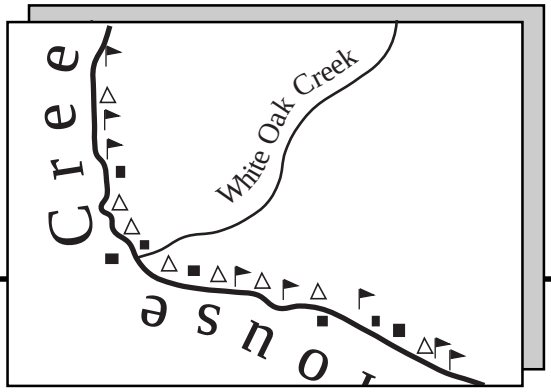
FHR Currents...

Region 5 Fish Habitat Relationships Technical Bulletin
Number 11 February 1993




# A Technique for Moving Existing Fish Habitat Data Sets Into The Spatial Environment Of A Vector Geographic Information System

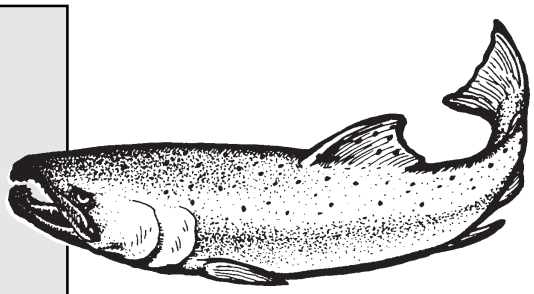
Michael F. Martischang
USDA Forest Service, Six Rivers National Forest
1330 Bayshore Way
Eureka, CA 95501

## Abstract

This article provides a general overview of techniques that were used to prepare existing fish habitat data sets so they could be transferred into the spatial analysis environment of a vector geographic information system (GIS). By refining the application of "address matching" software tools, a GIS was used to derive a new spatially defined habitat data set that indicated the location along a watercourse that each habitat unit occurred. Address matching (also known as "geocoding") is an automated process that compares addresses from two data sets for similarity. Typically, one data set has no spatial context (e.g., fish habitat data) and the other has a well defined spatial extent (e.g., a digitized watercourse). If addresses from the different data sets match, then relationships between the data sets can be established. Matched addresses result in the creation of a third data set that can be graphically displayed to show the location of fish habitat units along the watercourse. This newly derived habitat data set can then be manipulated in a GIS setting where it becomes possible to analyze the proximity and condition of habitat units in relation to other features and processes represented in the GIS (e.g., soils, geology, roads, vegetation, et cetera). The basic address data structures necessary for address matching also provide the foundation for conducting a variety of other network analysis involving soils, geology, vegetation, and road data themes. Data sets containing address structures can be used to model movement of such things as sediment, cold water, woody debris, and fish (both up- and down-stream) within a watershed unit, enabling assessment of interactive effects of these various themes on habitat quality, distribution, and use by fish.

*USDA*
*Forest Service*
*Pacific Southwest Region*

## *FHR Currents…*
## Purpose



The Fish Habitat Relationships (FHR) Program of Region 5, Pacific Southwest Region, USFS, has been established to research and develop information on fish ecology and to coordinate effective applications of this knowledge in managing and protecting our fisheries. By relating life stage requirements of specific species to physical habitat parameters, we are aiming at our main objective: developing a methodology to manage fisheries through the management of habitat.

## Submissions:

If you wish to submit a paper for publication in *FHR Currents*, please write Jerry Boberg, Dave Fuller (Technical Editors) or Stephanie Gomes (Editor/Layout) for information and guidelines at: Six Rivers National Forest, 1330 Bayshore Way, Eureka, CA 95501; or call (707) 442-1721.

# Introduction

## Addresses

The notion of “addresses” and “address matching” in relation to fish habitat may seem incongruous at first and warrants clarification to better establish the applicability of this concept and technique to fish habitat location determination.

The definition of the term “address” as used in this article is no different than that which we use to describe where we live, work or shop. Figure 1 is a schematic drawing of two common address data structures. A “2-number” address structure indicates that address values for each segment of a route are defined with one beginning value and one ending value. 2-Number address structures provide

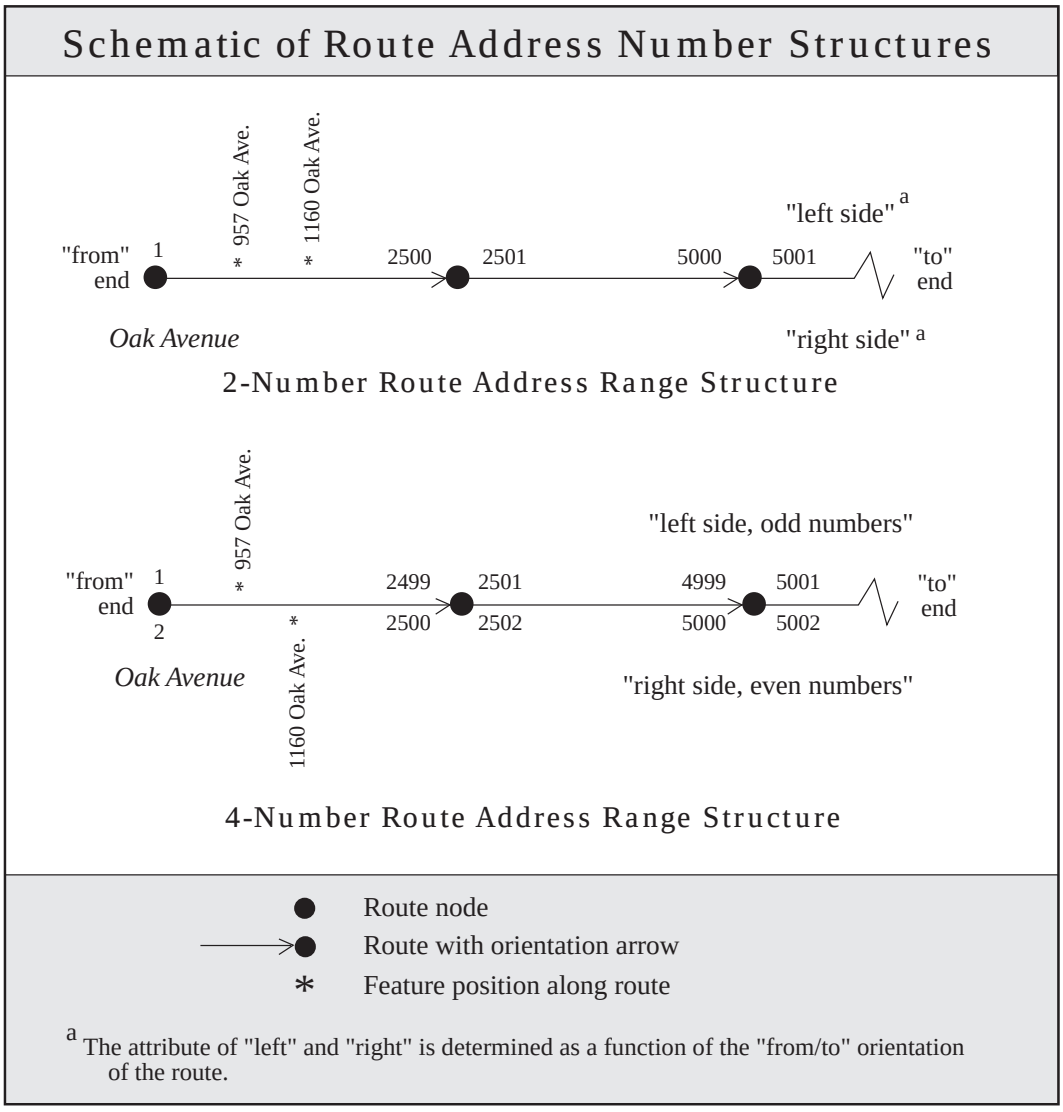



Figure 1. Schematic examples of 2- and 4-Number address range structures for a route (adapted from ESRI 1990). 2-Number addresses have two address values associated with each route node. 4-Number addresses have four address values associated witheach route node.

indications of relative location along the route. A "4-number" address structure indicates that each segment of a route is defined by four values: odd-valued beginning and ending numbers and even-valued beginning and ending numbers.

