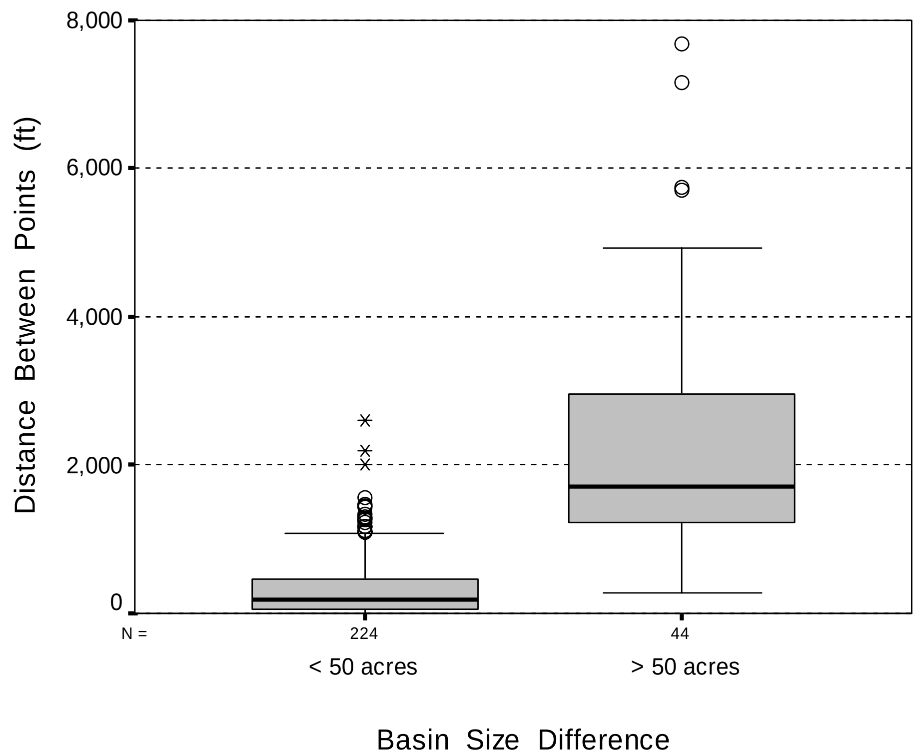
We conducted our initial analysis with a preliminary subset of the conflicting EOFP pair data using methods similar to those reported below. We had previously identified basin size (BASIZE) as the most important physical attribute used by the logistic regression models to distinguish Fish Absent points from Fish Present points. In our initial analysis we found that the majority of the BASIZE differences between conflicting pairs of EOFPs were less than 50 acres. The distance on the stream network between the conflicting EOFPs was generally much smaller when the BASIZE difference was 50 acres or smaller. For the final set of 268 conflicting EOFP pairs, the majority (76%) of the distances between the conflicting EOFPs were 500 ft or less when the BASIZE difference was less than 50 acres. In contrast, 86% of the distances between the conflicting EOFPs were greater than 1,000 ft when the BASIZE differences were greater than 50 acres (Appendix Figure A1).

When the differences in BASIZE between the conflicting EOFP pairs exceeded 50 acres there were some very large distances between the points on the stream network. The average distance between these points was almost one-half mile and the median distance was one-third of a mile (Appendix Table A2). Because of these large distances between points, and the potential influence of these points on the logistic regression model, we decided that conflicting pairs would be considered for use in the model only when the difference in basin size between the conflicting pair of points was 50 acres or less.

Appendix Table A2 Summary statistics for the distances between conflicting EOFP pairs on the stream network for pairs with basin size differences > 50 acres compared to pairs with basin size differences > 50 acres.

Summary Statistic	Difference in Basin Size between Conflicting EOFP Points	
	> 50 acres	> 50 acres
Sample Size	224	44
Mean (ft)	335.1	2,381.5
Standard Error	28.2	260.7
Median (ft)	191.4	1,709.9
Minimum (ft)	33.8	278.9
Maximum (ft)	2,600.5	7,686.2





Appendix Figure A1 Bar chart (top) comparing the distribution of the distances between conflicting EOFPs for the two basin size difference groups and box-and-whiskers plot (bottom) summarizing the distances for the groups.

Based upon the initial analysis, the following protocol was developed for processing conflicting EOFP pairs:

- if the difference in basin size between the pair was greater than 50 acres the pair was omitted from consideration,
- if the difference in basin size between the pair was > 50 acres one EOFP from the pair was randomly[21] selected and used to generate FAFP points on the DEM network.

The following sections describe the results of the analyses of the differences between EOFPs for the 224 conflicting EOFP pairs that were resolved using the above protocol and used to generate FAFP data. The analyses focus on the differences between EOFP pairs for groups of the data which provide information on differences between EOFP point types and about survey variability.

<u>Comparison of Conflicting EOFP Pairs with LFH and LF Point Types</u>

There were 41 cases where the conflicting pair of EOFPs consisted of a Last Fish Habitat (LFH) and a Last Fish (LF) point. In 39 of the 41 cases the EOFP was identified by the same survey agency during a single survey (i.e., the survey date for the two points was the same). In one case the same survey agency identified the EOFPs on different survey dates, and in one case different agencies identified the EOFPs on different survey dates (Appendix Table A3).

[21] The exception to the random selection occurred when one of the pair did not have any of the survey identifying information associated with it (items g, h, i, j, k, l, and m from Table 2). In that situation the EOFP with the survey identifying information specified was always selected.

Appendix Table A3 Summary of the sponsoring agencies for each conflicting EOFP pair with Last Fish Habitat and Last Fish EOFP types.

Number of Conflicting Pairs	EOFP Point Type	
	Last Fish Habitat	Last Fish
17	Quinault DNR	Quinault DNR
19	Washington Trout[a]	Washington Trout
4	Aquatic Tech	Aquatic Tech
1	Quinault DNR	Weyerhaeuser

[a] None of the Washington Trout data were collected using the " PRE " survey protocol.

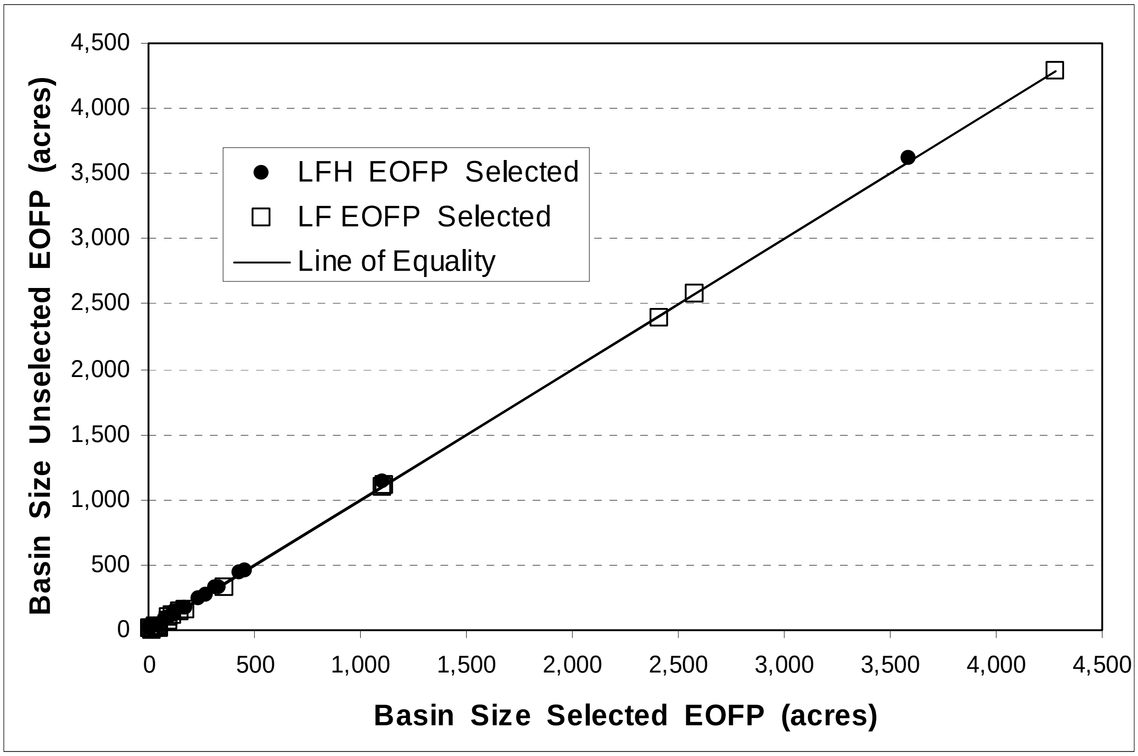
Summary statistics for the differences between the conflicting EOFP pairs in distance on the stream network and for the five GIS-derived physical attributes associated with each point are presented in Appendix Table A4. The mean DEM network distance between LFH and LF EOFP types that were generally identified by the same survey agency during the same survey was about 309 ft (median distance = 163 ft). The logistic regression (LR) procedure does not use location on the stream network. The LR model coefficients are based upon the five GIS-derived physical attributes associated with each EOFP. Therefore, the differences between these attributes give a better indication of the possible impact on the LR model from using a LFH EOFP as opposed to a LF EOFP to generate FAFP data.

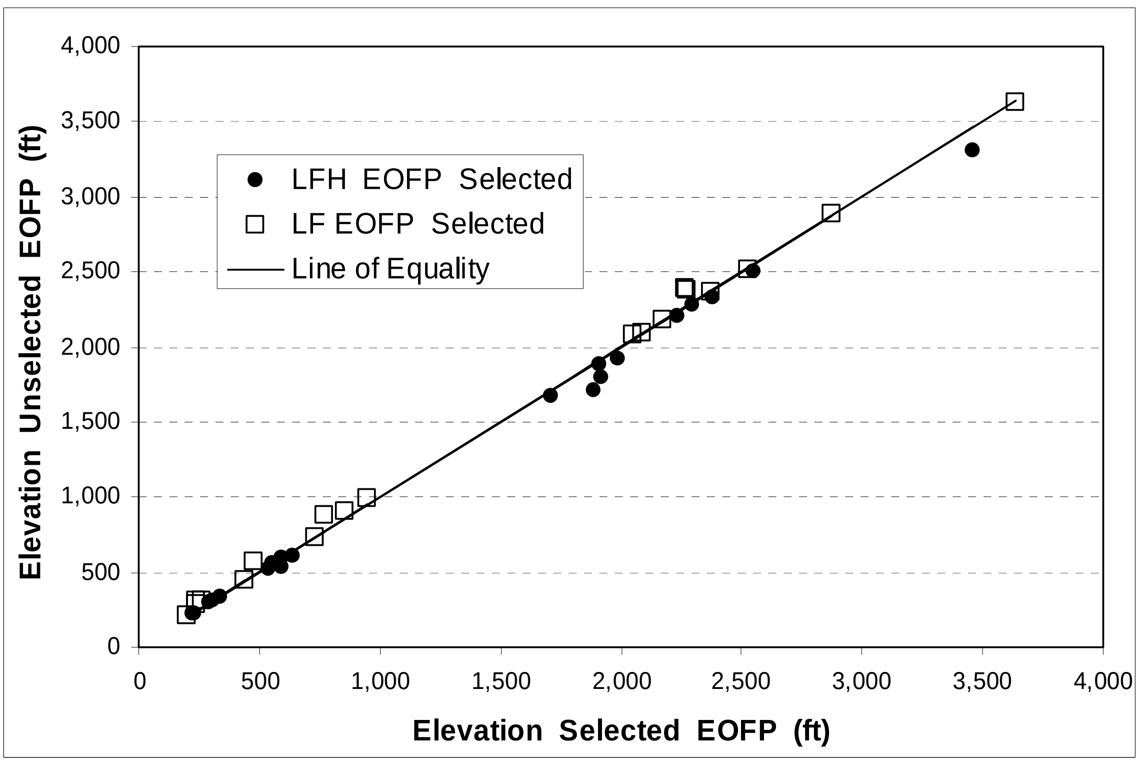
Appendix Table A4 Summary of differences in distance on the stream network and GIS-derived physical attributes between Last Fish Habitat and Last Fish surveys for conflicting EOFP pairs.

