59

Coryell A. Ohlander
6048 So. Lakeview Street
Littleton CO 80120
cohlander@polnow.net

April 9, 2000

To: Content Analysis Enterprise Team
USDA Forest Service
Bldg. 2 Suite 295
5500 W. Amelia Earhart Dr.
Salt Lake City, UT 84116 (cleanwater/wo_caet-slc@fs.fed.us)

FROM: Coryell A. Ohlander, hydrologist

RE: Good Science aspects in UFP
Part 3 regarding limiting factors

Goals "... in a unified and cost-effective manner."
Guiding principles: "...consistent and scientific approach...."
Agency Objectives: "...common science-based approach...."
"...test watershed assessment procedures...."
"...implement ... consistent with applicable legal authorities."
"...will base watershed management on good science."
"...science-based total maximum daily loads (TMDLs)."
"...sharing of scientific and technical resources;"
"...monitoring and evaluation...."

Summary Recommendation

In relationship to 6 stream health classes, it is essential that the FS/BLM work with metrics that define the habitat dimension for diversity, stability, and productivity as it comes from the law itself. Since many land use activities primarily affect habitat, the need is to routinely evaluate habitat characteristics that support biological integrity. Within the CWAP national focus, I recommend that UFP draw from S404 regulations (40 CFR 230) as a minimum standard for the habitat elements in aquatic assessments, even if S404 permits are not required.

There are three primary advantages: first, it is well organized with a depth of understanding carried over from the Water Resources Council procedures developed for federal water projects; second, it ties directly to the formal status given to the USF&WS for aquatic and wildlife assessment tools; and three, it exists in usable form NOW. All it would take is a UFP decision.

USF&WS Habitat Evaluation Procedures (HEP) and a subset called Habitat Suitability Index (HSI) models provide a ratio scale (0-1) based on carrying capacity and factors thought to be biologically limiting. A practitioner chooses the appropriate fish or wildlife HSI model, evaluates project changes on the factors, and interprets health based on the most limiting factor(s). Limiting factors used in this way avoids interpretation problems associated with multi-factor ordinal indexes and ordinary least-squares regressions.

Support

1) I have enclosed a paper by Jim Terrell and friends entitled "*Modeling stream fish habitat limitations from wedge-shaped patterns of variation in standing stock.*" Jim works for the National Biological Service, Mid-continent Ecological Science Center, which is now part of U.S. Geological Survey. At the time the paper was published, he was Supervisory Fish and Wildlife Biologist: modeling fish and wildlife habitat relationships with major responsibility for HSI. One value of the paper is that it provides an excellent rationale for the analysis of standing crop and habitat data as used in HSI. The paper challenges the use of ordinary least-squares as an indication of limiting factors.

In terms of good science, the analyses of wedge-shaped patterns for habitat relationships provides a strong logic that supports the HSI rationale. In addition, USF&WS is formally charged by law (16 USC 661 et seq.) with providing detailed guidance regarding pollution effects on fish and wildlife resources. As part of that mandate, as well as responsibility to identify and measure impacts from Federal water resource projects, the USF&WS started developing a comprehensive set of habitat models in the early 1980's. At last count, there are 156 models for fish and wildlife species that apply nationwide. The models have been converted to computer application and are available through the USGS web site. There would be no UFP downtime.

In terms of legal standing, the most detailed criteria for evaluation of biological resource impact from pollution are those associated with CERCLA (43 CFR 11). While CERCLA does not limit biological assessments to any particular method, one advantage to the use of HEP is that it is already listed as acceptable under CERCLA (43 CFR 11.71(l)) and carries similar clout under the 1990 Oil Pollution Act (OPA). OPA further allows "Habitat Equivalency Analysis" to establish an economic cost for habitat replacement (15 CFR 990.27) and does not depend on measuring population characteristics. This argument might help UFP get around the huge costs associated with population studies as a measure of biological diversity and productivity (some of the TMDL discussions re creditable evidence, i.e. Wyoming, have taken this direction).