Address numbers provide us with references to relative location. In both the 2- and 4-number schematics in Figure 1, the address number 957 comes "before" 1160. However, with a 4-number address structure we could further conclude that the address 1160 occurs on the "right" side of the route. The concepts of "before," "after," "left," and "right" are relative and are a function of which direction one is facing along the axis of the route. For the sake of consistency, these notions of relative location are usually referenced from the origin of the route (where the first address number is "1") and assume that the traveler is always facing toward the "end" of the route. While most streets permit two-way travel, it is the direction of address numbering along the street that implies the "direction" of the street. Thus, the route segment directional arrows in Figure 1 point in the direction of ascending address values.

A typical street address also contains a route name portion and we often include a zip code as part of an address to help resolve ambiguity. For instance, the address "1160 Oak Avenue" might actually exist at hundreds of locations throughout the United States. If we included a zip code with that address, say "95501," we eliminate ambiguity and uniquely identify the "1160" address location along Oak Avenue in Eureka, California.

Route names simply refer to the "conduit" along which "things" move to get to "places" along that conduit. Using mail delivery as an example, a letter ("thing") is delivered by a letter-carrier walking along Oak Avenue ("conduit") to a house on the right side of Oak Avenue at 1170 ("place"). Many interconnected conduits of the same kind (e.g., roads) constitute a "network." Given sufficient data about the network, it becomes possible to model movements of a wide variety of things through the network.

By stretching our imaginations we can substitute a wide variety of scenarios to this basic concept of the network model. The "conduit" could be a power line, delivering a "thing" called electricity to "places" called substations and residences. Watercourses are natural networks whose conduit is flowing water through which many things move (fish, sediment, woody debris, water) to many kinds of places (pool habitats, barriers, oceans).

By fashioning an appropriate address data structure and incorporating it to fish habitat data sets and digitized watercourse data sets, it becomes possible to use address matching tools of a GIS to derive the "places" where fish habitat are located along a watercourse. Furthermore, if enough data about the watercourse network is incorporated to the watercourse database (e.g., shade, velocity, discharge, gradient, et cetera), it could become possible to use the GIS for modelling movement of many kinds of "things" (cold water, fish, sediment, woody debris, toxic chemical spills, et cetera) in either direction along the network.

## Purpose of & Need for the Fish Habitat/GIS Project

Six Rivers National Forest (SRNF) fisheries staff was interested in refining techniques previously developed by Hemstrom (1989) that could be used to determine the locations of fish habitat units within watercourses that had been previously inventoried. Over 20 anadromous streams on the SRNF have been previously habitat typed (McCain et al 1990). These data are stored in personal computer databases and are extensively used by biologists to assist in habitat monitoring, enhancement, and restoration.

Within the database environment, any number of queries, statistical evaluations, and tests could be performed on the data to provide useful information to biologists. However, a grasp of "where" all these individual habitat units were in relation to each other and the surrounding landscape remained an intuitive process that usually required an intimate familiarity with a particular watercourse and the habitat data that had been collected from it. The large volumes of habitat data associated with an

entire watercourse system made graphic depiction of habitat unit locations and distributions, using either manual or computer-aided drafting methods, infeasible.

# Objectives

In the interests of preparing for the imminent implementation of Project 615[1] and being able to graphically represent habitat locations relative to features and processes beyond the watercourse channel, SRNF fisheries staff requested faculty of the College of Natural Resources and Sciences and staff of the California Cooperative Fishery Research Unit, both located at Humboldt State University in Arcata, California, to identify and refine techniques for representing fish habitat locations in a GIS using "off the shelf" technologies. Two objectives were identified to guide project development:

1. Identify procedures for preparing *existing* fish habitat data sets (consisting of both main and side-channel habitat descriptions) for input to and manipulation by the corporate GIS envisioned by the Forest Service integrated information management system, Project 615 (USFS 10/22/91, 1991b, 1992b; USFS/R6 1991; USFS/R8 n.d., 1989; USFS/R10 1989; Boberg, Stewart, pers. comms.)

2. Recommend revisions to the current habitat inventory and biological survey techniques that would facilitate more efficient and accurate integration of *future* data sets to the corporate GIS environment (Boberg, pers. comm.).

# Project Area

Habitat data from Grouse Creek were chosen to demonstrate the use of address matching methods. Grouse Creek is a 15.9 mile long tributary to the South Fork of the Trinity River and is located in the northwest corner of California, about 26 air miles east-southeast of Eureka (Figures 2 and 3). Within the context of the hydrologic unit code (HUC), Grouse Creek is designated as a "National Forest System (NFS) subwatershed." The HUC represents a standardized, hierarchically nested series of water catchments originally defined by the U.S. Geological Survey (USGS) and further subdivided by the Forest Service (Seaber et al. 1987; USFS, 6/1/90, 1989; Steinblums, pers. comm.). Table 1 briefly illustrates the hierarchical structure of the HUC as applied to the Grouse Creek NFS subwatershed. Most of the subwatershed falls within the administrative boundary of the SRNF, Lower Trinity District, with land ownership mixed between public (USFS, about 55 percent), corporate, and individual entities. The Grouse Creek subwatershed drains about 36,000 acres. Elevations range from about 900 to over 5600 feet above sea level. Mixed conifer forests are typical and dominated by Douglas fir and white fir with interspersed tan oak, madrone, incense cedar, pine, and chinquapin (Raines and Kelsey 1991). The natural biological and physical complexity of the subwatershed has been made even more complex over the past 45 years by different management practices of the various landowners (Boberg, Furniss, McRae, Morrison, Smith, pers. comms.).

## Previous Work

Hemstrom (1989; pers. comm.) had used pcARC/INFO to demonstrate how locations of fish habitat in Cummins Creek, Oregon could be determined using address matching techniques. As discussed in the Introduction, the integer portion of an address is usually a dimensionless value indicating the relative position of some feature along a route. However, in some instances the distribution network and the location of features along that network lend themselves to *distance* measurements referenced from unambiguous starting points. Hemstrom exploited the virtue of watercourse length as measured from the mouth of Cummins Creek and calculated address numbers for main-channel habitat units based on the distance of the downstream end-point of each unit from the mouth of the creek. When the GIS displayed the habitat unit markers derived as a result of address matching, their locations were distributed along the entire inventoried length of Cummins Creek as a function of their distance from the mouth of Cummins Creek.