Statistic	Distance (ft)	Basin Size (acres)	Elevation (ft)	Upstream Gradient	Downstream Gradient	Precip. (inches)
Mean	309.1	9.2	40.4	5.9	4.3	0.09
Standard Error	56.4	1.8	7.2	1.2	1.0	0.03
Median	163.4	2.9	24.5	1.6	1.3	0.00
Minimum	33.8	0.02	0.2	0.1	0.1	0.00
Maximum	1,307.2	41.9	183.0	28.3	28.5	0.84

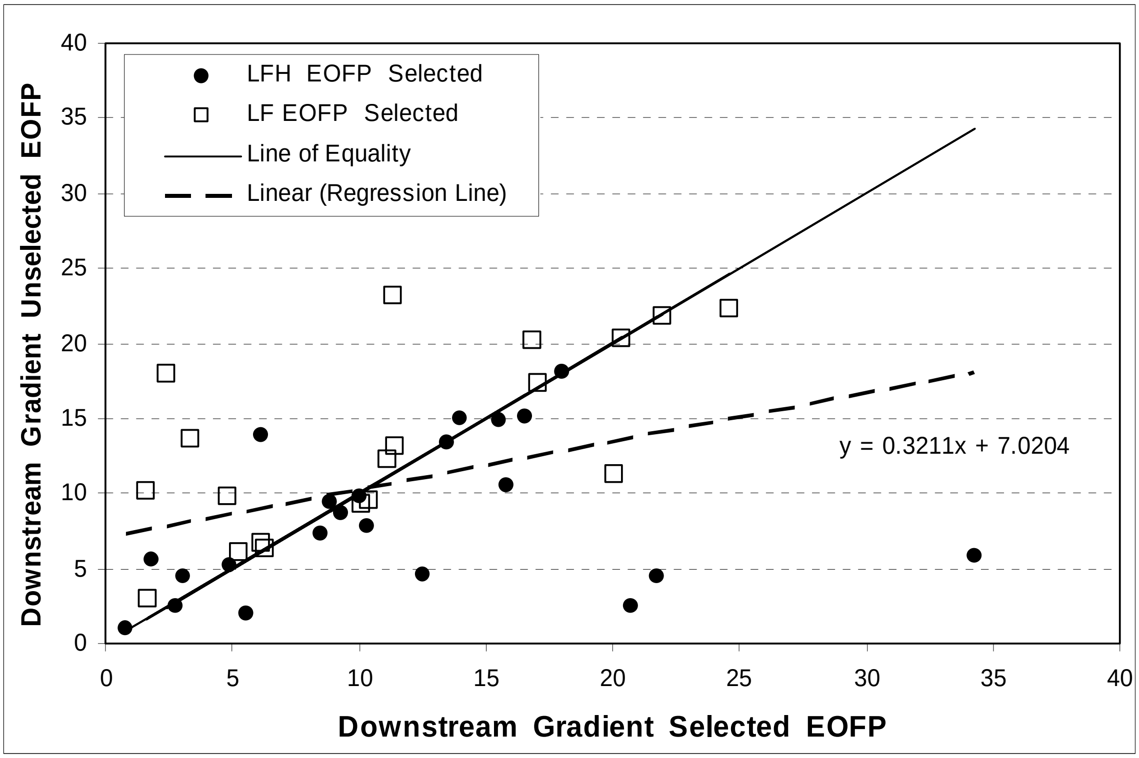
The mean differences between LFH and LF points were generally small. We visually compared the attribute values for these two point types using scatter plots with the value for the randomly selected EOFP on the x-axis and the value for the EOFP that was not selected on the y-axis. If the two values are similar they should cluster around a line with a slope of one (the line of equality). When regressed against each other, the slope of the line should not be significantly different from one (1.0). Appendix Figure A2 shows this relationship for BASIZE, ELEV, DNGRD, and PRECIP. UPGRD is not shown because it is similar to DNGRD and the upstream gradient attribute was rarely selected as a parameter in the LR models developed for this project. The slopes of the regression lines are not significantly different than one for BASIZE, ELEV, and PRECIP. For these three attributes, the y-intercept was not significant and was not included in the regression; the slope coefficient was very precisely estimated for all three of these attributes (coefficient of variation for the estimated slope less than 1%). The only departure from the line of equality is for the downstream gradient attribute. The sensitivity of the estimates of the probability of fish presence produced by the final logistic regression model to differences in the physical attributes within the range reported here are examined in Appendix B.

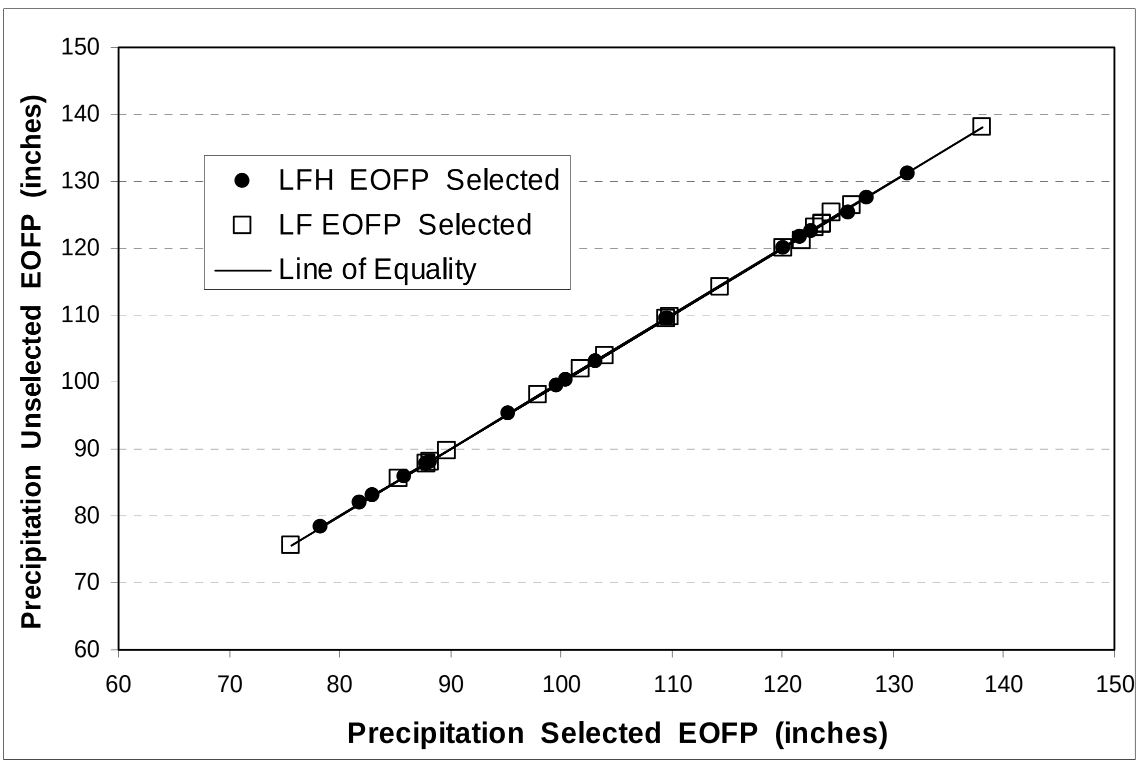
It is commonly assumed that a LFH EOFP will be located upstream of a LF EOFP when both are placed on the same stream. For 16 of the 41 cases (or 39% of the LFH|LF EOFP conflicts), however, the distance between the two EOFPs was less than 50 ft; this represents a difference in the placement of the points on the 10-m DEM network of essentially one 10-m cell. The downstream EOFP was selected for 20 of the 41 pairs when randomly selecting an EOFP from the conflicting LFH|LF EOFP pair. For 19 of the 41 conflicting LFH|LF EOFP pairs, the LF EOFP was selected.





Appendix Figure A2 Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with LFH and LF point types. Line of equality shown for reference.





Appendix Figure A2 (continued) Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with LFH and LF point types. Line of equality shown for reference.

Comparison of Conflicting EOFP Pairs with LFH and LS Point Types

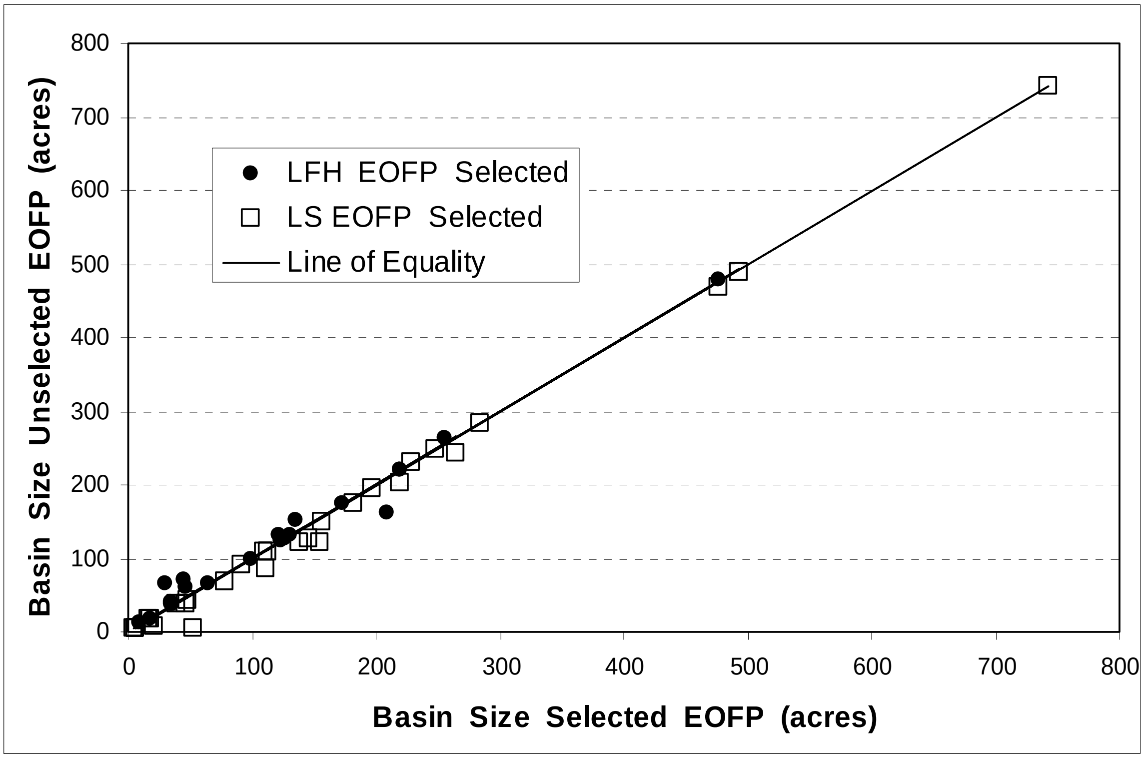
There were 50 cases where the conflicting pair of EOFPs consisted of a LFH and a Last Salmonid (LS) point. In all 50 cases the EOFP was identified by the same survey agency during a single survey (i.e., the survey date for the two points was the same). The surveying agency for all these conflicting EOFP pairs was Washington Trout; 21 of the surveys were conducted using the LS survey protocol and 29 were collected using the survey protocol that existed prior to the establishment of the emergency stream typing rules in 1996 (labeled the "PRE" protocol). Although we excluded EOFPs collected under the "PRE" protocol from the process used to generate FAFP points (see main report), they were included for this analysis. Regardless of the protocol, we felt that these points would supply valuable information on between EOFP type differences.

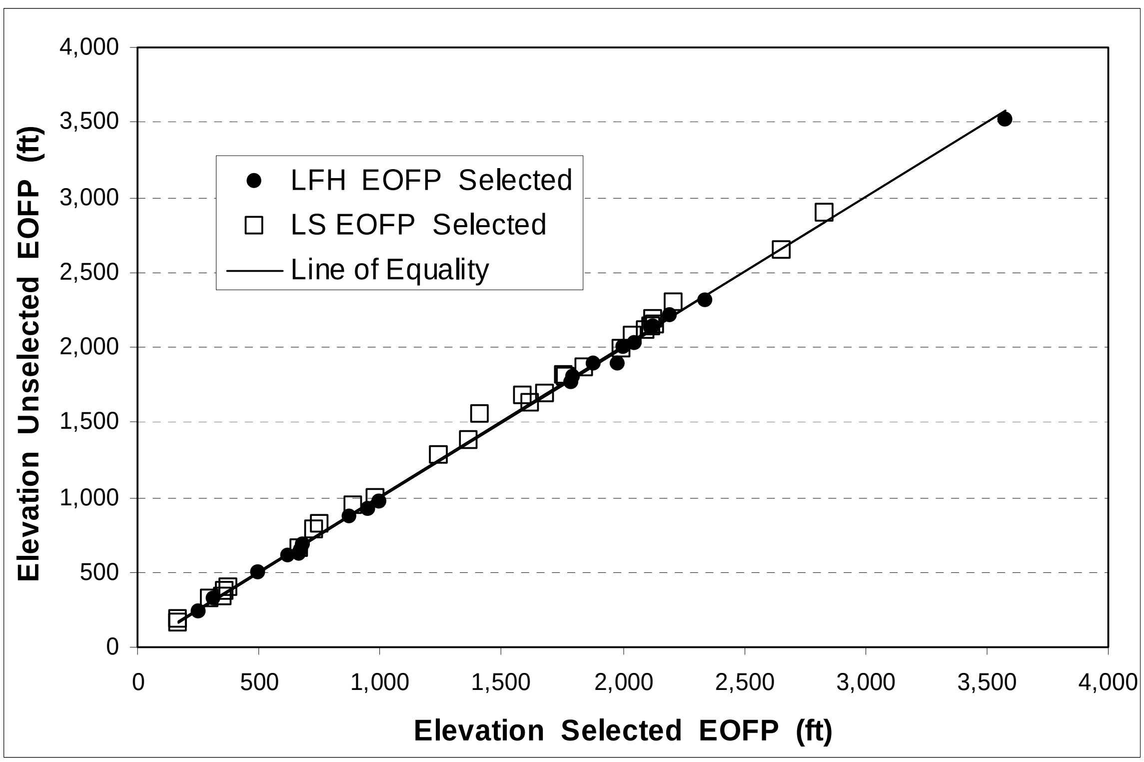
Summary statistics for the differences between the conflicting EOFP pairs in distance on the stream network and for the five GIS-derived physical attributes associated with each point are presented in Appendix Table A5. The mean DEM network distance between LFH and LS EOFP types that were identified by the same survey agency during the same survey was about 327 ft (median distance = 238 ft).