Based on a sampling of participants in the July 1999 T-Walk review, the FS has never used the HEP argument in connection with the TMDL issue. Even though the feds are faced with greater and greater water quality obligations and a continuing loss in technical people, there remains steadfast refusal to use existing tools. However, TMDL timetables do not allow time to continue fiddling around. It would be foolish for UFP to ignore the HEP storehouse especially in light of the fact that the TMDL issue will be resolved with or without the feds. It really is time to put aside internal discipline bickering and get on with the real job of watershed condition and aquatic health.

One additional nagging problem is the selection of reference condition. Reference conditions are needed from which to make a comparison and assign level of health. They are needed to set TMDL limits, define impairment, and provide a model of restoration. Take a moment to really think about what HSI models offer in terms of defining reference conditions. In areas that have few unimpacted reaches, and if the desired condition can be expressed in terms of species, then the reference condition can be inferred from the HSI models themselves. This is possible because the suitability index is a function of optimum carrying capacity which properly reflects the content of attainable designated aquatic uses. After selecting desired resident species, the HSIs would bring a description of the optimal conditions that could be used as reference for attainability. The factors with the lowest HSI would be those most in need of restoration.

2) Stream Health Classes and Habitat Suitability Indexes.

Pursuant to CWA S304, the Water Quality Standards Handbook includes major direction for Water Body Survey and Assessment Guidance for Conducting Use Attainability Analyses (WBA). EPA (WBA) recognizes Habitat Suitability models: stating that HSI models provide a basic understanding of species habitat requirements, have utility and applicability to use attainability analyses, and have been developed and published after review by species experts.

The HSI graph with a 0-1 suitability index (y-axis) is common to all factors and to all HEP models. The x-axis varies according to factors being evaluated. For example, rainbow trout HSI for dissolved oxygen has a line relationship between 4 mg/l at "0," 9 mg/l at "1," and 6.3 mg/l at "0.5". Because HSI suitability indexes are constructed as carrying capacity ratios, the proper interpretation is that 0.5 will carry only half of the potential at 1.0. This curve is based on literature that indicates that dissolved oxygen is a limiting factor. That the 0-1 HSI suitability index is a carrying capacity scale is a basic requirement for HEP. Among the different disciplines, there is no universally accepted definition of carrying capacity; therefore, the context of each HSI model has to be described in a way that has meaning for specific applications. Carrying capacity, as used for HSI, denotes a broad concept or theoretical expression of the population limit with respect to the habitat resources used in the model.

HEP and HSI models do not use narrative definitions that translate particular ranges of suitability into classes. But classes are necessary to respond to both policy issues as well as reporting on stream health for CWA 305 or 319 watershed assessments. Following a detailed review of three evaluation systems that use a 0 to 1 scale, the conclusion is that stream health classes and HSI match up as follows:

Robust would be	set at 1.0	on the	Suitability	Index	scale.
Adequate	0.8 - <1.0	"	"	"	"
Diminished	0.6 - <0.8	"	"	"	"
Impaired	0.4 - <0.6	"	"	"	"
Precarious	0.2 - <0.4	"	"	"	"
Catastrophic	<0.2	"	"	"	"

(The most interesting is the Environmental Evaluation System (EES), drafted by Battelle Columbus Laboratories for the Bureau of Reclamation in 1972. The conceptual framework is based on NEPA responsibilities triggered by large scale water resource development and is, by far, the most comprehensive and most internally consistent of the evaluation systems so far encountered. EES supports 4 groups and 18 components: ecology (species & populations, habitats & communities, ecosystems); environmental pollution (water, air, land, noise); esthetics (land, air, water, biota, man-made objects, composition); and human interest (educational/scientific, historical interest, cultures, mood & atmosphere, life patterns). EES also deals with large ecosystems explicitly. Page 109 provides a comprehensive procedure for indexing environmental effects from complex projects. Page 113 on ecosystems; page 164 on spatial and temporal concerns. EES provides a structure for "red flags," fragile elements, and potentially sensitive environmental situations that lack adequate data to make a reasoned judgement. The red flags also serve as limiting factors. Ecological red flags were considered major if the negative change was 10% or more. For non-ecological major red flags, the change had to exceed 0.1 EQ unit and associated parameter change greater than 30%.)