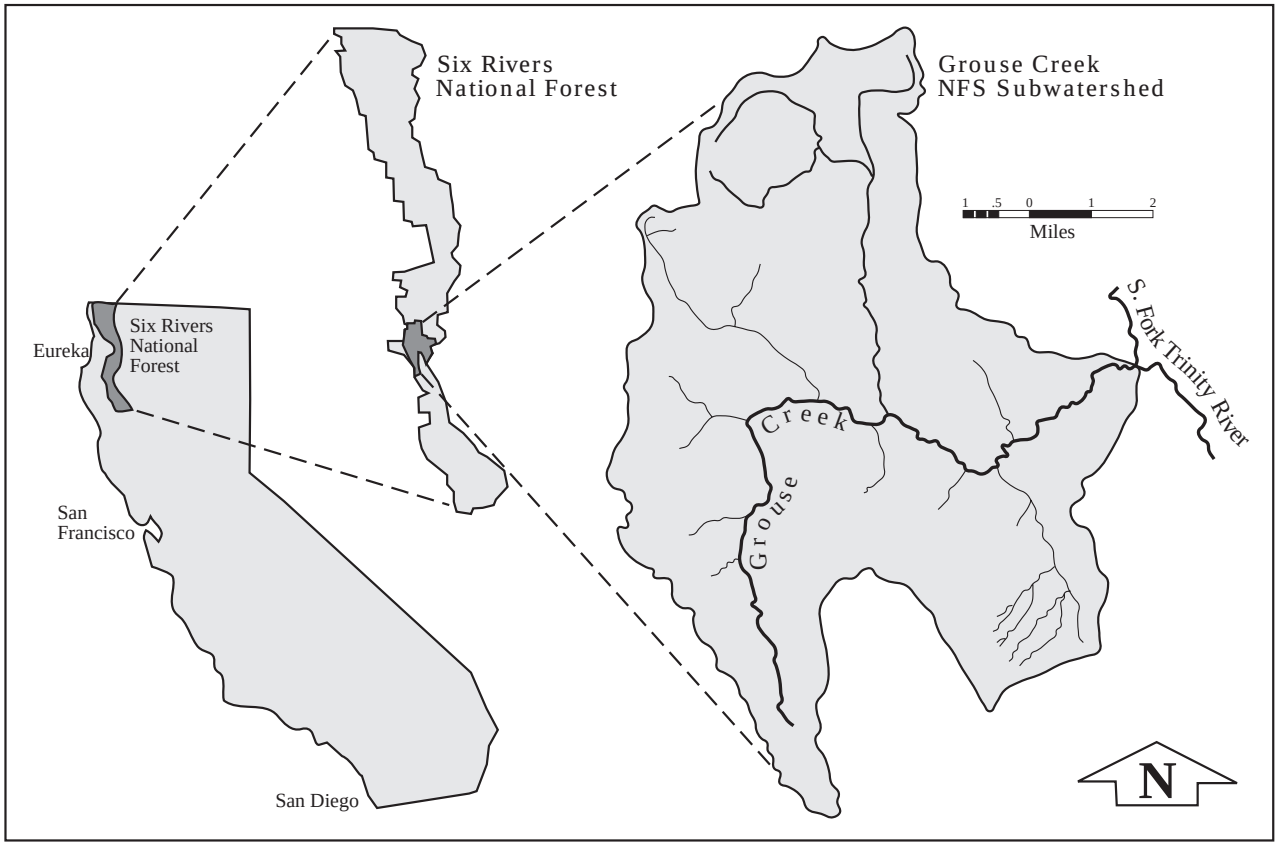



Figure 2. Vicinity maps highlighting locations of the Six Rivers NF, Grouse Creek subwatershed and the mainstem of Grouse Creek.

The variable distances between unit markers were indicative of the variable habitat unit lengths.

For example, refer back to the 2-number address schematic in Figure 1 and substitute "Cummins Creek" for "Oak Avenue" and assume that the integer portion of addresses are in units of meters. Thus the marker symbol at "1160" would indicate the downstream end-point of a habitat unit that is located 1,160 meters upstream from the mouth of Cummins Creek. Within the GIS database, descriptive habitat data were associated with each point marker, making it possible for Hemstrom to query the newly created fish habitat data layer and have only those habitat units that met criteria of the query displayed. Rather than a traditional tabular summary of the query results, the GIS provides a graphic "map" of the query results. When viewing habitat unit locations along with vegetative, soils, and valley form data themes, it became possible to identify spatial relationships between specific habitat units and specific features in the surrounding landscape.

## Data Necessary for Address Matching

To accomplish the objective of moving existing habitat data into a GIS environment, three data sets were required: fish habitat data, watercourse data, and watershed data. Using address matching commands of the GIS, individual habitat unit addresses would be matched against address ranges of watercourse arc segments. The watershed data set was developed to delineate areas that would define "wildland zip code zones" -- discrete basins within which features occur.

## Fish Habitat Data

A complete Grouse Creek fish habitat database file was provided by SRNF personnel. Total file size was 524KB, with 251KB of this consisting of a "COMMENTS" attribute with text entries. Over 70,000 feet of the physical in-stream fish habitat in Grouse Creek had been classified during 1988-89 into 868 fish habitat units. These habitat units were represented as 868 records in the database file. They were arranged in an "upstream" fashion such that the first record described the first habitat unit beginning at the mouth of Grouse Creek, and so on. All habitat units were broadly identified as being "main-channel" or "side-channel". Main-channel features consisted of 778 contiguous units and side-channel features consisted of 90 discontinuous units. Within the database, side-channel records were scattered among the main-channel records. Habitat unit lengths ranged from six feet to over 1100 feet. Intermittent azimuth data had been collected. The data set was devoid of any georeferences tied to absolute earth coordinates, although occasional location calls identified tributary confluences. Forty-five descriptive attributes were associated with each unit (e.g., stream dimension, cover complexity, substrate composition, channel morphology) (USFS/R5 1990; McCain et al. 1990; Boberg, Fuller, Kenfield, Ober, pers. comms.).

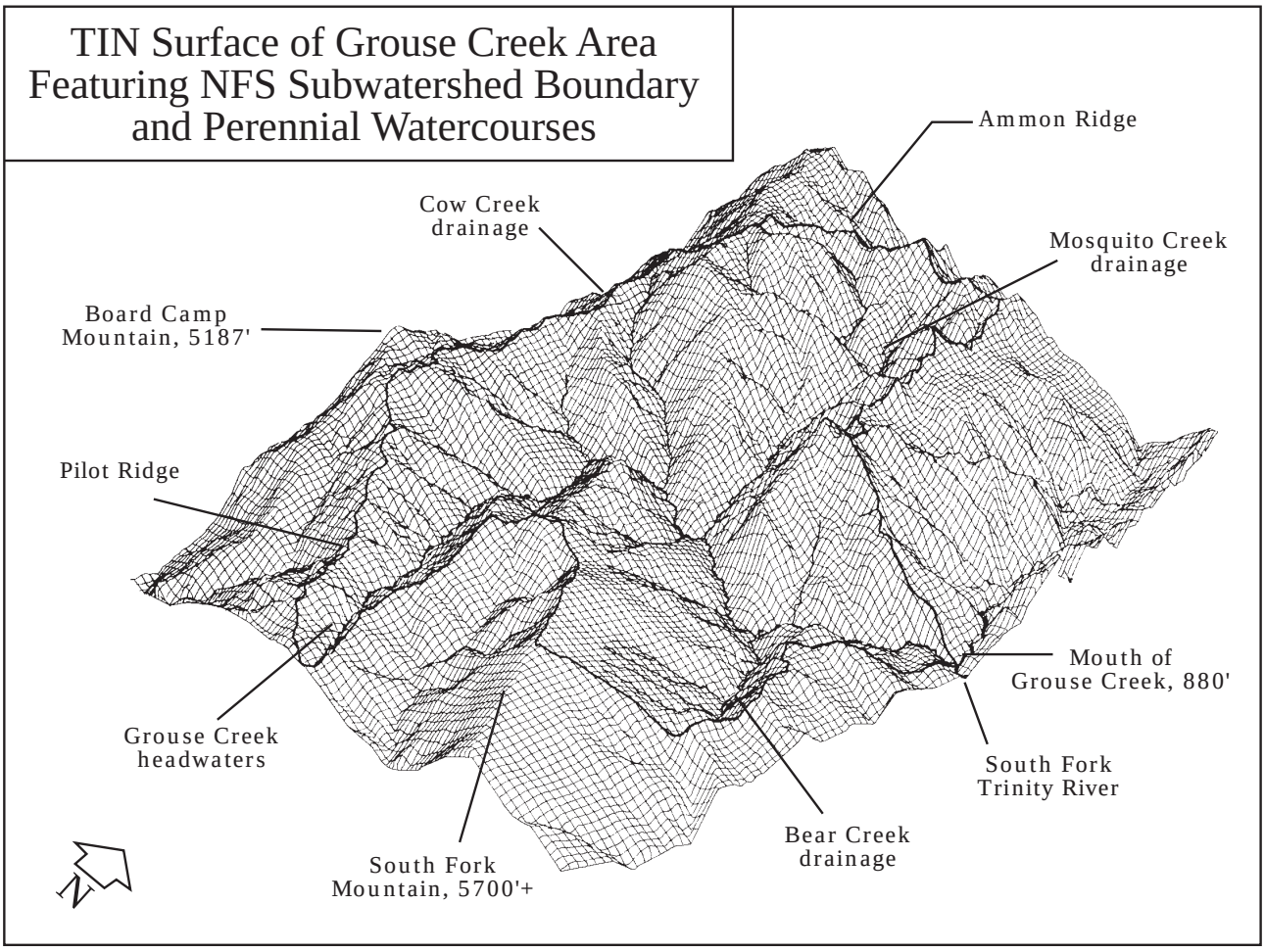
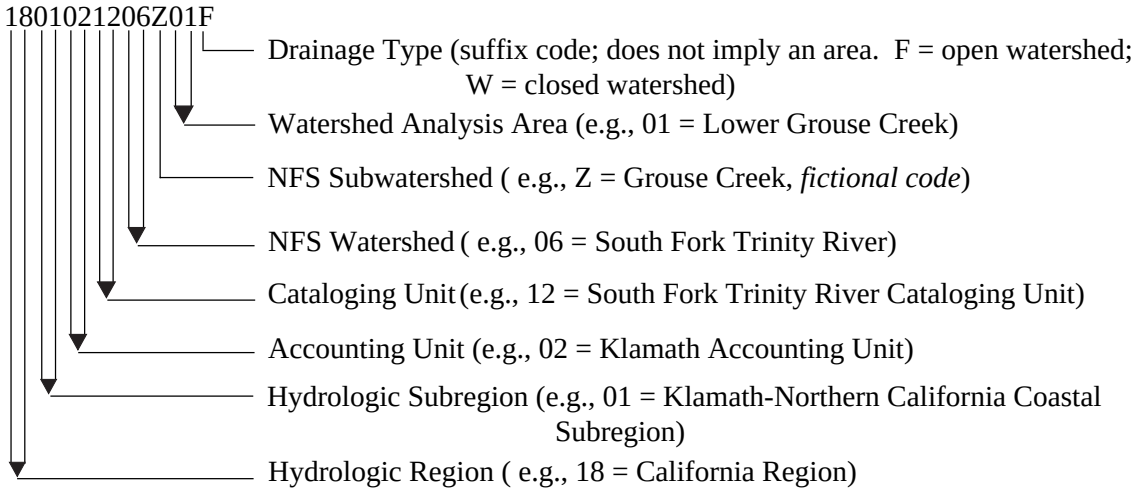