Appendix Table A5 Summary of differences in distance on the stream network and GIS-derived physical attributes between Last Fish Habitat and Last Salmonid surveys for conflicting EOFP pairs.

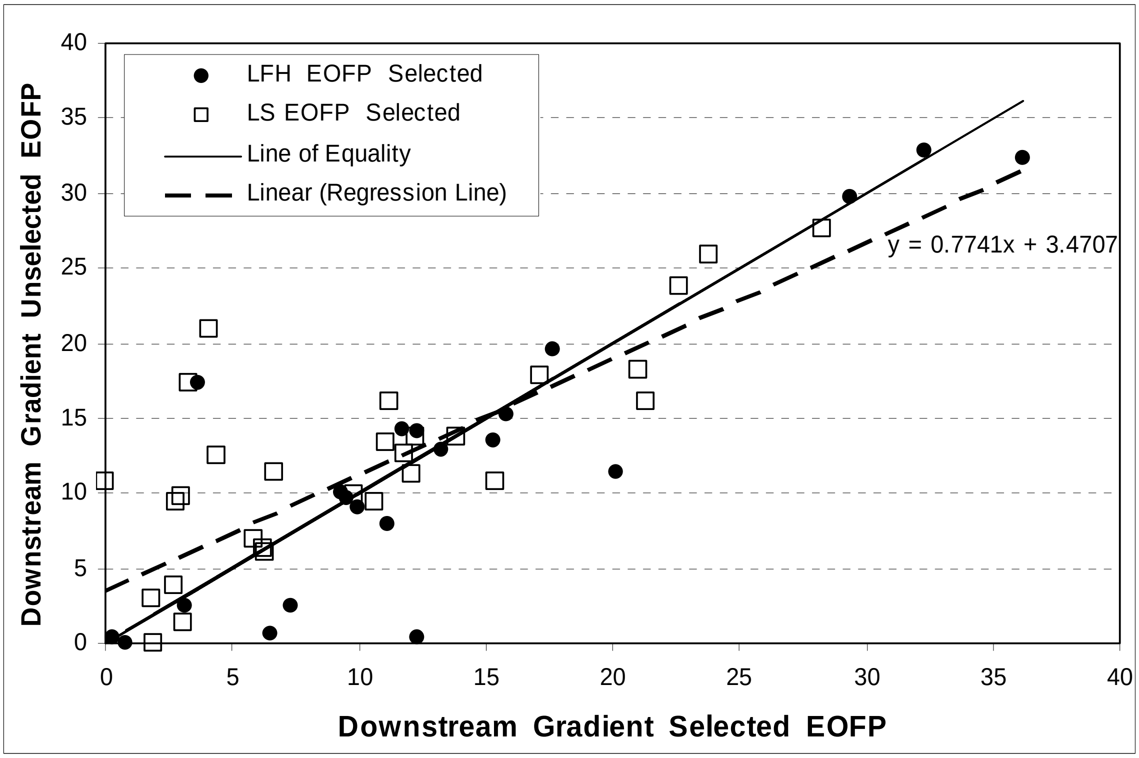
Statistic	Distance (ft)	Basin Size (acres)	Elevation (ft)	Upstream Gradient	Downstream Gradient	Precip. (inches)
Mean	326.5	8.6	29.1	4.7	3.4	0.07
Standard Error	50.2	1.7	4.1	0.8	0.6	0.04
Median	238.1	2.4	17.0	2.3	1.7	0.00
Minimum	33.8	0.03	2.0	0.2	0.04	0.00
Maximum	1,554.7	47.1	139.7	26.6	16.9	1.54

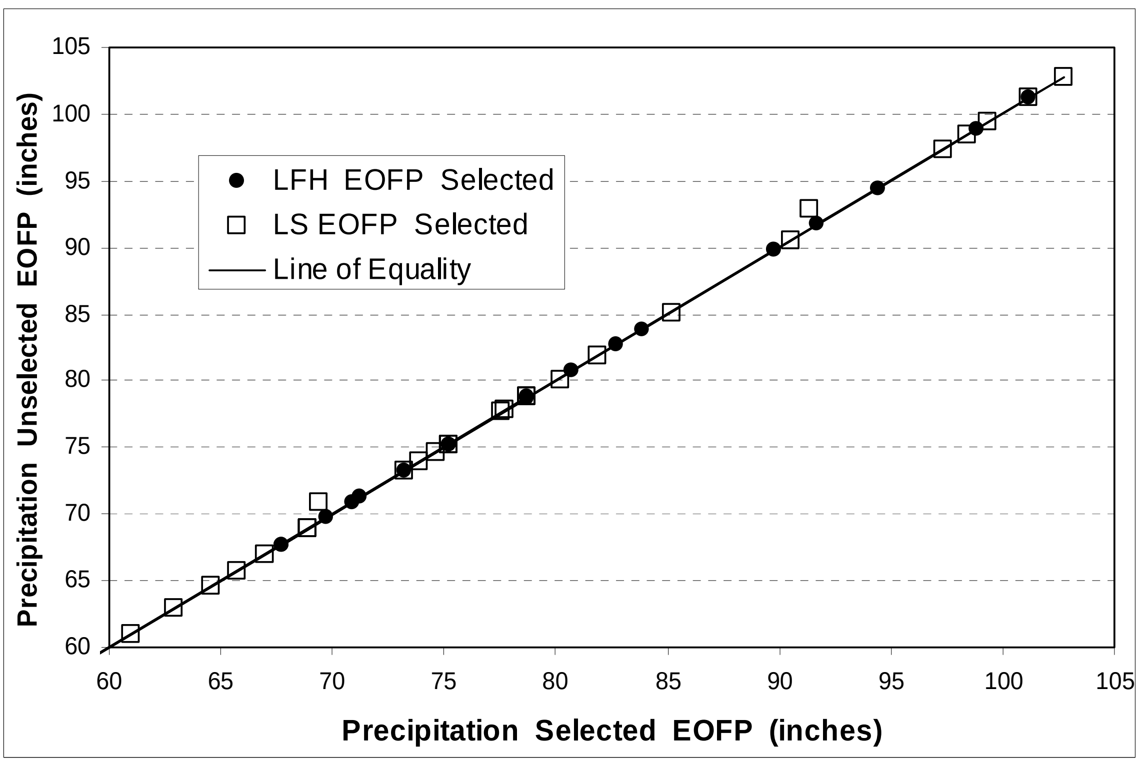
The mean differences between LFH and LS points were generally small. Although the mean distance between conflicting EOFPs was slightly larger than for the LFH|LF comparison, the mean differences for the five GIS-derived physical attributes were all smaller for the LFH|LS pairs compared to the LFH|LF EOFP pairs. Appendix Figure A3 shows the relationship of the LFH|LS data for BASIZE, ELEV, DNGRD, and PRECIP relative to the line of equality. The slopes of the regression lines are not significantly different than one (1.0) for BASIZE, ELEV, and PRECIP. For these three attributes, the y-intercept was not significant and was not included in the regression; the slope coefficient was very precisely estimated for all three of these attributes (coefficient of variation for the estimated slope less than 1%). Similarly to the LFH|LF analysis, the only departure from the line of equality was for the downstream gradient attribute.





Appendix Figure A3 Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with LFH and LS point types. Line of equality shown for reference.





Appendix Figure A3 (continued) Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with LFH and LS point types. Line of equality shown for reference.

It is commonly assumed that a LFH EOFP will be located upstream of a LS EOFP when both are placed on the same stream. For 17 of the 50 cases (or 34% of the LFH|LS EOFP conflicts), however, the distance between the two EOFPs was less than 50 ft; this represents a difference in the placement of the points on the 10-m DEM network of essentially one 10-m cell. The downstream EOFP was selected for 30 of the 50 pairs when randomly selecting an EOFP from the conflicting LFH|LF EOFP pair. For 29 of the 50 conflicting LFH|LS EOFP pairs, the LS EOFP was selected.

Comparison of Conflicting EOFP Pairs with Identical Point Types

There were 14 cases where the conflicting pair of EOFPs had the same point type but had been collected by different agencies and/or on different days. For six cases the conflicting pair had a LSH point type, for six cases the conflicting pair had a LFH point type, and the remaining two pairs were both LF point types. For seven of the conflicting pairs the surveys were conducted on the same day and for seven of the pairs the surveys were conducted on different days.

Appendix Table A6 Summary of the sponsoring agencies, survey date, and EOFP types for conflicting EOFP pairs with identical types.

Number of Pairs	Survey Agency #1	Survey Agency #2	Survey Date	Point Type
6	Wa. DNR	Tulalip DNR	Same	LSH
6	Quinault DNR	Hoh Fisheries	Different	LFH
1	Quinault DNR	Quinault DNR	Different	LF
1	Weyerhaeuser	Weyerhaeuser	Same	LF

Summary statistics for the differences between the conflicting EOFP pairs in distance on the stream network and for the five GIS-derived physical attributes associated with each point are presented in Appendix Table A7. The mean DEM network distance between conflicting EOFP pairs with the same point type was about 151 ft (median distance = 41 ft).

Appendix Table A7 Summary of differences in distance on the stream network and GIS-derived physical attributes between conflicting EOFP pairs with the same point type.

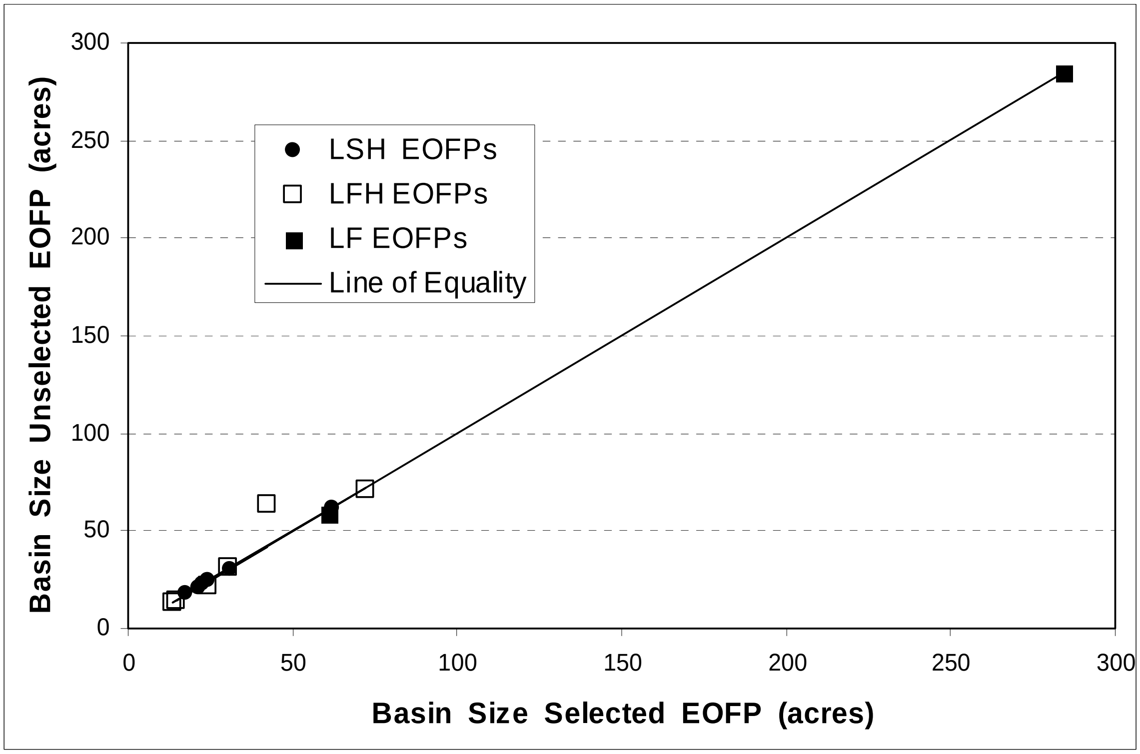
Statistic	Distance (ft)	Basin Size (acres)	Elevation (ft)	Upstream Gradient	Downstream Gradient	Precip. (inches)
Mean	150.8	2.0	22.2	1.7	1.4	0.01
Standard Error	88.0	1.5	13.0	0.5	0.4	0.01
Median	40.9	0.3	8.4	0.9	1.2	0.00
Minimum	3.8	0.03	3.3	0.2	0.12	0.00
Maximum	1,271.0	21.4	188.9	6.8	5.5	0.11