3) Since several States have started to move toward ecosystems for their own CWA evaluations, I previously recommended that UFP adapt EPA's stream health structure to the riparian corridor to encompass both vegetation and wildlife as a means of anticipating the trend toward ecosystem based State programs. S404 already provides a well defined structure for watershed level cumulative effects as well as a public interest review that encompasses the entire watershed.

Except for CWA S404 Permits for Dredged or Fill Material, the guidelines are not required; they are, however, well written and provide a solid foundation for evaluating aquatic ecosystems in general. There are enough land management situations, including watershed condition and risk assessments, and water related special use permits, to make the effort at standardizing basic aquatic evaluation techniques cost-effective and worthwhile. Some of the rationale used in T-Walk includes:

1) Many project related decisions involving transportation planning, road construction and maintenance just barely skirt S404 involvement by the COE, while other projects require direct COE and EPA S404 approval. The most stringent requirements and best management practices for roads are found in the S404 regulations. Whether or not forest roads are exempt from the S404 permitting process depends on construction and maintenance practices that meet the "no impairment" criteria. The question of impairment is tested with these guidelines. It is not tested with the either NEPA or NFMA regulations.

2) The guidelines are used by the COE and EPA in their S404 evaluation process of water related projects and either agency may take control of the decision itself if S404 is involved. For example, if the plan and supporting environmental assessment for a water resource project does not satisfy the guidelines, then S404 permits are denied. Adopting the guidelines as a foundation, then adding to it where other laws or regulations add emphasis should result in fewer chances of lacking necessary information and miscommunication.

3) The S404(b)(1) guidelines are valuable as a protocol for both NEPA determinations of significant impacts on aquatic ecosystems and as a measure of NFMA's restriction of adverse sedimentation impacts on fish habitat and channel conditions. The guidelines also provide a logical framework to document permit conditions, mitigation, monitoring, or enforcement as might be specified in the decision.

4) Review of hundreds of EA's and EIS's suggests the need for a more rapid, yet more legally complete, evaluation process for aquatic systems. The widely variable formats encountered in R-2 suggests that for each new analysis, the tendency has been to start over and rehash choices about formats, what data to collect, what to analyze, how it should be displayed, etc. The routine use of these guidelines would provide a more efficient data and analysis structure for both the report and supporting evaluation and, even without contributions from other disciplines, at least cover the requirements for the aquatic ecosystems and pertinent requirements of CWA.

4) To the extent practical and reasonable, the scope of a S404 Factual Determination is structured to determine potential short-term, long-term, and cumulative effects of proposed activities and related pollution on the physical, chemical, and biological components of the aquatic environment. Both primary and secondary cumulative effects are included (40cfr230.11(g & h)).

The analysis includes the character of the aquatic environment as well as the pollution process and availability of pollutants. The factors would include current patterns, water circulation, wind and wave action; water depth, flow velocity, direction, and variability; turbulence and stratification; water column physical and chemical properties; and changes in the hydrologic regime including diversions and dams (flow obstructions). As related to pollution, the specifics would include information on technology and field methods; discharge location, volumes, rates, and duration; grain size and expected chemistry of the discharge material; settling velocities; the shape, size, and water chemistry of the drift plume. The combination of factors are more or less structured into these 5 general topics:

Physical substrate -- nature and degree of effect, including duration and physical extent, on the characteristics of the substrate, such as similarity in particle size, shape, and degree of compaction of the discharge material compared to the substrate at the disposal site, and any potential changes in substrate elevation and bottom contours, including changes outside of the disposal site (off-site damage) which may occur as a result of erosion, slumpage, or other movement of the discharged material.

Water circulation, fluctuation, and salinity -- nature and degree of effect on water, current patterns, circulation including downstream flows, and normal water fluctuation; and on water chemistry, salinity, clarity, color, odor, taste, dissolved gas levels, temperature, nutrients, and eutrophication plus other appropriate characteristics.

Suspended particulate/turbidity -- nature and degree of effect in terms of changes in the kinds and concentrations of suspended particulate or turbidity.

Contaminants -- degree to which the discharge material will introduce, relocate, or increase contaminants. Compare expected results with applicable state water quality standards including beneficial use classification.