Figure 3. A quasi 3-dimensional perspective view of the Grouse Creek NFS subwatershed boundary and the perennial watercourse channels therein. Some prominent geographic landmarks and major tributaries to Grouse Creek are identified.

| Attribute Position[a] | Attribute Width & Data Type | Attribute Name | "Average" Hydrologic Area[b] (Square Miles) | Attribute Steward |
|---|---|---|---|---|
| 1st | 2, integer | Hydrologic Region | 50,000 | USGS |
| 2nd | 2, integer | Hydrologic Subregion | 25,000 | USGS |
| 3rd | 2, integer | Accounting Unit | 15,000 | USGS |
| 4th | 2, integer | Cataloging Unit | 1,000 | USGS |
| 5th | 2, integer | NFS Watershed | 750 | USFS/WO |
| 6th | 1, character | NFS Subwatershed | 80 | USFS/WO |
| 7th | 2, integer | Watershed Analysis Area | 8 | USFS/R6 |
| 8th | 1, character | Drainage Type | n/a | USFS/R6 |

[a]Example of the HUC for the Lower Grouse Creek Watershed Analysis Area (WAA) (3, 810 acres) separated into its component parts to illustrate code structure. The example code below uniquely identifies the Lower Grouse WAA in the United States.




[b]Hydrologic area varies within wide ranges. Greater topographic relief results in subdivision to more units of smaller area.

Table 1. Attribute structure and "stewardship" of the 14-character hydrologic unit code (HUC) as used for the Grouse Creek GIS project. EPA and USGS hydrologic data sets are resolved to the cataloging unit. The USFS has extended the HUC for resolution to smaller catchments. The HUC is conceptually similar to the postal zip code, with two primary differences: hydrologic unit boundaries are generally based on natural, physiographic features and they consist of a nested hierarchical structure that enables unique identity of various size areas.

### Watercourse Data

Using pcARC/INFO, all watercourse features were hand-digitized from four 1:24,000 original mylar primary base series (PBS) quadrangles (i.e., 7.5 minute by 7.5 minute quads). Had cartographic feature files (CFFs) been available, they would have been used instead[2]. All watercourse arcs were digitized in an upstream direction, resulting in a contiguous "from/to" network topology that imparted consistent attributes of "left" and "right" to watercourse arcs (Kiser, Lienkaemper, pers. comms.). Watercourse arc end-points, or nodes, were set at perennial/intermittent stream breaks (as depicted on the PBS), dangling (headwaters) arcs, and confluences. The upstream orientation of arcs corresponded with the upstream orientation of the fish habitat data set. Grouse Creek itself was represented with 25 line segments in the watercourse layer database. Note that references to "left" or "right" are relative to watercourse arc orientation and *do not* imply any relative position relationship of fish habitat units to one another.

### Watershed Data

The hierarchically-ordered HUC, as extended by the Forest Service, was used to provide the structure and syntax for defining watershed boundaries, names, and unique identification codes (refer to Table 1). The HUC consisted of 14 characters in eight numeric and alpha fields. Fields one through four were defined by the U.S. Geological Survey (USGS) (Seaber et al. 1987); fields five and six were defined by the USFS (6/1/90, 1989); fields seven and eight were defined by the Siskiyou National Forest (SNF), of the USFS (Steinblums, pers. comm.). Fields seven and eight identify a "watershed analysis area" (WAA) roughly 3,000 to 7,000 acres in size. The HUC can be truncated to suit various resolutions of analysis. The Environmental Protection Agency (EPA) and USGS typically reference their hydrologic data sets to the cataloging unit (CU) extent (Dulaney 1991; Hansen, pers. comm.). Forest Service needs warranted higher resolution, hence the code extensions implemented by the SNF to enable unique identification of smaller catchments.

The HUC was incorporated to digitized watercourse and fish habitat data sets as an address "zone" attribute and was used to resolve address ambiguity, much as the nine digit zip code is by the U.S. Postal Service. Boundary delineation and coding of watersheds to the WAA level had not been accomplished on the SRNF as this project began. On the advise of the SRNF hydrologist, Grouse Creek drainage area boundaries that had been defined by Raines and Kelsey (1991) were transferred to the mylar PBS quads for digitization and thus used to delineate WAAs (Furniss, pers. comm.). Due to this impromptu delineation, it was necessary to assign one fictitious code, "Z," to the sixth field ("NFS Subwatershed") of the HUC. Fields one through five, seven and eight contain correct values.

## Habitat and Watercourse Address Calculations

Detailed descriptions of address attributes and address calculation procedures are tedious and beyond the scope of this article (contact the author for more detailed information). For now, it will be sufficient to briefly describe the two basic address calculation steps that were necessary to prepare for address matching by the GIS.

The desire to represent main and side-channel habitat units in a GIS suggested that address structures of higher resolution than those used by Hemstrom (1989) would be necessary to facilitate more robust and basin-level spatial analysis. One particular challenge to designing address data structures for fish habitat data would be the ability to represent the discontinuous occurrence of side-channels in correct relationship to the continuous main-channel habitat they were adjacent to. For a GIS to accomplish address matching, there must be address attributes in both the spatial data set (the watercourse layer) and the aspatial data set (the habitat data file).

First, address values for each digitized watercourse segment were calculated as cumulative, odd and even-numbered ranges. This approach represented adoption of a 4-number address structure which could be used to model the locations of both main and side-channel habitat units. The digitized version of Grouse Creek consisted of 25 arc segments.

Second, address values for each habitat unit in the habitat database were calculated. The habitat unit address values were slightly different than the cumulative length values because it was necessary to proportionally adjust each unit's address to reconcile the discrepancy between the digitized length of Grouse Creek and its length as measured during habitat inventory.

## Result: Address Matching to Create a Fish Habitat Layer

Once address attributes and addresses were present in both the habitat data file and the digitized watercourse database, a single GIS command was used to address match them, resulting in the creation of an entirely new "derived" data layer that consisted of point marker symbols (ESRI 1990). The GIS read each habitat unit's discrete address value and then found the appropriate watercourse arc segment whose address range bracketed the discrete value. Since all address values were indicative of distance from the mouth of Grouse Creek, unit marker positions reflect their appropriate distance from the mouth of Grouse Creek. By virtue of the odd or even value of addresses (i.e., main or side-channel status), the GIS placed markers to either the left or right side of the Grouse Creek arc (as if facing upstream).

Figures 4, 5, and 6 focus on the same portion of Grouse Creek and were composed using the watercourse layer and the newly derived fish habitat point layer. These figures illustrate some of the cartographic functionality of the fish habitat layer in the GIS.