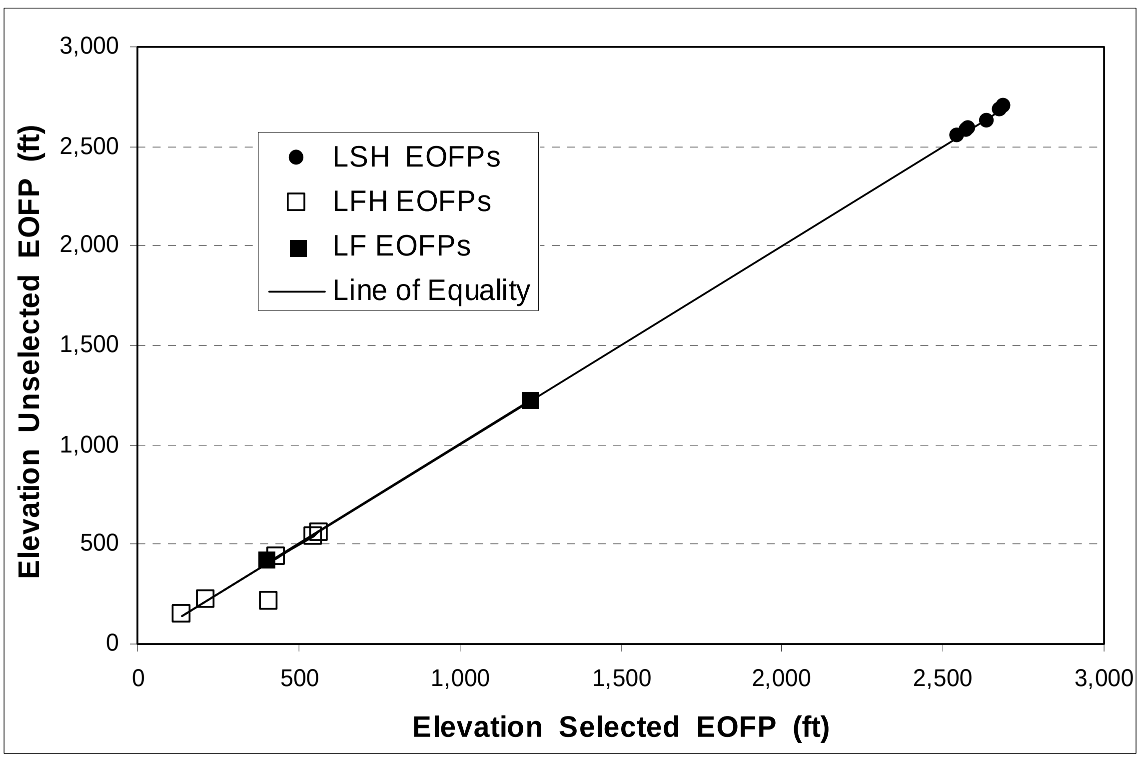
The mean differences between the conflicting EOFP pairs with the same point types were generally very small. Not surprisingly, given that the point types for these conflicting EOFP pairs were the same, the mean differences between the points in the pair were the smallest of the differences examined so far. Appendix Figure A4 shows the relationship of these data for BASIZE, ELEV, DNGRD, and PRECIP relative to the line of equality. The slopes of the regression lines were not significantly different than one (1.0) for all four physical attributes. For these attributes, the y-intercept was not significant and was not included in the regression; the slope coefficient was very precisely estimated for all four of these attributes (coefficient of variation for the estimated slope less than 4%).

For 11 of the 14 cases (or 79% of the identical point type EOFP conflicts), the distance between the two EOFPs was less than 50 ft; this represents a difference in the placement of the points on the 10-m DEM network of essentially one 10-m cell. The downstream EOFP was selected for 10 of the 14 pairs when randomly selecting an EOFP from the conflicting pair.

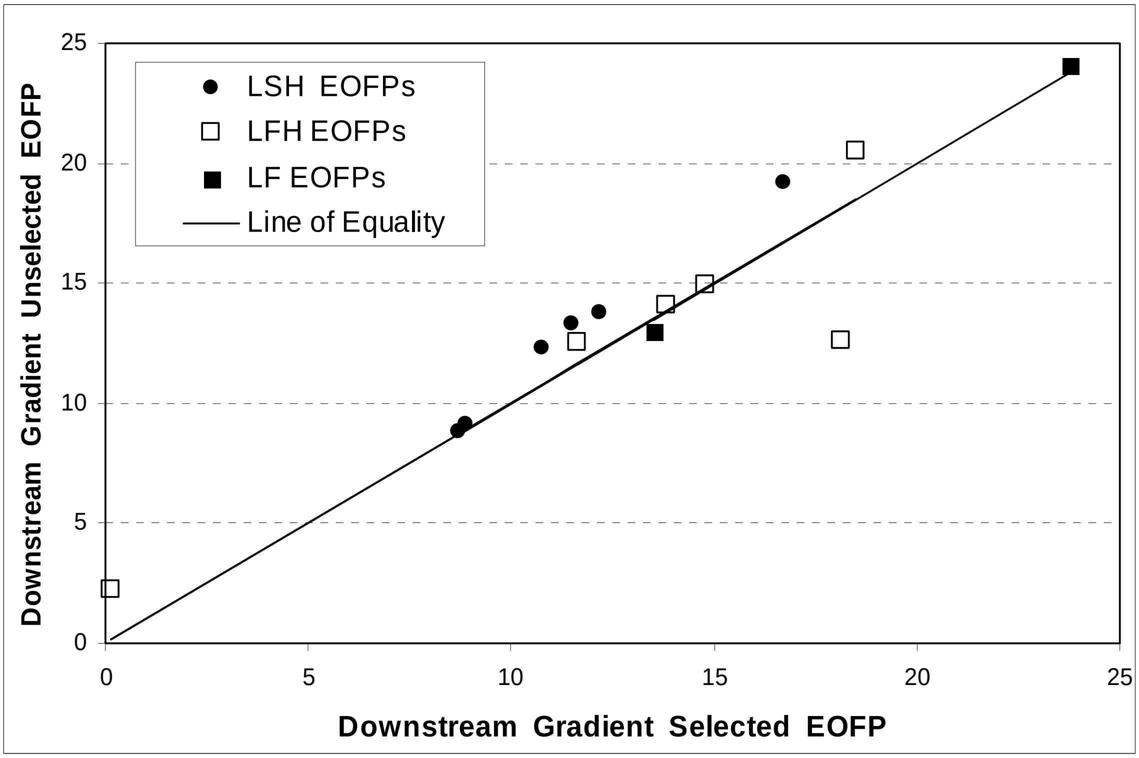
Comparison of Conflicting EOFP Pairs Where Information is Missing from One Survey

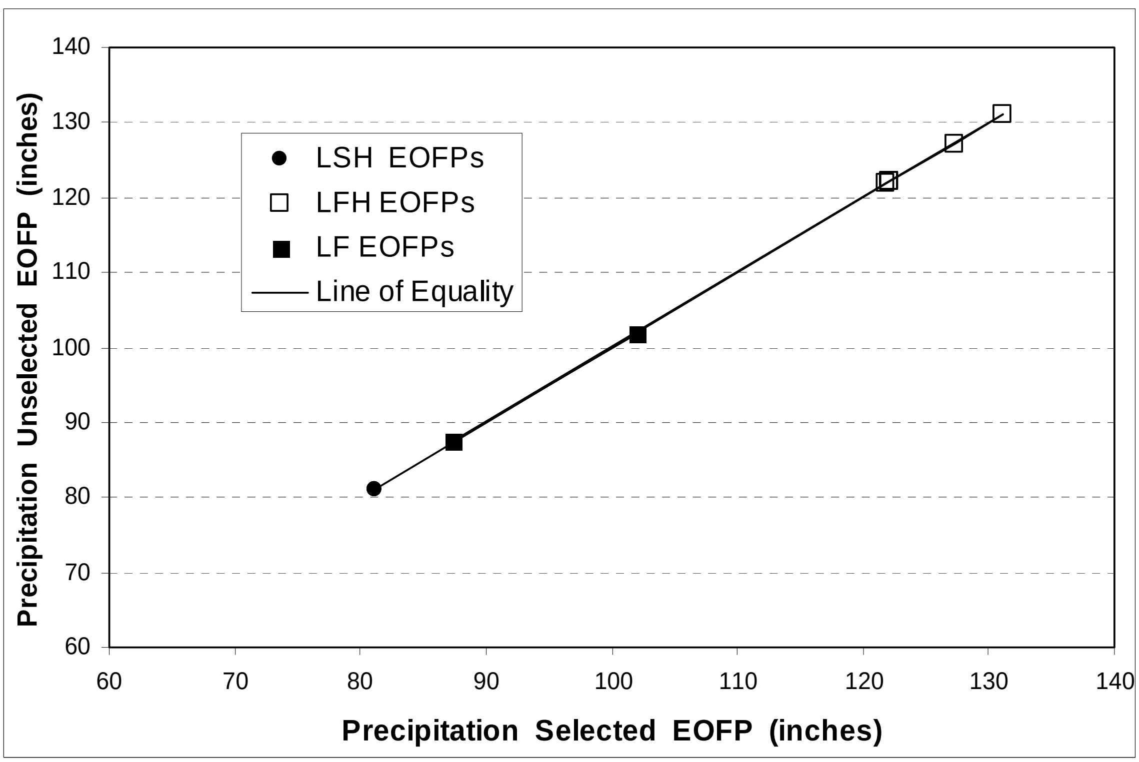
There were 112 cases where there was no identifying information associated with one of the EOFPs in the conflicting pair. For these cases, the EOFP with the identifying information was always used in the process to generate FAFP points. Ten different agencies conducted the surveys for the points which had the identifying information. The majority of the points were collected by Washington Trout (51 points or 46% of the conflicting pairs) and the Quinault Dept. of Natural Resources (31 points or 28%). Appendix Table A8 summarizes the survey agency and survey protocol used for these 112 pairs of conflicting EOFPs. Fifteen of these conflicting EOFP pairs were collected by Washington Trout using the "PRE" protocol. Although we excluded EOFPs collected under the "PRE" protocol from the process used to generate FAFP points (see main report), they were included for this analysis because we felt that these points would supply valuable information on between-EOFP differences. The majority of the points for the EOFPs with information were LFH point types (64 or 57%); there were 30 LF point types (27%), 15 LS point types (13%) and three LSH point types (Appendix Table A9).





Appendix Figure A4 Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with the same point types. Line of equality shown for reference.





Appendix Figure A4 (continued) Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs with the same point types. Line of equality shown for reference.

Appendix Table A8 Summary of the sponsoring agencies and survey protocol for conflicting EOFP pairs where one of the pair has no identifying information.

Survey Agency	Survey Protocol Used					Total
	LF	LFH	LS	LSH	PRE	
Aquatic Technical	2	1				3
Crown Pacific		1				1
Wa. Dept. Fish and Wildlife			1			1
Wa. Dept. Natural Resources		6		1		7
Hoh Fisheries Dept.		6				6
Olympic Environmental	1					1
Olympic Resources		8				8
Quinault Dept. Natural Res.	1	30				31
Washington Trout		15	21		15	51
Weyerhaeuser	3					3
Total	7	67	22	1	15	112

Appendix Table A9 Summary of the sponsoring agencies and end-of-fish point type for conflicting EOFP pairs where one of the pair has no identifying information.

Survey Agency	End-of-Fish Point Type				Total
	LF	LFH	LS	LSH	
Aquatic Technical	2	1			3
Crown Pacific		1			1
Wa. Dept. Fish and Wildlife				1	1
Wa. Dept. Natural Resources	2	4		1	7
Hoh Fisheries Dept.		6			6
Olympic Environmental	1				1
Olympic Resources	6	2			8
Quinault Dept. Natural Res.	5	23	3		31
Washington Trout	11	27	12	1	51
Weyerhaeuser	3				3
Total	30	64	15	3	112

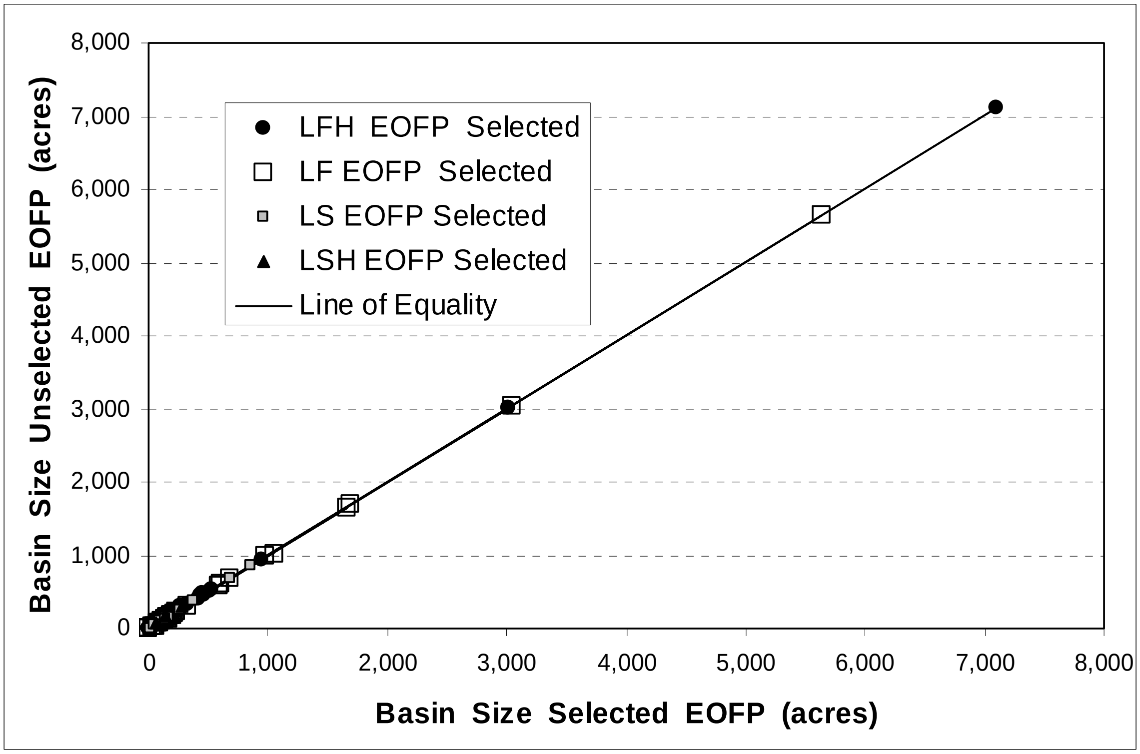
Summary statistics for the differences between the conflicting EOFP pairs in distance on the stream network and for the five GIS-derived physical attributes associated with each point are presented in Appendix Table A10. The mean DEM network distance between conflicting EOFP pairs where one of the pair has no identifying information was about 343 ft (median distance = 214 ft).