Aquatic ecosystem and organisms -- nature and degree of effect that the discharge material will have on the structure and function of the aquatic ecosystem and on organisms. This is designed to account for changes in substrate, water or substrate chemistry, nutrients, currents, circulation, fluctuation, and salinity, on the recolonization and existence of indigenous aquatic organisms or communities.

5) In terms of detail, an aquatic resource impact analysis can start from 40 CFR 230 Subparts C-G offered as main headings:

Substrate. (40cfr230.20)
Suspended particulates/turbidity. (40cfr230.21)
Water chemistry. (40cfr230.22)
Current patterns and water circulation (40cfr230.23)
Normal water level fluctuations. (40cfr230.24)
Salinity gradients. (40cfr230.25)
Threatened and endangered species. (40cfr230.30)
Aquatic Food Web (40cfr230.31 Fish, crustaceans, mollusks, and other food web organisms)
Wildlife (40cfr230.32)
Sanctuaries and refuges. (40cfr230.40)
Wetlands. (40cfr230.41)
Mud flats. (40cfr230.42)
Vegetated Shallows. (40cfr230.43)
Riffle and pool complexes (40cfr230.45)
Municipal and private water supply (40cfr230.50)
Recreational and commercial fisheries. (40cfr230.51)
Water-related recreation. (40cfr230.52)
Aesthetics (40cfr230.53).
Parks and Preserves (40cfr230.54 Parks, national and historical monuments, national seashores, wilderness areas, research sites, and similar preserves).
Chemical contamination (40cfr230.60 & 61 General evaluation)

6) Subpart H (40 CFR 230.70-.77) Actions Taken To Minimize Adverse Effects

S404 also provides a structure that can be used in the TMDL context and as a basis for restoration planning. For example, projects could identify directly which of the following ways listed in Subpart H have been selected for the project:

Actions relating to location:
a) Discharge is located and confined to a minimum area.
b) Periodic water inundation patterns are not disrupted.
c) Disposal sites have been used before.
d) Disposal site substrates are similar to discharge material (i.e. sand on sand or mud on mud).
e) Disposal site, discharge point, and the method of discharge will minimize the drift plume.
f) Disposal sites have been designed to prevent the creation of standing water in areas of normally fluctuating water levels.
g) Activities have been designed to prevent/minimize draining areas subject to fluctuating water levels.

Actions relating to discharged material:

a) Operations will retain existing physiochemical conditions and reduce the potency and availability of pollutants.
b) Limits on solid, liquid, and gaseous components of discharge material will be applied to each disposal site.
c) Treatment substances will be added to the discharge material.
d) Chemical flocculent will be added to ensure that suspended particulates are controlled.

Actions relating to control of material after discharge:

a) Erosion and minor slumping of materials into the surrounding aquatic ecosystem will be controlled by containment levees, sediment basins, and revegetation.
b) Leaching of materials into the surrounding aquatic ecosystem will be controlled by lined containment areas.
c) Contaminated material will be capped in-place with clean material.
d) Contaminated material will be capped in-place by selectively placing the most contaminated material down first and capped with lesser contaminated material.
e) Project has been designed and discharged material placed to reduce long term vulnerability from natural processes that generate point and nonpoint pollution.
f) Activities will be suspended during climatic or hydrologic events when damage can not be prevented. This will include wet weather or high water flows.

Actions relating to method of dispersion:

a) Discharge material will be distributed widely in a thin layer at the disposal site to maintain environmentally sensitive natural substrate contours and elevation.
b) Deposits will be oriented to prevent/minimize undesirable obstruction to the water flow and patterns.
c) Deposits will be located to fill in natural depressions to minimize the size of the mound.
d) Silt screens/other appropriate methods will be used to confine suspended particulate/turbidity to a small area where settling/removal can occur.
e) Water column turbidity increases will be minimized by using a submerged diffuser or pipeline to release materials near the bottom.
f) Discharges will be confined to minimize the release of suspended particulates over a larger area.
g) Design includes limits on amount of material released per unit of time.
h) Design includes limits on amount of material released per unit of volume of receiving water.