Figure 4 shows all the fish habitat data plotted to illustrate data density, the effect of the 100-foot offset, and how main and side-channel features are depicted by virtue of their odd or even address integer values. Without the offset, habitat point markers would have been plotted precisely along the Grouse Creek watercourse arc position.

Figure 5 used a lookup table (LUT) to depict all the habitat units as classified into the four basic categories of fish habitat (i.e., cascade, pool, riffle, and run). A LUT is a database file that is used relationally by the GIS to automatically assign various symbology and/or colors to the features in a layer based on the values contained in a specified attribute of the layer's database file. A layer can have a variety of LUTs written for it to facilitate repeatable, consistent, rapid production of various special purpose cartographic products.

Figure 6 depicts a subset of the 868 habitat units in the fish habitat point layer that met the criteria for "quality habitat." The query "select for HAB_CATEG = 'pool' and PRCT_COVER > 39 and PRCT_FINES < 20 and MAX_DEPTH 3" produced the results in Figure 6.

## Discussion

Data sets developed within the context of a 4-number address structure proved to be a suitable approach for deriving absolute positions of main and side-channel habitat features along Grouse Creek. Habitat unit addresses, as composed of an address number, watercourse name, and the HUC, provide a powerful and logical mechanism for establishing the unique identity of any particular habitat unit anywhere in the continental United States. An address data structure lends itself well to facilitating data relationships with the myriad of data layers that will ultimately reside within a typical National Forest GIS, as well as to other agency hydrologic databases.

As other GIS layers are developed (e.g., transportation networks, soils, geologic parent material, existing vegetation, land ownership, et

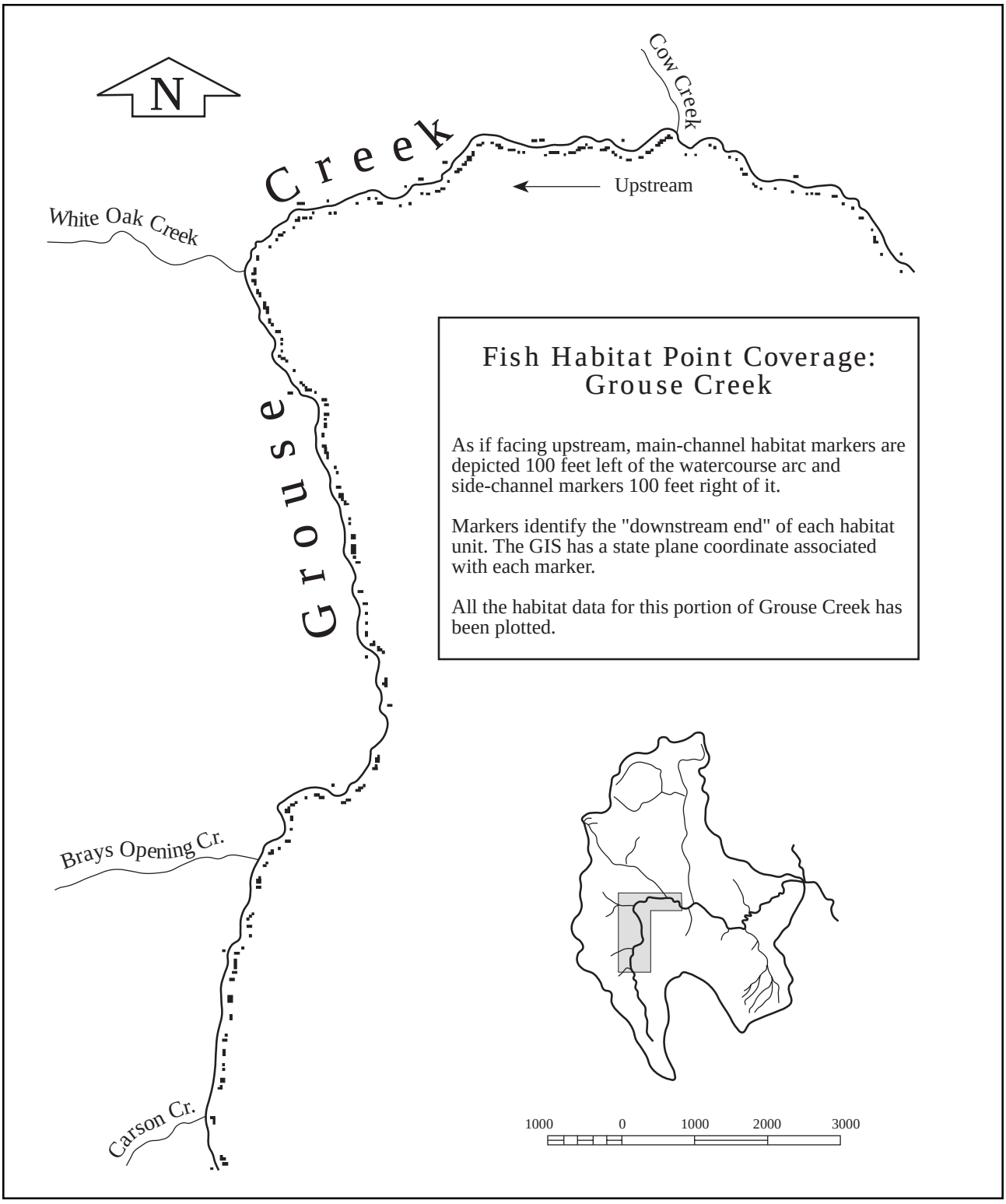



Figure 4. Cartographic results of address matching. All the habitat data has been plotted to either the left or right of the Grouse Creek arc by virtue of the habitat unit address that was calculated for each unit.