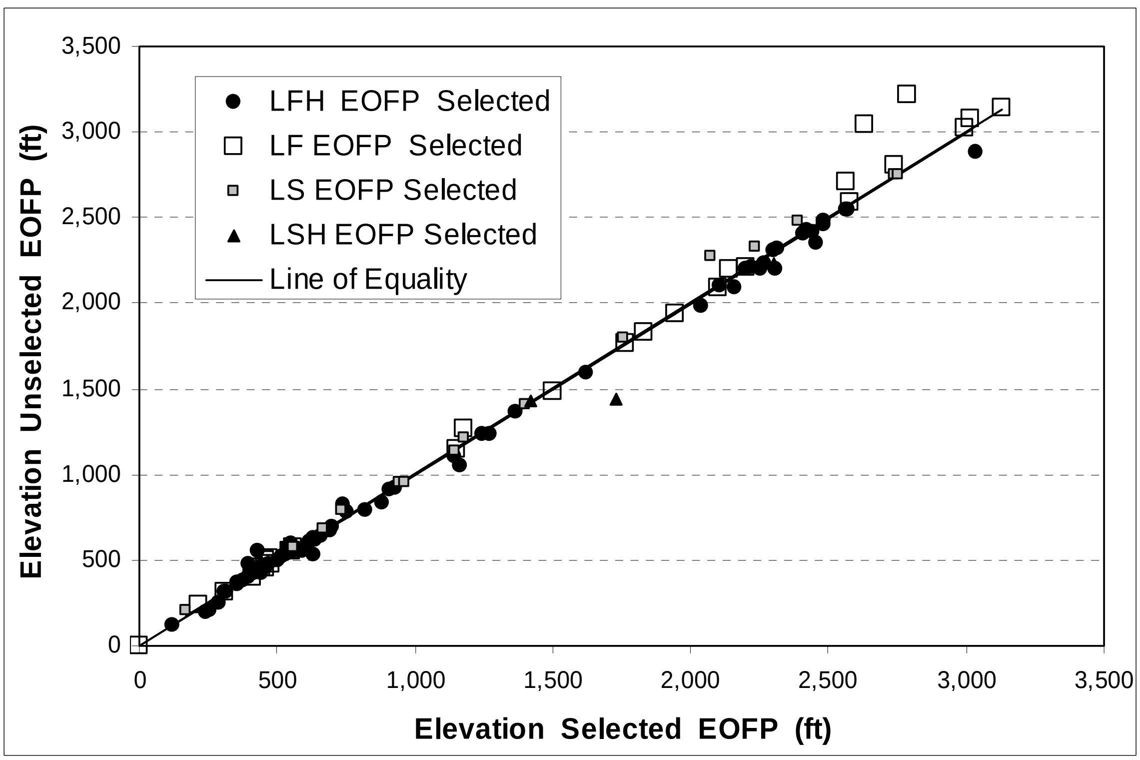
Appendix Table A10 Summary of differences in distance on the stream network and GIS-derived physical attributes between conflicting EOFP pairs where one of the pair has no identifying information.

Statistic	Distance (ft)	Basin Size (acres)	Elevation (ft)	Upstream Gradient	Downstream Gradient	Precip. (inches)
Mean	343.0	8.3	38.8	3.6	3.3	0.04
Standard Error	40.8	1.1	6.5	0.5	0.5	0.01
Median	214.2	2.7	13.0	1.9	1.4	0.00
Minimum	33.8	0.03	0.0	0.0	0.0	0.00
Maximum	2,201.1	46.7	433.0	29.5	30.0	0.80

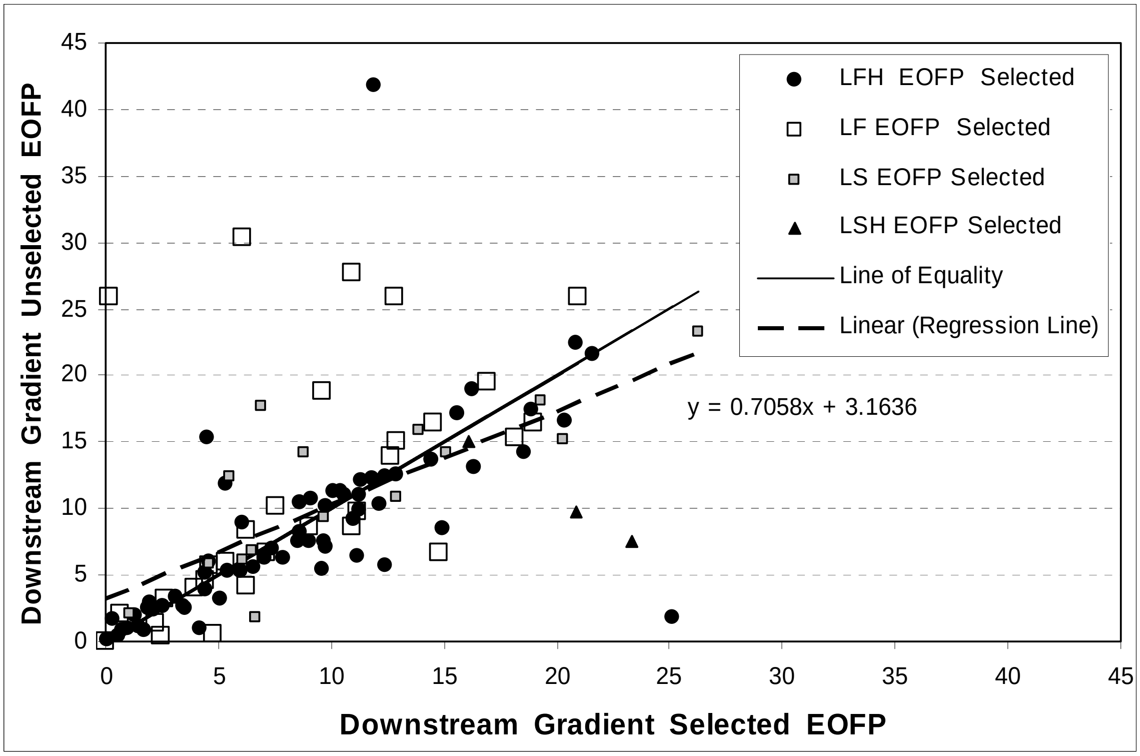
The mean differences between the conflicting EOFP pairs were generally small. The mean differences between the points in the pair were generally similar to those observed for the LFH|LF and LFH|LS conflicting pairs. Appendix Figure A5 shows the relationship of these data for BASIZE, ELEV, DNGRD, and PRECIP relative to the line of equality. The slopes of the regression lines were not significantly different than one (1.0) for BASIZE, ELEV, and PRECIP. For these three attributes, the y-intercept was not significant and was not included in the regression; the slope coefficient was very precisely estimated for all three of these attributes (coefficient of variation for the estimated slope less than 1%). Similarly to the LFH|LF and LFH|LS analyses, the only departure from the line of equality is for the downstream gradient attribute.

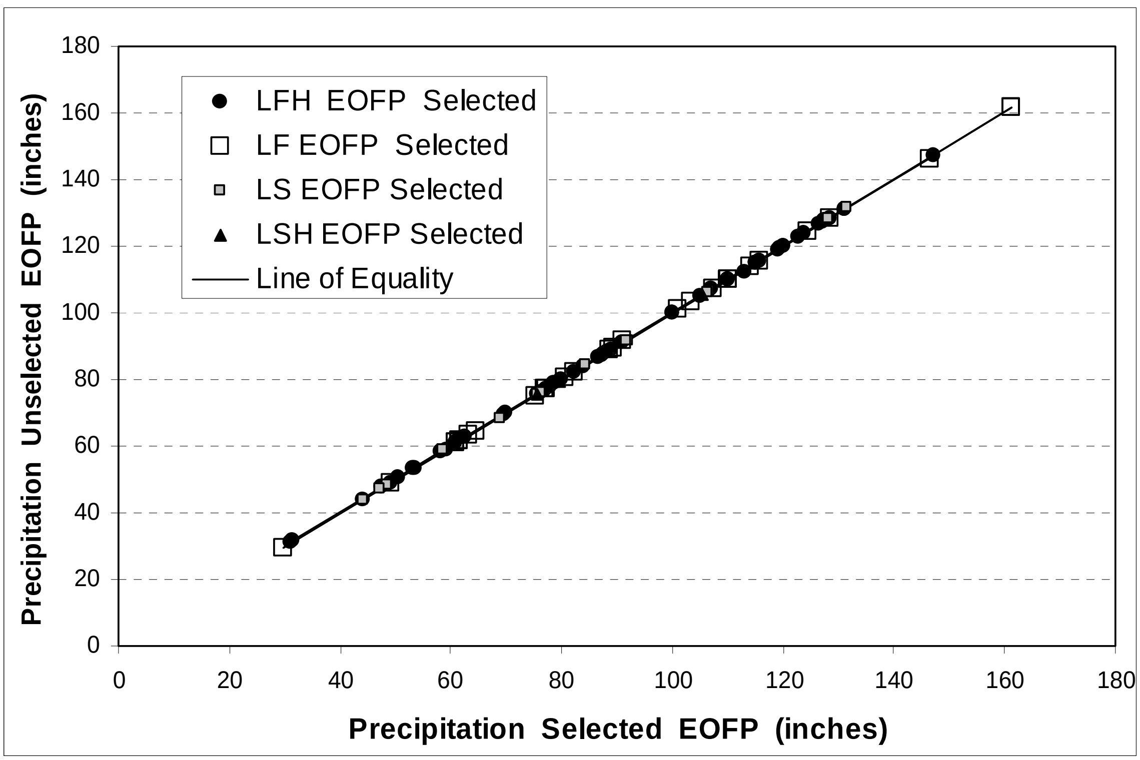
For 47 of the 112 cases (or 42% of conflicting EOFP pairs where one of the pair has no identifying information), the distance between the two EOFPs was less than 50 ft; this represents a difference in the placement of the points on the 10-m DEM network of essentially one 10-m cell. Selecting the EOFP with information resulted in the downstream point being selected in 54 of the 112 (48%) conflicting EOFP pairs.





Appendix Figure A5 Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs where one of the pair has no identifying information. Line of equality shown for reference.





Appendix Figure A5 (continued) Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting EOFP pairs where one of the pair has no identifying information. Line of equality shown for reference.

Comparison of Miscellaneous Conflicting EOFP Pairs

The seven remaining conflicting EOFP pairs were not members of the previous four groups so we combined them into a "Miscellaneous" category. These pairs were a mixture of EOFP point types and survey agencies. The majority of the cases (5) were collected by the same agency on the same survey date but were of different point types (Appendix Table A11). Four of the six cases had LSH|LS conflicting point types.

Appendix Table A11 Summary of the sponsoring agencies, survey date, and EOFP types for conflicting EOFP pairs in the "Miscellaneous" group.

Number of Pairs	Survey Agency #1	Survey Agency #2	Survey Date	Point Types
2	Wa. DNR	Wa. DNR	Same	LSH|LS
2	Tulalip DNR	Tulalip DNR	Same	LSH|LS
1	Champion	Champion	Same	LSH|LF
1	Quinault DNR	Hoh Fisheries	Different	LF|LS
1	Wa. Trout	Wa. DNR	Different	LFH|LSH

Summary statistics for the differences between the conflicting EOFP pairs in distance on the stream network and for the five GIS-derived physical attributes associated with each point are presented in Appendix Table A12. The mean DEM network distance between conflicting EOFP pairs with the same point type was about 788 ft (median distance = 665 ft).

Appendix Table A12 Summary of differences in distance on the stream network and GIS-derived physical attributes between conflicting EOFP pairs in the "Miscellaneous" group.

Statistic	Distance (ft)	Basin Size (acres)	Elevation (ft)	Upstream Gradient	Downstream Gradient	Precip. (inches)
Mean	787.7	19.3	68.5	3.4	6.5	0.17
Standard Error	320.6	6.9	14.5	0.7	1.8	0.13
Median	665.2	16.3	78.5	4.0	4.6	0.00
Minimum	33.8	0.03	6.2	0.3	2.0	0.00
Maximum	2,600.4	47.1	111.0	5.5	15.2	0.94

Except for the UPGRD attribute, the mean differences between the conflicting EOFP pairs in this group were the largest observed. Appendix Figure A6 shows the relationship of these data for BASIZE, ELEV, DNGRD, and PRECIP relative to the line of equality. The slope of the regression lines was not significantly different than one (1.0) for the BASIZE, ELEV, and PRECIP physical attributes. For these attributes, the y-intercept was not significant and was not included in the regression; the slope coefficient was very precisely estimated for all three of these attributes (coefficient of variation for the estimated slope less than 6%). There was not a significant linear relationship for the DNGRD comparison ($P = 0.627$).