Actions related to technology:

a) Appropriate equipment and related protective devices have been selected for the activity, operations, and site.
b) Appropriate maintenance and operation of such equipment has been ensured by adequate training, staffing, and working procedures.
c) Equipment and techniques especially designed to reduce damage to wetlands will be used. This includes machines equipped with devices that scatter rather than mound excavated materials, machines with specially designed wheels or tracks, and the use of mats under heavy machines to reduce wetland surface compaction and rutting.
d) Access roads and channel spanning structures, including culverts, open channels, and diversions, will provide for fluctuating water levels, pass low and high water, and maintain circulation and faunal movement.
e) Appropriate equipment will be used to transport material.

Actions related to plant and animal populations:

a) Project does not modify water current and circulation patterns which would interfere with the movement of animals.

b) Project does not create habitat conducive to the development of undesirable predators or species which have a competitive edge ecologically over indigenous plants or animals.
c) Project avoids sites with unique habitat or other value, including habitat of threatened or endangered or sensitive species.
d) Project provides habitat development and restoration to produce a new or modified environmental state of higher ecological value by displacement of some or all of the existing environmental characteristics.
e) Project provides habitat development and restoration to minimize adverse impacts and to compensate for destroyed habitat.
f) Project uses techniques that have been demonstrated to be effective in circumstances similar to those under consideration.
g) Project uses new and untested development and restoration techniques; such techniques will be used on a small scale to allow corrective action if unanticipated adverse impacts occur.
h) Activities that create unpreventable sedimentation or turbidity increases will be scheduled to avoid spawning or migration seasons and other biologically critical time periods.
i) Project avoids the destruction of remnant natural sites.

Actions related to human use:
a) Activity and operations prevent/minimize damage to the aesthetically pleasing features of the aquatic site and visual aspects of water quality.
b) Project avoids the use of valuable natural aquatic areas.
c) Activity and operations avoid the seasons or periods when human recreational activity associated with the aquatic site is most important.
d) Activity and operations use procedures which avoid or minimize the disturbance of aesthetic features of an aquatic site or ecosystem.
e) Activity and operations are not detrimental nor will the project increase incompatible human activity, or require the need for frequent dredge or fill maintenance activity in remote fish and wildlife areas.
f) Disposal sites are not located near public water supply intake.

Other related actions:
a) Runoff and related sediment and other pollutants from activities and operations are controlled.
b) Water releases from dams are designed to accommodate the needs of fish and wildlife.
c) Desired water quality of return discharge from a dredge operation has been maintained and meets scientifically defensible pollutant concentration levels and any applicable water quality standards.

7) I made mention of the fact that much of the forest transportation system is actually determined in relationship to S404. Quite apart of the State BMP exercise, there are mandatory requirements for roads listed at 33 CFR 323.4(a)(6). The S404 exemption for roads is the **only** place in the law where a standard is actually written - and it is a no-impairment standard. The following list are mandatory (paraphrased):

1) limit road and trail system to minimum feasible number, width, and total length consistent with the specific operations, topographic, and climate;
2) except at crossings, all roads shall be located sufficiently far from streams or other water bodies to minimize discharges;
3) crossings shall not restrict the passage of expected flood flows;
4) fills shall be stabilized during and after construction to prevent erosion;
5) minimize equipment disturbance in "waters" outside construction zone;

6) minimize vegetative disturbance in "waters" during and after construction;
7) road crossings shall not disrupt the movement of resident aquatic species;
8) take borrow material from upland sources whenever feasible;
9) the discharge shall not take, or jeopardize the continued existence of, a T&E species, or adversely modify or destroy critical habitat;
10) avoid discharges into migratory waterfowl habitat, spawning areas, and special aquatic sites;
11) discharge shall avoid areas in or near public water supply intake;
12) discharge shall avoid areas of concentrated shell fish production;
13) discharge shall avoid National Wild and Scenic River System reaches;
14) discharge material will be free from toxic pollutants in toxic amounts;
15) all temporary fills will be removed and restored to original elevation.