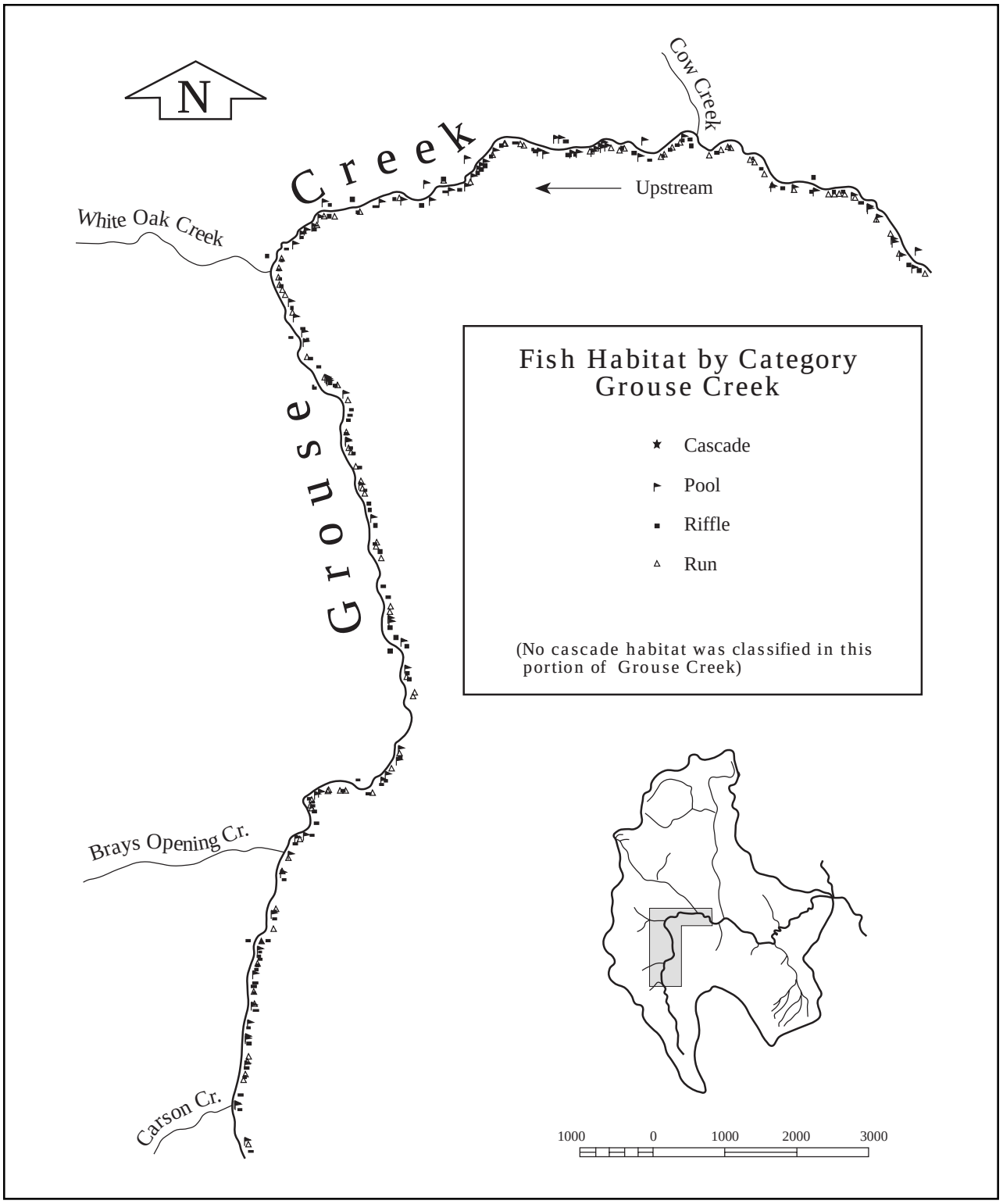



Figure 5. Cartographic results demonstrating use of a lookup table (LUT) to assign various marker symbols toll the habitat units in the fish habitat point coverage, as determined by the "HAB_CATEG" each unit occurs in. Data resolution (e.g. 244 units <40 feet in length), symbol size and plotter output scale occasionally conspire to make it difficult to discern individual units.

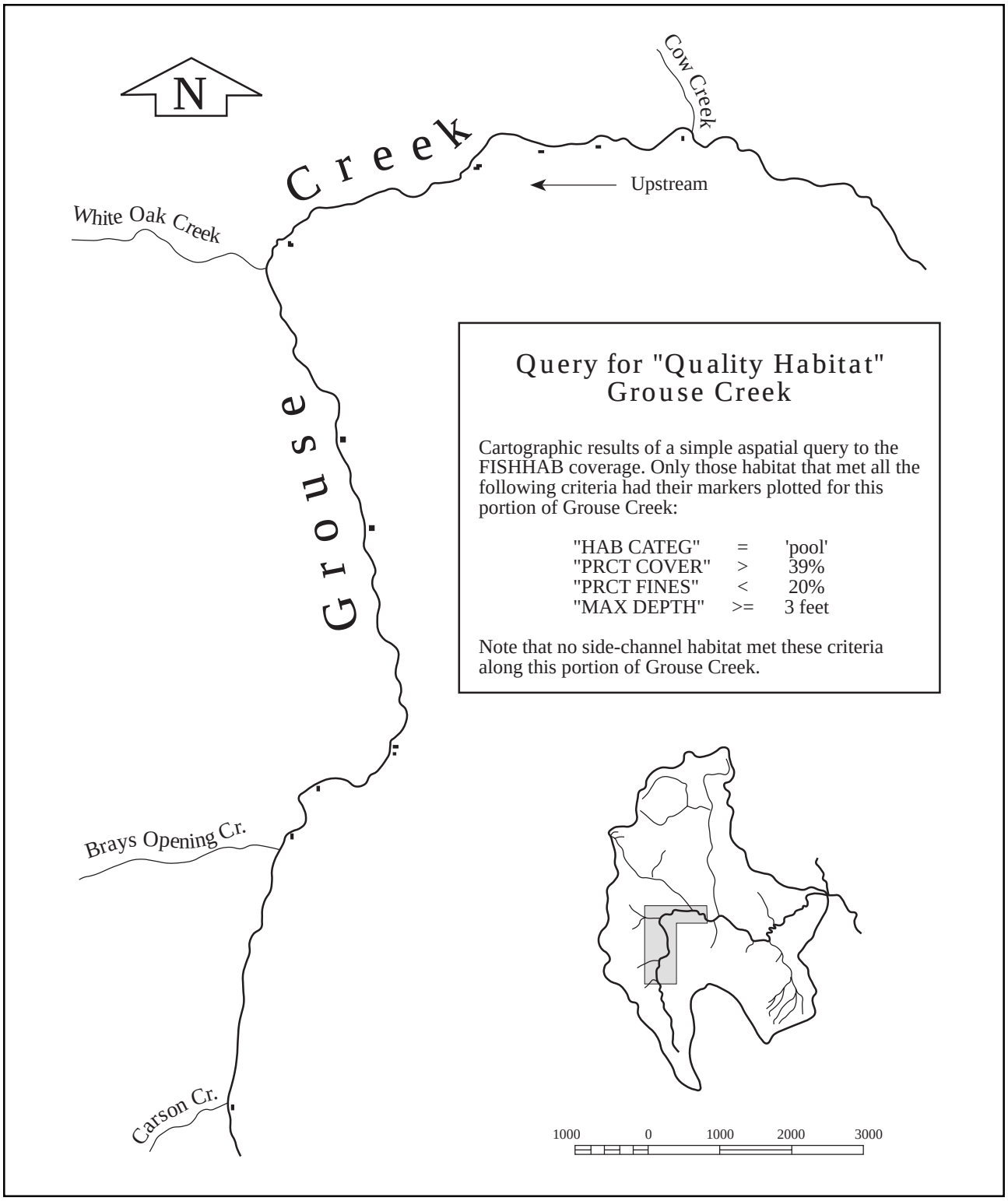



Figure 6. Map composition displaying results of a simple, aspatial query to the fish habitat point coverage for "quality habitat". As other data themes are developed, more powerful analytical tools of the GIS can be invoked to explore relationships between features and processes in the watershed basin and in-stream fish habitat.

cetera), biologists can exploit the full analytical power of the GIS to interactively model and display complex basin-level time/distance cause and effect relationships on in-stream fish habitat conditions and distribution. The watercourse arc network topology necessary for address matching can be applied to modeling fish migration and spawning as well as other hydrologic processes. Regardless of the analysis complexity, by virtue of the four-number address structure, it will always be possible to readily distinguish between main and side-channel habitat.

## Habitat Position Accuracy

For the Grouse Creek GIS project, it is estimated that any given habitat unit position derived by the GIS was represented within 125 feet of its true position. Position quality of future habitat data sets could be easily improved to ±50 feet or better by implementing several revisions to data collection procedures (as presented in the Recommendations section).

Any effort to improve the absolute position accuracy of watercourse features in a GIS must acknowledge that "accurate" stream position data begins to degrade immediately. Processes of aggradation and degradation are continually altering stream channel position and the condition and distribution of fish habitat in it. Storms that produce discharges in excess of bankfull produce significant enough change in habitat structure and distribution to warrant reassessment (Boberg 1991, Trush 1991, pers. comms.). In the coastal northwest, events of this magnitude or greater can occur as frequently as every five years. Among other biological considerations, the costs and methods of achieving spatially accurate fish habitat data sets should be balanced against the frequency and magnitude of data obsolescence.

## Data Standards

As the scope and complexity of environmental analysis continues to broaden and take on bio-regional extents of evaluation, it becomes important that standardized data definitions be implemented in spatial databases to facilitate "vertical" and "horizontal" information sharing and joining. The Forest Service (USFS 6/1/90, 10/22/91, 1989, 1991b, 1992a, 1992b; USFS/R5 1990; USFS/R6 1990, 1991, 1992; USFS/R8 n.d.; USFS/R10 1989) and some states (e.g., California in Flosi and Reynolds 1991) have invested considerable time and resources to the design of "sharable" databases. Within the scope of this project, every effort was made to assemble data sets to known data standards.