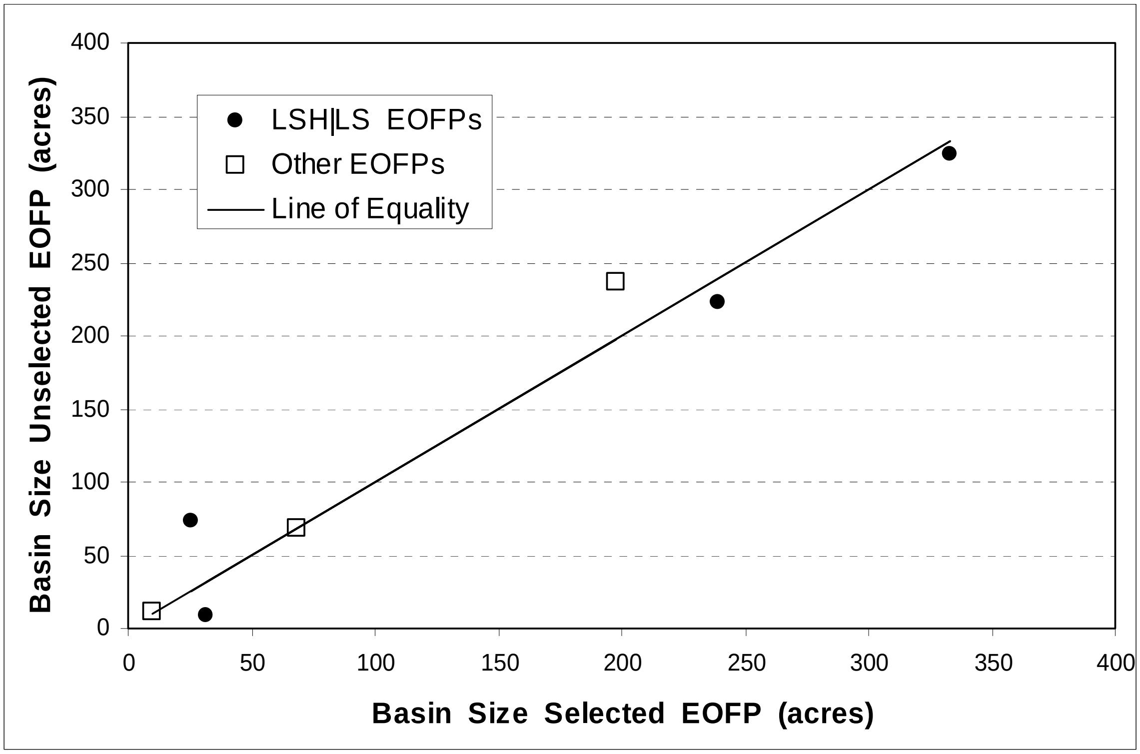
Only one of the seven cases had a distance between the two EOFPs less than 50 ft. The downstream EOFP was selected for three of the seven pairs when randomly selecting an EOFP from the conflicting pair.

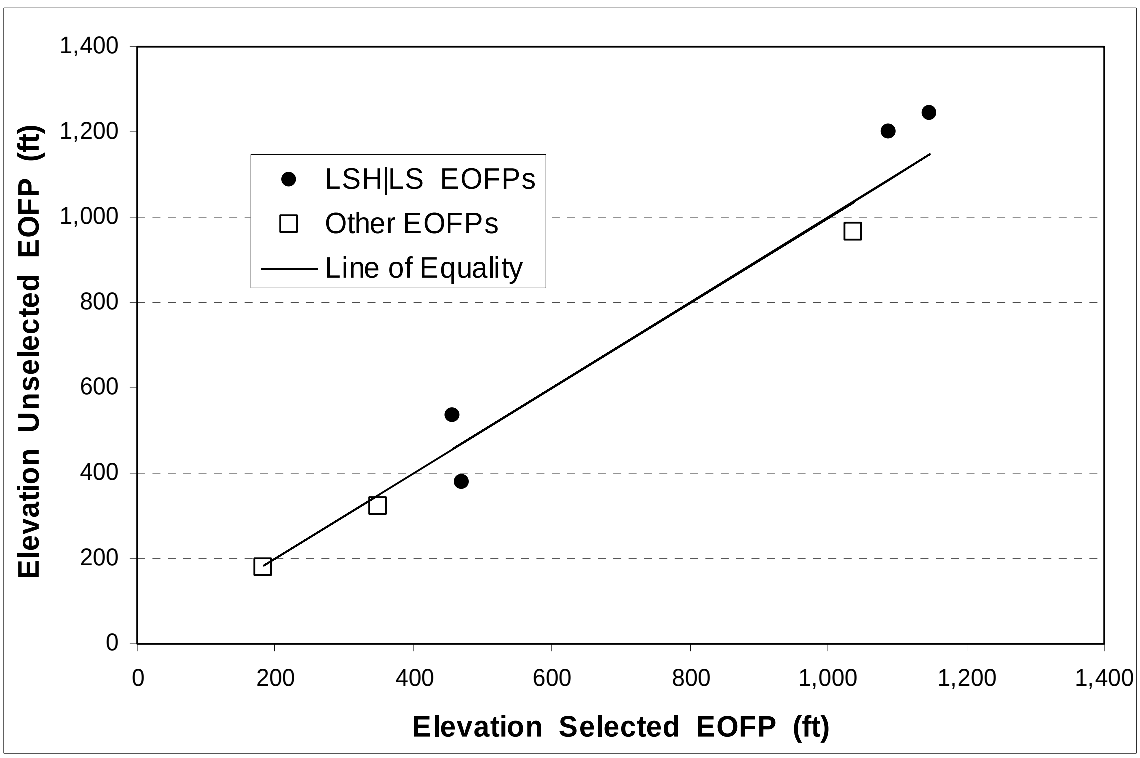
Summary and Discussion

It is important to recognize that the identification of any end-of-fish point during a field survey is subject to measurement variability. Measurement variability meaning that if the field survey was repeated using the same survey protocol but, for example, by a different survey crew or by the same survey crew on a different day, the exact placement of the EOFP might be different. Each survey protocol is subject to potential measurement variability but the sources of this variability are different among the EOFP types.

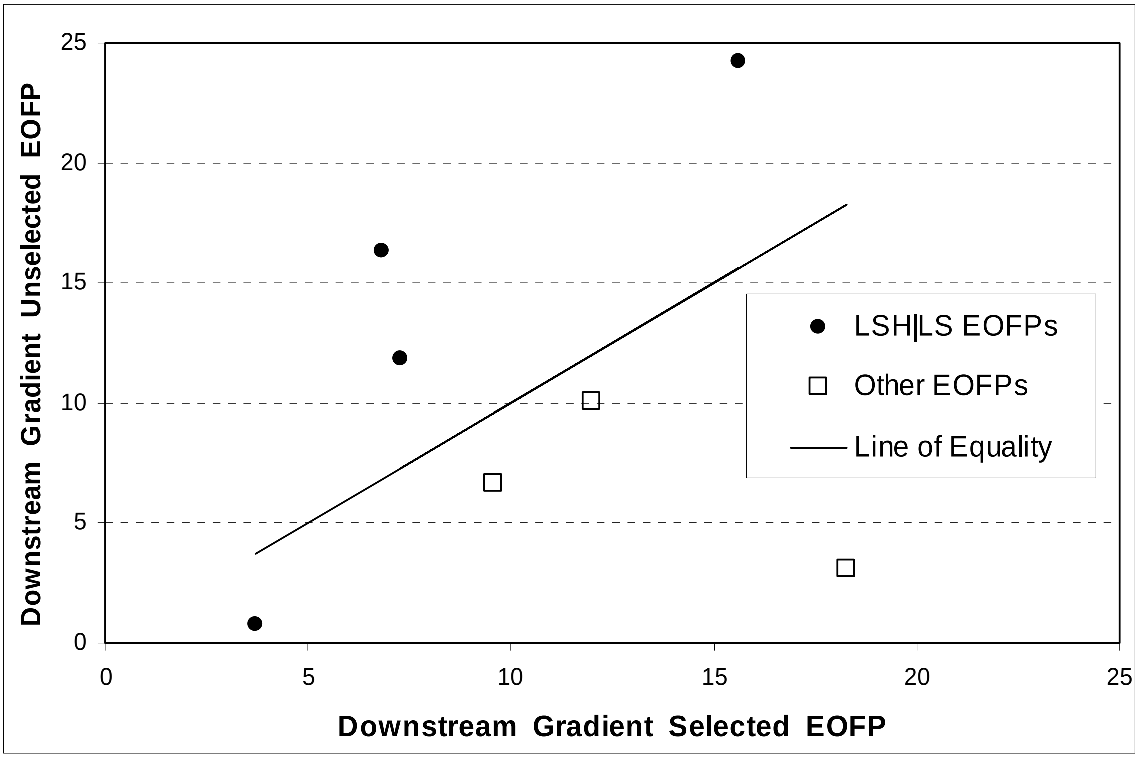
We are not aware of any studies that have examined the potential sources of measurement variability associated with any of the EOFP types or tried to estimate the amount of variability inherent in any field identification of an end-of-fish point. Nor are we aware of any studies that have quantified the differences in location between end-of-fish points identified on the same stream using different survey protocols. However, the data presented in this appendix allow a rudimentary examination of this variability - with one important caveat; these data were not collected following any experimental design and cannot be considered a random sample. Therefore, we must be cautious when interpreting the results of any comparison of the data.

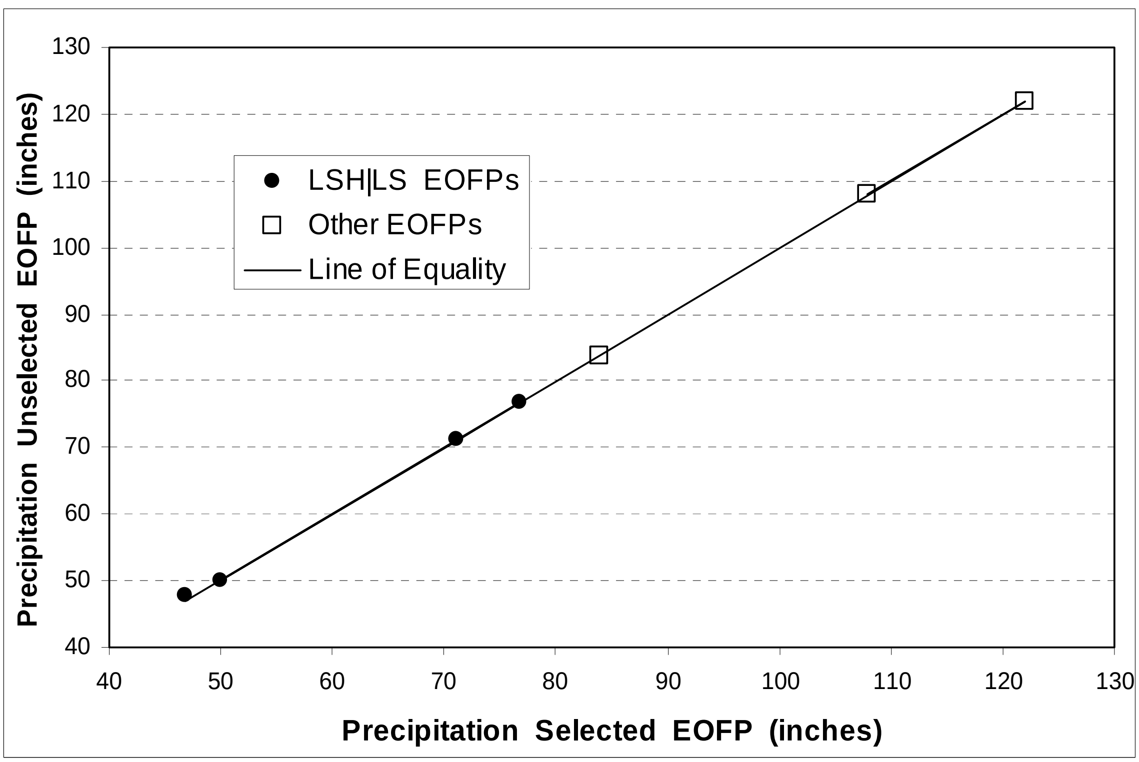
In general, there was very little difference between the conflicting pairs of EOFPs for the values of three of the physical attributes: basin size, elevation, and precipitation. The sizes of the differences for these attributes were not large enough to have a major influence on any logistic regression estimated from the data if one protocol were consistently chosen over another. The only physical attribute for which there were consistent differences in values between the conflicting EOFPs was the downstream gradient attribute. However, the logistic regression model is least sensitive to differences in the value of this attribute (see Appendix B).





Appendix Figure A6 Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting "Miscellaneous" EOFP pairs. Line of equality shown for reference.





Appendix Figure A6 (continued) Scatter plots comparing values of BASIZE, ELEV, DNGRD, and PRECIP for conflicting "Miscellaneous" EOFP pairs. Line of equality shown for reference.