----

8) Similarly, the Public interest review (33cfr 320.4) would serve as an immediate structure for environmental assessment as required by NEPA. These public policy issues are also part of that review:

Aesthetics
Conservation
Economics - national and regional
Energy conservation and development
Erosion and deposition
Effects on Fish and wildlife
Flood hazards
Floodplain management
Food and fiber production
General environmental benefits
Effects on historical, cultural, scenic, recreation
Land use; other Federal, state, local requirements
Mineral needs
Navigation
Effects on property ownership
Recreation
Safety of impoundment structures
Special aquatic sites
Transportation system
Effects on water quality
Effects on water supply and conservation
Effects on wetlands

End of Part 3 letter.

With regards

Coryell A. Ohlander

Coryell A. Ohlander
Scott Taylor Trust, trustee

Comment dated 4/9/00

59

Transactions of the American Fisheries Society 125:104–117, 1996
American Fisheries Society 1996

Modeling Stream Fish Habitat Limitations from Wedge-Shaped Patterns of Variation in Standing Stock

JAMES W. TERRELL, BRIAN S. CADE, JEANETTE CARPENTER, AND JAY M. THOMPSON[1]

National Biological Service, Midcontinent Ecological Science Center 4512 McMurry Avenue, Fort Collins, Colorado, 80525, USA

Abstract.—A wedge-shaped pattern of variation in stream fish standing stock estimates relative to a habitat variable, in which range of standing stocks increases as a function of the variable, is consistent with the concept that the habitat variable is a limiting factor for fish populations. This pattern of variation complicates interpretation of parameter estimates and significance of ordinary least-squares (OLS) regression models of conditional mean standing stock; slopes of these regression models may have little or no relation to slopes of models describing standing stock limits. We modeled standing stock limits by testing for homoscedastic error distributions, screening plots of coordinate pairs for evidence of a wedge-shaped pattern of data, and estimating 90th regression quantiles for simple linear models. Application of this technique to data sets supporting 35 previously published OLS regression models of stream fish standing stocks led to rejection of homoscedasticity ($P < 0.10$) in 13 of the 35 data sets. Eight of these heteroscedastic data sets had wedge-shaped patterns of variation in standing stock and slopes of 90th regression quantiles that differed from slopes of OLS regression models. For three of these eight data sets, tests rejecting homoscedasticity were more significant than tests rejecting zero slope parameters in OLS regression models. In a separate exercise, analysis of simulated standing stock data generated from known distributions indicated that our technique can detect heteroscedastic error distribution patterns and yield 90th regression quantile models of standing stock limits from data sets characterized by OLS regression as having no correlation between mean standing stock and a habitat variable. Identification of correlations between habitat variables and standing stock by OLS regression is a common method of determining whether a variable is to be used for habitat assessment. Application of our technique to data sets that display wedge-shaped patterns of variation should help identify variables that may be limiting standing stock from data sets that do not yield significant OLS regression models of mean standing stock.

Numerous ordinary least-squares (OLS) regression models have been published that predict mean standing stock of fishes in streams from measures of macrohabitat variables such as stream width, average velocity, substrate composition, and total dissolved solids. Fausch et al. (1988) reviewed 99 of these models and found that for small data sets (especially those with less than 20 df) the models seemed to be precise predictors of mean standing stock and had coefficients of determination (r^2) that were greater than 0.75. However, for larger data sets, r^2 values were usually less than 0.75. They concluded that relatively precise models (those with a high r^2) lacked the generality (i.e., were not derived from a large sample size) required for efficient application to a wide variety of habitat conditions. Models published in *Transactions of the American Fisheries Society* and the *North American Journal of Fisheries Management* subsequent to the time period reviewed by Fausch et al. (1988) indicate a continuing trend for greater model precision to be associated with smaller data sets. All of the significant ($P < 0.05$) OLS models described by Layher and Maughan (1985), Lanka et al. (1987), Layher and Maughan (1987a, 1987b), McClendon and Rabeni (1987), Pajak and Neves (1987), Wesche et al. (1987), and Hubert and Rahel (1989) that had r^2 greater than 0.60 had 22 df or fewer.