## Connectivity with EPA's Reach File 3

Forest Service data standards suggest incorporation of Reach File 3 (RF3) watercourse segment codes to FS hydrologic databases (USFS 1989; USFS/R6 1991). RF3 data sets represent a national hydrologic network database maintained by the EPA to assist in the monitoring of water quality and were developed from 1:100,000 digital line graph (DLG) data. They depict between 75 and 90 percent of the watercourse arcs that are present on 1:24,000 quads (Dulaney 1991; Hansen, pers. comm.). RF3 files for the South Fork Trinity River Cataloging Unit (which contains the Grouse Creek NFS subwatershed) were acquired from EPA (Hansen, Veisze, pers. comms.) to explore the feasibility of matching RF3 segments with watercourse segments derived from 1:24,000 quads. The crossover was easy to discern and it took only several minutes during an interactive arc editing session to assign RF3 segment codes to arcs in the Grouse Creek watercourse layer. With RF3 segment codes attached to each arc in the watercourse layer, along with the presence of the "HUC" attribute, a vast amount of EPA and USGS hydrologic data becomes relationally accessible (e.g., DAMS, IFD, GAGES, NWIS, PCS, STORET, and WBS databases).

## Extended Role and Utility of the HUC

While this project intended to use the HUC to resolve water feature name ambiguities during address matching and access to other databases, it also offers potential utility as a hierarchical key capable of aggregating (or disaggregating) other data sets to various drainage basin extents. In

essence, the HUC provides a natural, physiographic mechanism that can define a "wildland resource zip code zone".

Hydrologic unit boundaries (i.e., ridgetops) are often coincident with soils, geologic, vegetative, political, and administrative boundaries, and in many cases represent the "truest" boundary definition. Arcs from hydrologic unit boundaries will inevitability find their way into many other spatial data sets as the issues of vertical integration and sliver polygons are addressed during development of national forest GIS'. This suggests that delineation and coding of hydrologic units should be a high priority for GIS data development on national forests.

## Role of the Global Positioning System

If the x,y coordinate position of downstream endpoints of habitat units could be acquired during habitat inventory, most of the tedious process of habitat unit address calculation could be foregone, and a more accurate representation of fish habitat could be assembled by the GIS. With the global positioning system (GPS) satellite constellation nearly complete, 24-hour 3D positioning will soon be available almost anywhere in the world. However, in many instances a watercourse thalweg represents some of the most adverse conditions (narrow visible horizons and dense vegetative canopy) in which to successfully operate a GPS receiver. Skillful integration of GPS data with traditional metes and bounds measurement methods to the habitat inventory procedure could easily result in habitat data sets whose absolute position accuracy significantly exceeds the NMAS[3] for 1:24,000 quads (±15 feet verses ±40 feet).

## A Revised Address Matching Paradigm

Using the address matching paradigm developed during this project, insertion, deletion, or updating either arc or habitat unit address data would be a task of modest effort for a stream the length of Grouse Creek. However, for larger river systems (>5 miles in length) address insert, delete, or update operations could become very burdensome. This project calculated addresses based on the total cumulative length of the entire stream -- the "fundamental entity" was the entire length of Grouse Creek. Shifting the paradigm to *watercourse arc segments* as the fundamental entities against which address matching occurs has the potential to eliminate countless hours of address recalculation and facilitates both temporally and spatially discontinuous data collection and correction.

Furthermore, if habitat in a watercourse could have positions predominantly determined using GPS methods (with some metes and bounds methods used to fill in the blind spots) it could become possible to shift the address matching paradigm to an even smaller fundamental entity (i.e., the downstream end-point coordinate of individual habitat units). Indeed, if habitat coordinates could be unmistakably associated with their correct watercourse arc segment, the address matching paradigm could be abandoned completely.

## GIS Enhancement Opportunities

"Dynamic segmentation" features of recently released UNIX-based GIS offer powerful tools that would improve analysis, modelling, and representation of fish habitat (ESRI 1991; Intergraph 1992). Dynamic segmentation enables analysts to "virtually subdivide" a line segment by virtue of attributes that describe that line feature. This eliminates the need to physically subdivide arcs by inserting nodes. Any number of "segmentation tables" can be assembled for any number of line features, to suit any particular analysis scenario. A cartographic benefit of dynamic segmentation is that portions of lines can be represented with different line symbols. As might be applied to fish habitat, units could be depicted as lines of various length rather than as points.

# Recommendations

Many recommendations are interrelated, but are presented as they occur in three broad categories: 1) field methods; 2) topologic improvements; and 3) database considerations. These recommendations apply, regardless of the paradigm used to incorporate fish habitat data to a GIS (address matching or GPS). It is assumed that watercourse arcs are defined from the CFFs with all arcs oriented upstream (i.e., with a network topology).

## Field Methods

1. **Use metric units of measure to describe stream dimensions**. The corporate GIS adopted by the USFS will be based on the de facto standard UTM coordinate grid to take full advantage of DEM and satellite imagery data (Lillis and Kiefer 1987; USGS 1990).

2. **Reference all in-stream measurements to the channel thalweg**. In the interests of consistency and accuracy, measure length parallel, width perpendicular to, and confluences at the intersection of watercourse thalwegs.

3. **Conduct inventories on the basis of common-segments**. Program habitat inventory and population survey work schedules to completion of data collection between pre-established nodes (i.e., common-segment end-points).

4. **Physically mark common-segment nodes on the ground**. With common-segment nodes marked on the ground, field crews will have unmistakable evidence of their location at the critical common-segment end-points.

5. **Reference population surveys and habitat inventory to the same common-segments**. This will assure a strong spatial link between data sets that are gathered at two different times.

6. **Add some cadastral quality positions along inventoried watercourses**. Where PLSS monumentation is sparse along watercourses containing valuable habitat, coordinate with cadastral survey engineers to "densify" monumentation to provide tie points for in-stream inventories.

7. **Use electronic data collection devices to record habitat data**. Much of the USFS Region 5 habitat inventory form could be adapted to bar code input, offering an opportunity to speed inventory and indirectly contribute toward improvement of spatial accuracy.

8. **Revise attribute definitions in FSH 2609.23 (USFS Region 5 Fish Habitat Evaluation Handbook) to reflect FS data standards**. Nearly all of the "new" or revised attributes implemented for this project have standardized definitions that should become part of everyday use in the interests of assembling vertically and horizontally integrated spatial data sets (USFS/R5 1990).

## Topologic Improvements

9. **Delineate watershed boundaries to at least the NFS subwatershed level**. This recommendation cannot be over-emphasized. Watershed boundaries represent such a fundamental and pervasive data element in a corporate GIS that their accurate delineation, on the first effort, warrants priority attention.

10. **Insure that watercourse arcs are all oriented in the same direction**. With watercourse arcs all oriented upstream, two significant benefits are realized: a network topology is created, and data "orientation" matches that of other agency databases. Network topologies enable analysts to model movement of objects through a route system.

11. **Begin densification of the tic registration grid**. Tic registration coordinates are what a GIS uses to vertically align layers of data. In a wildland setting a tic grid based only on 7.5 minute quad corners is insufficient for conduct of project level spatial data collection. Identification of PLSS monuments of acceptable position quality for inclusion to a forest's master tic file would help improve data quality.

### Database Considerations

12. **Eliminate the "memo" data type from the fish habitat database**. While very convenient, a "COMMENTS" attribute can wreak havoc in a true relational database environment. Diligent data normalization would identify additional attributes and relational data tables that could be used to eliminate the ambiguous "COMMENTS" attribute.

13. **Transfer fish habitat data files to the Oracle RDBMS**. Oracle provides a very powerful database environment that will most probably support the GIS procured by the USFS. It also offers powerful database design, integrity control, and normalization tools that are unavailable in dBASE.

14. **Adopt the habitat address coding scheme as the data standard for unique identity of fish habitat**. Regardless of the paradigm used, the "number," "street name," and "zone" attribute set provides a powerful compound key that is capable of providing unique identity to any fish habitat unit in the United States.

15. **Assign unique watercourse arc segment codes to all arcs on the PBS**. Within the extent of a cataloging unit, every watercourse arc segment in the CFF needs a unique ID number assigned to it. Once hydrologic unit areas have been delineated, the GIS could be used to accomplish this task.

16. **Seek definition of a "corporate" registration tic coding scheme**. As analysis of environmental issues expands to larger, bio-regional extents, consistently and accurately registering a wide range of data from a variety of sources becomes a serious issue. A coding scheme that provides for unique identity of tics is needed.

## Acknowledgments

The author feels very fortunate to have had the opportunity this project has offered and wishes to express his appreciation and thanks to the following people for their support and encouragement: Jerry Boberg, Jim Davis, Gene Graber, Kathy Heffner, Jack Kahl, John Larson, and Leah Stewart of the Six Rivers National Forest; Steve Carlson and Larry Fox of Humboldt State University and Roger Barnhart, Tom Hassler, and Delores Neher of the California Cooperative Fishery Research Unit, U.S. Fish and Wildlife Service.