APPENDIX B

Final Logistic Regression Model Sensitivity Analysis

A sensitivity analysis (Swartzman and Kaluzny 1987) was conducted which examined the sensitivity of the final logistic regression model output (estimated probability of fish presence) to the four physical attributes associated with each point on the DEM network that were model inputs: basin size, elevation, downstream gradient, and precipitation. The final LRM was used to estimate the probability of fish presence for a range of values for each attribute defined by a ±50% change from an initial baseline value. The change from the baseline value was done for only one attribute at a time while the values of the other attributes remained at their baseline value.

<u>Sensitivity Analysis Methods and Results</u>

Because inputs were altered by a percentage of their baseline value, the starting values for the attributes had some influence on the degree of change in the estimated probability of fish presence. For example, the range of elevations examined (and the range in estimated probabilities of fish presence from the LRM) would be greater if the baseline elevation was 1,500 ft compared to 500 ft. Therefore, four different sets of baseline starting values were used in the analysis. The four different baselines were:

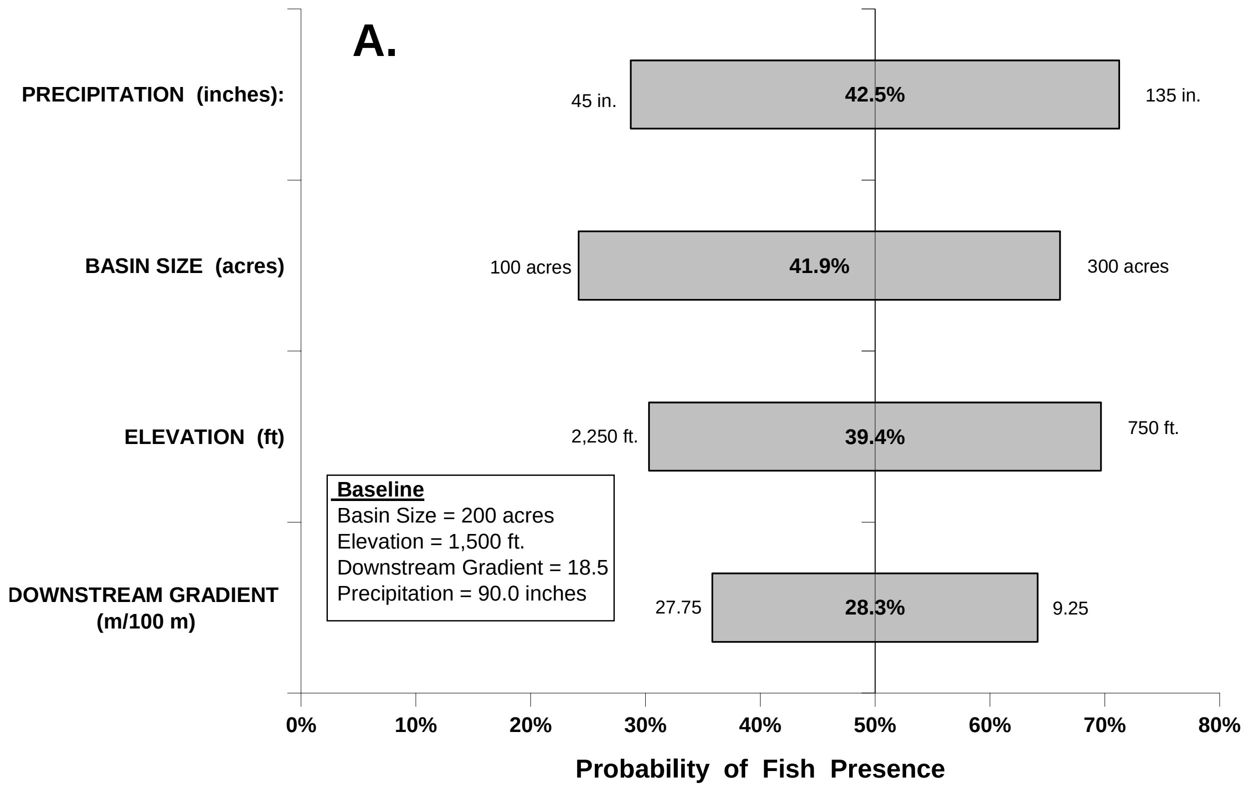
- Baseline A resulted in an estimated probability of fish presence of 50% and had attribute values of:
    - Basin size = 200 acres
    - Elevation = 1,500 ft
    - Downstream gradient = 18.5
    - Precipitation = 90.0 inches.
- Baseline B resulted in an estimated probability of fish presence of 8.5%. Its attribute values were the mean values for all fish absent points in FOSBs containing EOFPs with end types 1 or 2 only (Appendix Tables 2, 3, 4, and 6). Baseline B had attribute values of:
    - Basin size = 62.9 acres
    - Elevation = 1,793 ft
    - Downstream gradient = 20.5
    - Precipitation = 88.6 inches.
- Baseline C resulted in an estimated probability of fish presence of 98.2%. Its attribute values were the mean values for all fish present points in FOSBs containing EOFPs with end types 1 or 2 only (Appendix Tables 2, 3, 4, and 6). Baseline C had attribute values of:
    - Basin size = 734 acres
    - Elevation = 614 ft
    - Downstream gradient = 4.2
    - Precipitation = 88.1 inches.

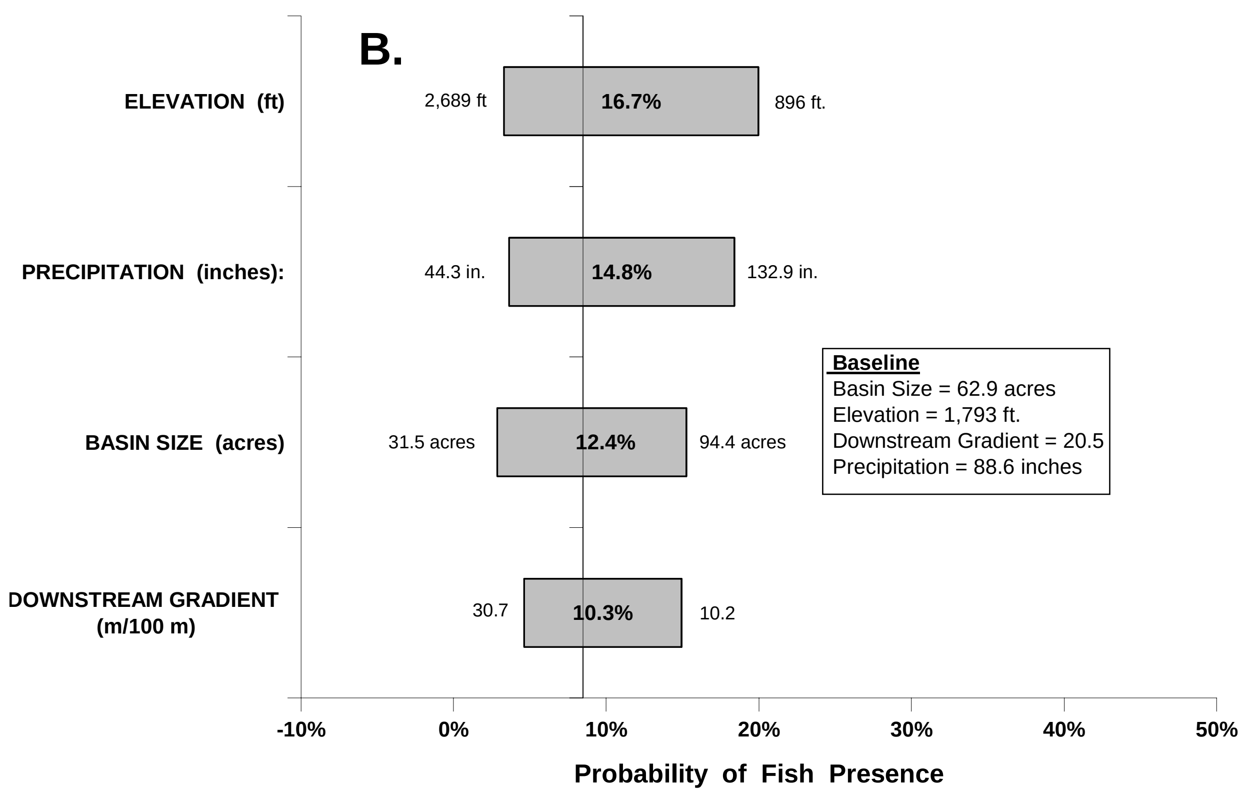
- Baseline D resulted in an estimated probability of fish presence of 41.9% and had attribute values of:
  - Basin size = 50 acres
  - Elevation = 500 ft
  - Downstream gradient = 5.0
  - Precipitation = 90.0 inches.

For each baseline, two values for each physical attribute were sequentially entered into the final LRM to estimate two probabilities of fish presence. First a value that was 50% of the baseline value was used in the final LRM and then a value that was 150% of the baseline value. All other physical attributes were kept at their baseline value. This procedure was followed for each attribute. The result was a range of probabilities of fish presence associated with a ±50% change from an initial baseline value for each physical attribute. The absolute differences between the two resulting probabilities for each attribute were calculated and the attributes were then ordered from greatest difference to smallest difference.

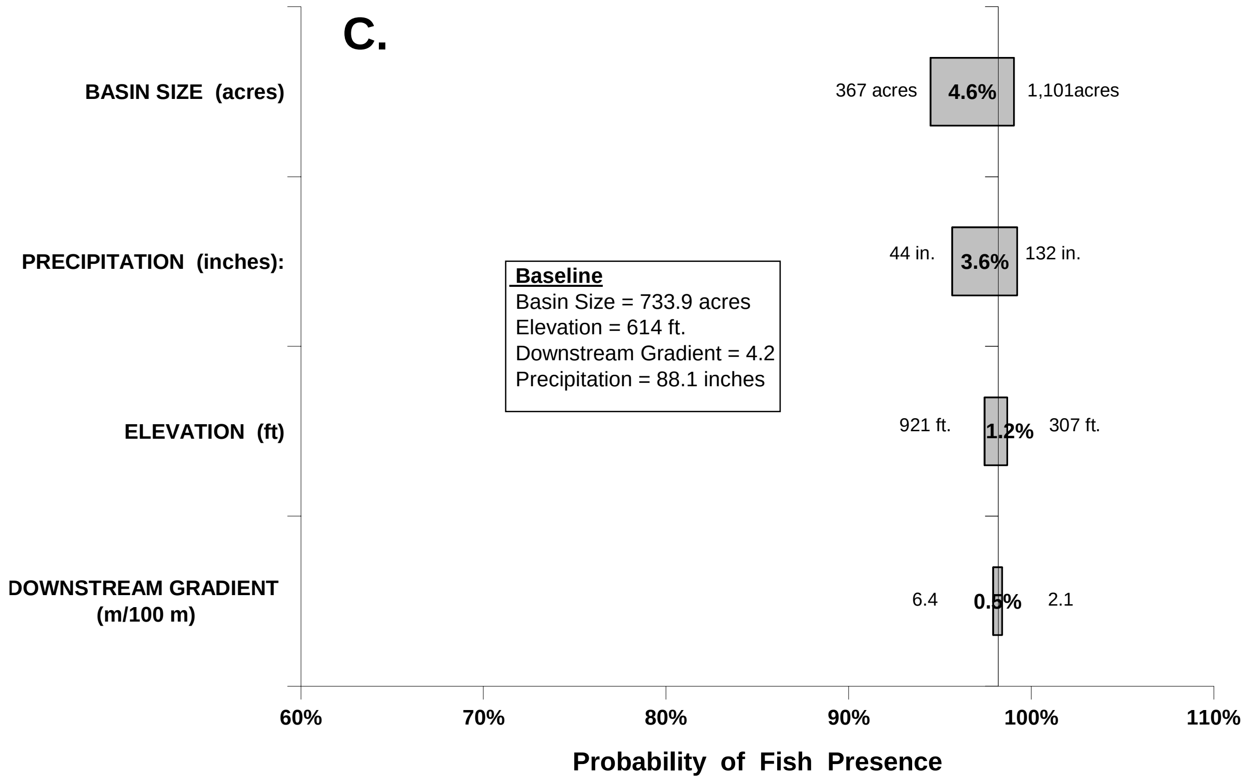
These differences were then graphed (Appendix Figure B1). For each physical attribute, the graphs display the range in estimated probabilities of fish presence defined by the range in attributes values from the baseline value (shaded box) and the absolute difference between the two estimated probabilities (percentage in the shaded box). The attribute values used to define the range are given next to the shaded box for each attribute. The attributes are ordered with the attribute the model was most sensitive to highest on the y-axis (i.e., the attribute with the largest difference between estimated probabilities of fish presence). The estimated probability of fish presence for the baseline values is denoted by the vertical axis bisecting the shaded boxes. A graph was produced for each baseline.