Bowlby and Roff (1986) noted that regression analysis is an objective approach for identifying limiting habitat factors in streams. However, the application of regression to stream macrohabitat data is producing few, if any, models based on large or geographically diverse data sets. Authors commonly recommend limiting the geographic area of model usage and repeating the model building effort for different geographic areas (e.g., McClendon and Rabeni 1987; Wesche et al. 1987). Perhaps the strongest warning on model usage was presented by Layher et al. (1987), who concluded that precise and transferable models of mean standing stock for spotted bass *Micropterus punc-*

[1] Present address: U.S. Bureau of Land Management, Box 1009, Glenwood Springs, Colorado 81602, USA.

tulatus probably would never be achieved because of the variability of fish populations and stream perturbations. We postulate that patterns of variability in stream fish populations that violate the assumptions of OLS regression analysis may be hindering development of transferable habitat models by traditional data analysis techniques. Alternative techniques need to be evaluated.

Fausch et al. (1988) noted that a major assumption of regression models of stream fish standing crop is that fish populations are limited by the set of habitat variables included in the model. They stated that accurate models cannot be constructed unless factors not in the model that might limit standing crop below what the environment can support are eliminated or accounted for. Because standing stock is reduced by these other factors, it should be possible to account for them with a data analysis technique that models the upper limit (instead of the mean) of standing stock as a function of measured habitat variables. Standing stocks below a hypothesized limit attributed to habitat variables are not evidence that the limit has been incorrectly defined. Unequal weighting of high and low standing stocks associated with similar values of a habitat variable could provide a systematic approach for accounting for unmeasured "other factors" that limit standing stock and for identifying values of a habitat variable sufficient to limit fish populations. The relation of standing stock limits to a habitat variable could be very different than the relation of standing stock central tendency to the same habitat variable.

Aho et al. (1986) indicated that a wedge-shaped pattern of data distribution similar to Figure 1 will occur when a progressively lower value of a habitat variable is a sufficient condition to limit standing stocks to progressively lower values. Persons and Bulkley (1984) plotted biomasses of rainbow trout *Oncorhynchus mykiss* and cutthroat trout *O. clarki* against habitat suitability for instream cover and substrate. Their data have a similar wedge-shaped pattern. Several unpublished sets of standing stock data for Wyoming streams exhibit a similar distribution (W. A. Hubert, Wyoming Cooperative Fish and Wildlife Research Unit, personal communication). The wedge-shaped distribution is consistent with the hypothesis that habitat is capable of limiting populations and that numerous "other factors" often further limit populations below what the habitat could support.

The variation in standing stock described in Figure 1 is a function of the independent variable. This dependency violates the OLS regression assumption of homoscedastic (identical) error variances. Consequently, OLS regression estimates of mean standing stock conditional on a habitat variable do not have minimum sampling variance (precision), although they may be unbiased (Neter et al. 1989:418). Interpretation of r^2 and P-values is tenuous. Weighted least-squares regression (where weights are inversely proportional to error variances) or transformations (e.g., logarithmic) can be used with heteroscedastic data to estimate regression model parameters of central tendency that minimize sampling variance (Neter et al. 1989: 418). The resulting models are more precise estimates of mean (or median) standing stock but may not represent the relationship between upper limits of standing stock and habitat variables. Slopes of weighted regression models (mean or median) may be zero when patterns of heteroscedastic variances indicate variation in standing stock is functionally related to the habitat variable. Models of wedge-shaped or other patterns of unequal variation in stream fish standing stock relative to a habitat variable could help formulate testable predictions of changes in standing stock limits in response to changes in a habitat variable.