## Footnotes

[1]Project 615 is the title of the pending national implementation by the Forest Service of a fully integrated "corporate" digital information management system. One component of this information management system will be a full-featured GIS. A corporate information management system implies use of common data definitions and structures throughout the organization to enable vertical and horizontal sharing of information (Date 1990; Elmasri and Navathe 1989).

[2]CFFs are digital versions of the PBS prepared to National Map Accuracy Standards by the Forest Service Geometronics Service Center in Salt Lake City. CFFs eliminate error-prone, tedious hand digitizing procedures and are already edgematched with line-work on adjacent quads (Holland 1991; USFS 1991a).

[3]Relative to horizontal accuracy of maps with published scales of 1:24,000, the NMAS *statistically* defines "economic and expeditious" map accuracy as no more than 10 percent of the "well-defined points" on a map sheet being in error by more than 0.02 inches of their absolute position. "Well defined points" are those that are easily recoverable on the ground and generally plottable on the scale of map being produced, to within 0.01 inches (ACSM and ASCE 1978; Holland 1991; Muehrcke 1986).

# Personal Communications

Boberg, Jerry. 1990-92. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Fuller, David. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Furniss, Mike. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Hansen, Olof. 1992. U.S. Environmental Protection Agency, Region IX, 75 Hawthorne Street, San Francisco, CA 94105.

Hemstrom, Miles. 1991. USFS, Rocky Mountain Region, 11177 W. 8th Ave., Lakewood, CO 80225.

Kenfield, Karen. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Kiser, Jim. 1991. Oregon State University, College of Forestry, Corvallis, OR 97331.

Lienkaemper, George W. 1991. USFS, Pacific Northwest Forest and Range Experiment Station, Forestry Sciences Lab, 3200 SW Jefferson Way, Corvallis, OR 97331.

McRae, John. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Morrison, Sam. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Ober, Nan. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Smith, Mark. 1991. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Steinblums, Ivars. 1991. USFS, Siskiyou NF, Box 440, Grants Pass, OR 97526.

Stewart, Leah. 1990-92. USFS, Six Rivers NF, 1330 Bayshore Way, Eureka, CA 95501.

Trush, William. 1991. Fisheries Department, College of Natural Resources, Humboldt State University, Arcata, CA 95221.

Veisze, Paul. 1992. State of California, Department of Fish and Game, Inland Fisheries Division, 1416 9th Street, Sacramento, CA 95814.

# References

ACSM and ASCE (American Congress on Surveying and Mapping and the American Society of Civil Engineers). 1978. Definitions of Surveying and Associated Terms. American Congress on Surveying and Mapping, Bethesda, Maryland.

Date, C.J. 1990. An Introduction to Database Systems, 5th Edition. Addison-Wesley Publishing Co., Inc., Reading, Massachusetts.

Dulaney, Richard A. 1991. The U.S. EPA River Reach File 3: A National Hydrographic Database for GIS Analyses. Pages 201-210 *in* Proceedings of the Eleventh Annual ESRI User Conference, Volume 1. Environmental Systems Research Institute, Inc., Redlands, California.

Elmasri, Ramez and S.B. Navathe. 1989. Fundamentals of Database Systems. Benjamin/Cummings Publishing Co., Redwood City, California.

ESRI (Environmental Systems Research Institute, Inc.). 1990. pcNETWORK Users Guide. Environmental Systems Research Institute, Inc. Redlands, California.

ESRI (Environmental Systems Research Institute, Inc.). 1991. Dynamic Segmentation. Pages 81-98 *in* Workshop Proceedings of the Eleventh Annual ESRI User Conference. Environmental Systems Research Institute, Inc. Redlands, California.

Flosi, Gary, and F.L. Reynolds. 1991. California Salmonid Stream Habitat Restoration Manual. California Department of Fish and Game, Inland Fisheries Division, Sacramento, California.

Hemstrom, Miles. 1989. Integration of Riparian Data in a Geographic Information System. Pages 17-22 *in* Practical Approaches to Riparian Resource Management: An Ecological Workshop. American Fisheries Society, Bethesda, Maryland.

Holland, Carolyn K. 1991. Cartographic Feature Files: A Synopsis for the User. United States Forest Service, Engineering Staff, EM7140-21. Washington, DC.

Intergraph Corporation. 1992. MGE Segment Manager. (A product specification publication). Intergraph Corporation, Huntsville, Alabama.

Lillisand, Thomas M. and R.W. Kiefer. 1987. Remote Sensing and Image Interpretation. John Wiley and Sons, New York, New York.

McCain, M., D. Fuller, L. Decker, and K. Overton. 1990. Stream Habitat Classification and Inventory Procedures for Northern California. Pages 1-15 in Region 5 *FHR Currents* Technical Bulletin 1. United States Forest Service, Pacific Southwest Region, San Francisco, California.

Muehrcke, Phillip C. 1986. Map Use: Reading, Analysis, and Interpretation. JP Publications, Madison, Wisconsin.

Raines, Mary A. and H.M. Kelsey. 1991. Sediment Budget for the Grouse Creek Basin, Humboldt County, California. Master Thesis. Western Washington University, Bellingham, Washington.

Seaber, Paul R., F.P. Kapinos, G.L. Knapp. 1987. Hydrologic Unit Maps. United States Geological Survey, Water-Supply Paper 2294. Washington, DC.

USFS (United States Forest Service) 6/1/90. Watershed and Air Management, Watershed Planning. Forest Service Manual, Title 2500, Chapter 2510, Washington, D.C.

USFS (United States Forest Service). 10/22/91. Wildlife, Fish and Sensitive Plant Habitat Management. Forest Service Manual, Title 2600, Washington, DC.

USFS (United States Forest Service). 1989. Interim Resource Inventory Glossary. United States Forest Service, Washington, DC.

USFS (United States Forest Service). 1991a. Geometronics Service Center. United States Forest Service, Engineering Staff Poster EM-7140-22, Geometronics Service Center, Salt Lake City, Utah.

USFS (United States Forest Service). 1991b. National Geographic Information Structure. United States Forest Service, Computer Sciences and Telecommunications, Washington, DC.

USFS (United States Forest Service). 1992a. GIS Evaluation Report. United States Forest Service, Computer Sciences and Telecommunications, Washington, DC.

USFS (United States Forest Service). 1992b. Information Management: A Framework for the Future. United States Forest Service, Washington, DC.

USFS/R-5 (United States Forest Service, Region 5). 1990. DRAFT Fish Habitat Assessment Handbook. Forest Service Handbook, Chapter 2609.23, United States Forest Service, Pacific Southwest Region, San Francisco, California.

USFS/R-6 (United States Forest Service, Region 6). 1990. Chapter 2 Stream Inventory Handbook, version 4.0. United States Forest Service, Pacific Northwest Region, Portland, Oregon.

USFS/R-6 (United States Forest Service, Region 6). 1991. Standard Terms and Definitions for Region 6 Land Based Resource Information. United States Forest Service, Pacific Northwest Region, Portland, Oregon.

USFS/R-6 (United States Forest Service, Region 6). 1992. Guidelines for Map Manuscript Preparation for Digitizing. United States Forest Service, Pacific Northwest Region, Portland, Oregon.

USFS/R-8 (United States Forest Service, Region 8). n.d. Digital Data Base Development Issues and Approaches, Version 2. United States Forest Service, Southeast Region, Atlanta, Georgia.

USFS/R-8 (United States Forest Service, Region 8). 1989. Geographic Information System Implementation Plan, Region 8. C.W. Dull, Team Leader. Internal document, United States Forest Service, Southern Region, Atlanta, Georgia.

USFS/R-10 (United States Forest Service, Region 10). 1989. Resource Information Management Data Dictionary: Alaska Region. United States Forest Service, Alaska Region, Juneau, Alaska.

USGS (United States Geological Survey). 1990. Digital Elevation Models: Data Users Guide 5. United States Geological Survey, National Cartographic Information Center, Reston, Virginia.

**Recycled/Recyclable Paper**