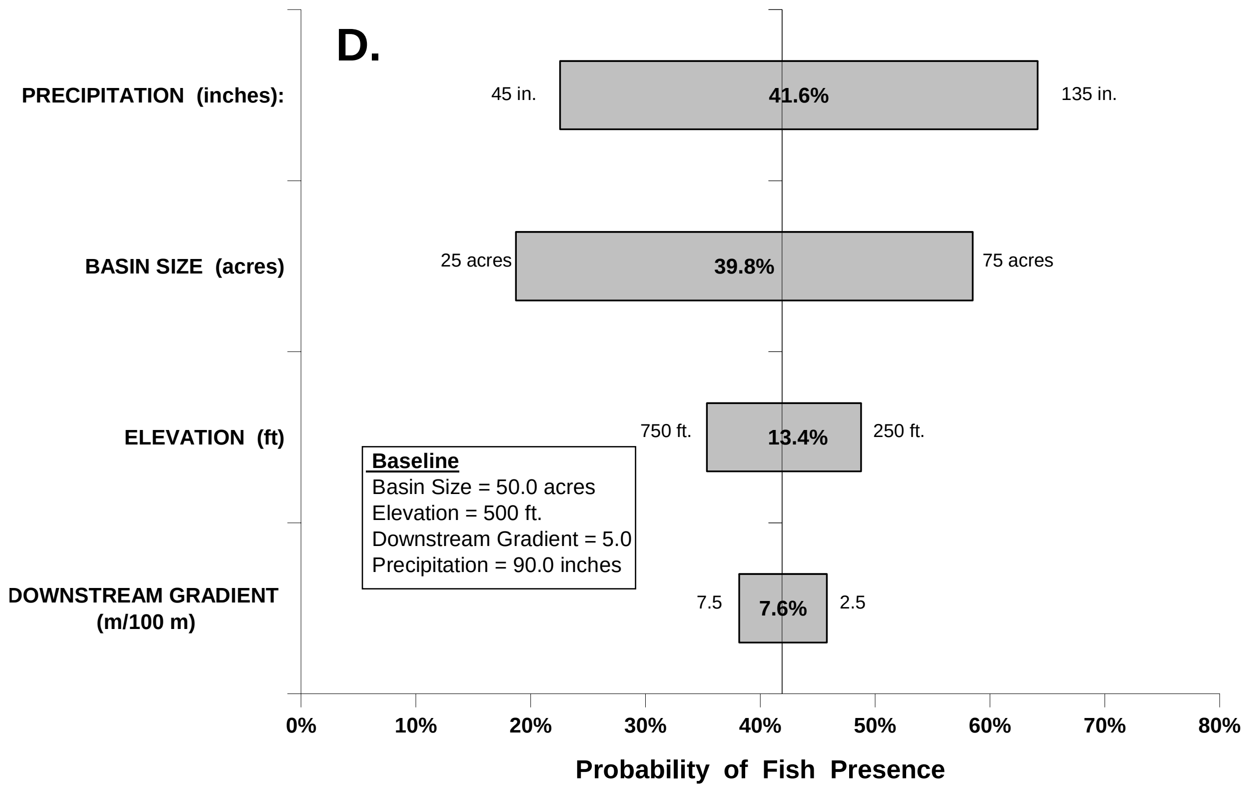
The one consistency among the results for the four baselines examined was that smallest range in probabilities was associated with the downstream gradient attribute. Therefore, the final logistic regression model appears to be least sensitive to changes in the downstream gradient attribute relative to the other attributes. At relatively high elevations (> 1,000 ft) and small basin sizes (less than 200 acres), the model is about equally sensitive to the basin size, elevation, and precipitation attributes (baselines A and B). At lower elevations (<1,000 ft), the model is about equally sensitive to the basin size and precipitation attributes (baselines C and D).





Appendix Figure B1 Graphs comparing the sensitivity of the estimated probability of fish presence to a ±50% change to each physical attribute (one attribute at a time) from the specified baseline values (figure continued on next page).





Appendix Figure B1 Graphs comparing the sensitivity of the estimated probability of fish presence to a ±50% change to each physical attribute (one attribute at a time) from the specified baseline values (continued).

APPENDIX C

Process for Reconciling Inconsistencies and Eliminating Duplicated Error in the Random Error Assignment Process

This appendix explains the steps for processing an entire section of the stream network, such as a fourth order sub-basin (FOSB). The process begins by applying the logistic regression model and stopping rule to an entire section of the stream network and predicting end-of-fish points (EOFPs) throughout the section. Error distances are calculated for predicted end-of-fish points associated with survey-based EOFPs. Error distances for EOFPs which have no associated survey-based EOFP are randomly assigned. These steps are followed to implement the total sub-basin error assessment for selected FOSBs described in the section on validation in the report.

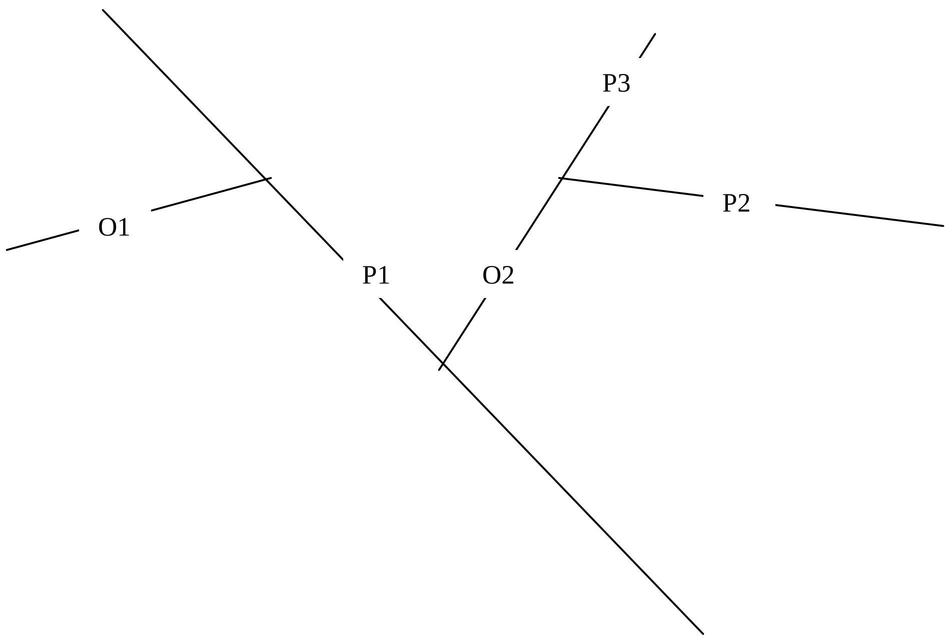
1) Apply the logistic regression model and stopping rule to the GIS network data for a sub-basin (or larger section of the stream network).
2) Calculate an error distance for each survey-based EOFP present in the data.
    a. For each survey-based EOFP search up and downstream until you find a predicted EOFP. When searching upstream always follow the branch with largest basin area (i.e., the mainstem).
    b. Calculate the error distance between the observed and predicted EOFP. If the predicted EOFP is downstream of the survey-based EOFP and, when moving down to the predicted EOFP you encounter a junction with a larger stream, a predicted EOFP is added at the junction and the error distance and covariates (described in c) are based on the new predicted EOFP (see Example 1).
    c. Record the basin area, boundary type, and other potential covariates for the predicted EOFP.
3) Assign an error distance for each predicted EOFP in the basin that is not associated with a survey-based EOFP.
    a. For each predicted EOFP draw a random error distance based upon the basin area and boundary type (A, B, or C).
    b. If the random error selected is upstream of the predicted EOFP (a negative error is drawn):
        i. Move upstream until you reach the error distance or the end of the mainstem.
        ii. For each tributary you encounter, repeat this process, i.e.:
            1. randomly draw an error from the error pool for lateral points with comparable basin area.
            2. move upstream the appropriate distance (and recursively deal with any additional tributaries you encounter in the same way).
        iii. Tally the error distances to create a total for that predicted point.
    c. If the random error selected is downstream of the predicted point (a positive error is drawn)
        i. Move downstream until you reach the error distance or the junction with a larger stream.

<u>Dealing with survey-based EOFPs in the random error assignment process.</u>

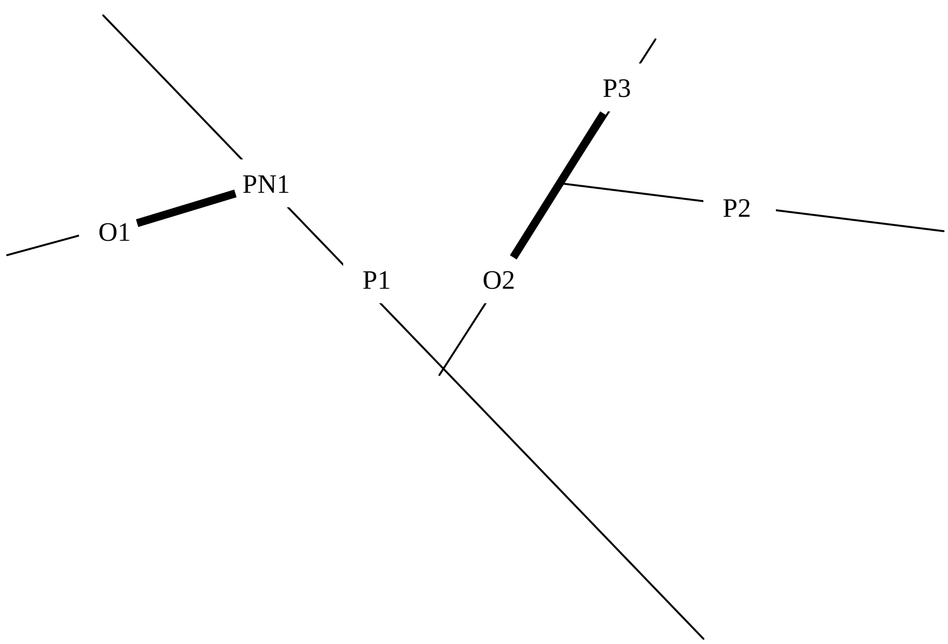
   d. If there is a survey-based EOFP below the predicted EOFP, the error distance is the distance to the survey-based EOFP or to the nearest junction with a larger stream (whichever comes first).
   e. If there is one or more survey-based EOFPs above a predicted EOFP:
      i. Calculate the mainstem error distance.
         1. if one of the survey points is on the same mainstem as the predicted point then the error is just the distance between the two.
         2. if there is no survey point on the same mainstem then move upstream until you reach the junction with the last tributary that has a survey point. Draw an upstream error distance and add it to the distance to the predicted point to get the total error distance.
      ii. Using the calculated error distance, follow the procedure for no survey points (3 a.) except when you hit a tributary with a survey point (in which case you should repeat 3e).

4) Sum all of the error distances. This should provide a total that does not include any overlap.
5) Repeat the whole random error assignment process several times to get an idea of the variability in the error estimate.

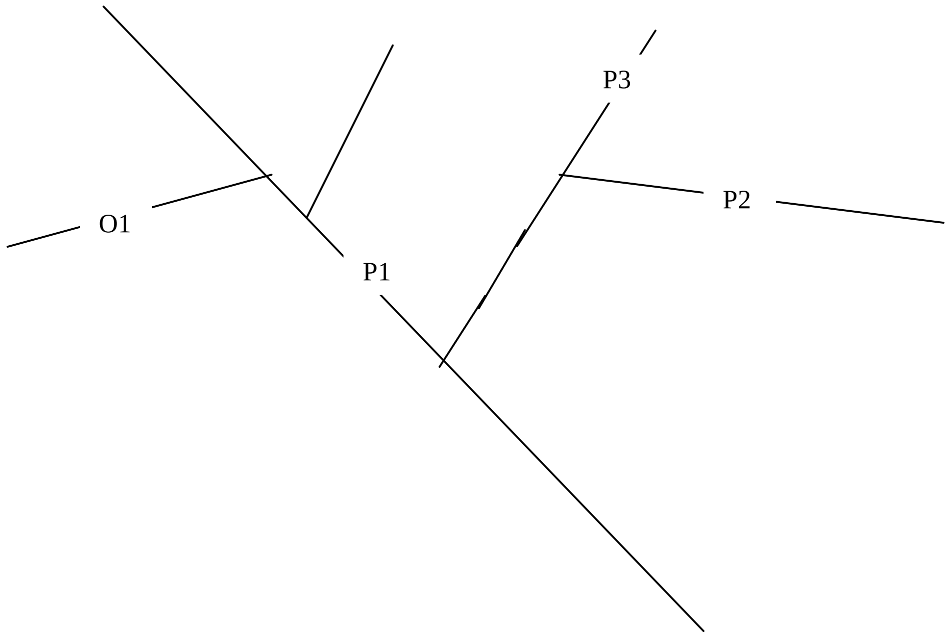
Example 1.



For survey-based EOFP 1 (O1), the prediction (P1) is downstream and there is a junction with a larger stream before you reach the prediction. Therefore a new prediction is placed at the junction (PN1 below) and the error distance calculated for O1 is between O1 and that new prediction. For observation 2 (O2), the upstream prediction (following the mainstem) is P3 and the error distance is between O2 and P3.



Example 2.



Now apply step 3 (step 2 is applied in Example 1). We have essentially the same network with O2 removed and an extra tributary added. For P1 calculate the error distance upstream using the procedure in 3e to make sure that the randomly drawn error distance is greater than the distance to the junction with the tributary with O1 (e.g., 300m). So go 300m up the mainstem. For the first tributary randomly draw another error distance from the error pool for lateral points with comparable basin area. For this example, say it is 10m. When you get to the second tributary there is an observed point on its mainstem so that point is used to calculate the error distance. For P3, say the randomly drawn error distance is downstream 500m and the downstream junction with the larger stream is 300m. Then the 300m distance to the downstream junction is used. For P2, the randomly drawn error distance is downstream 10m.