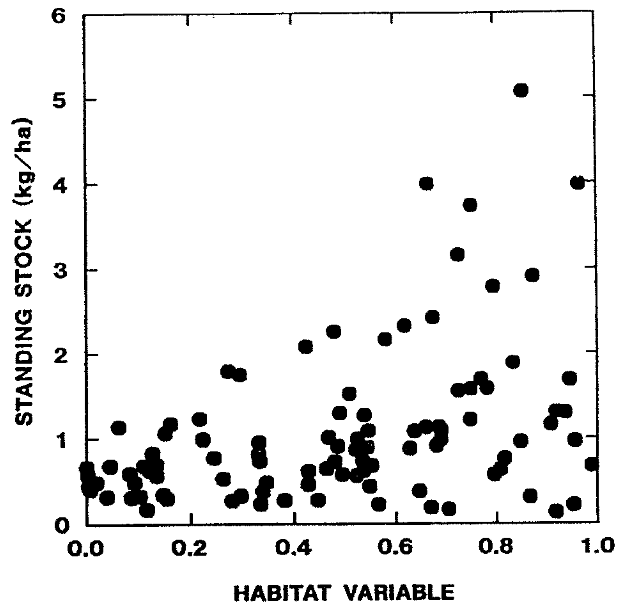

FIGURE 1.—An example of fish standing stock estimates in which both central tendency and variation increase as a linear function of a habitat variable.

We describe a three-step technique to identify data sets for which linear models of standing stock upper limits have different slopes than linear models of standing stock central tendency. The steps consist of testing for homoescedasticity by using residuals from an initial 50th regression quantile model, screening data to determine if heteroscedasticity is due to a wedge-shaped pattern of variability as a linear function, and estimating and comparing the slope of a 90th regression quantile to slopes of 50th regression quantiles and OLS regression. We demonstrate the three-step technique using two types of data sets: field data sets that were originally used to produce 35 published univariate OLS regression models of mean fish standing stock and simulated data sets that include wedge-shaped patterns of variation that yield models of central tendency with slopes equal to zero.

TABLE 1.—Independent habitat variables used in selected simple OLS regression models of fish standing stock.

Abbreviation	Variable or statistic
Kozel and Hubert (1989: brook trout)	
OVCOVA	Overhanging vegetation (%)
WDC	Woody debris control (%)
BPROOT	Backwater pools due to rootwads (%)
AQVEGCOV	Instream aquatic vegetation (%)
PAQVEG	Pools with aquatic vegetation (%)
GLIDE	Glide (%)
LGSED	Large fine sediment (%)
SMSED	Small fine sediment (%)
EMBEDD	Embeddedness (%)
BPDEBRIS	Backwater pools due to woody debris (%)
ELEV	Elevation (m)
Hubert and Rahel (1989: white sucker, creek chub, longnose dace, common shiner)	
MCRUN	Main-channel run (%)
SHADE	Shade (%)
TURBID	Turbidity (Jackson turbidity units)
WTEMP	Water temperatures (°C)
LGWOOD	Large woody debris (%)
CVCURREN	Coefficient of variation in current velocity
COVSV	Submerged aquatic vegetation (%)
ACURRENT	Mean current velocity (m/s)
ALLMCP	Main-channel pools (%)
PH	pH
RMCRUN	Rating of main-channel run
OVCOV	Overhanging cover (%)
ADEPTH	Shoreline water depth ≥15 cm(%)
ALLBP	Backwater pools (%)
MWIDDEP	Width-to-depth ratio
WREACH	Mean wetted width (m)
SILT	Silt substrate (%)
Layher and Maughan (1985, 1987a, 1987b: channel catfish, slenderhead darter, green sunfish)	
RISI	Suitability index for percent riffle
CASI	Suitability index for calcium hardness
MXWSI	Suitability index for maximum width

Methods

Field data sources.—Authors furnished us copies of computer files (Hubert and Rahel 1989; Kozel and Hubert 1989) or printouts (Layher and Maughan 1985, 1987a, 1987b) of original data analyzed in five refereed papers. The papers present significant ($P \leq 0.05$) OLS regression models for predicting conditional mean standing stocks from habitat variables. We repeated the univariate OLS regression analysis procedures described by the original authors to verify that we were using the same data sets described in their publications. We replicated 31 of the significant, single-variable OLS regression equations reported in the original publications and developed 4 additional univariate equations that had only minor discrepancies in the regression constant or variable coefficient. Some of the P and r^2 values associated with these 35 equations were slightly different than their published counterparts. For clarity, our analyses are based on these 35 univariate OLS equations and descriptive statistics (Tables 1, 2) instead of the original published equations and statistics.

The field data were collected in prairie streams in Kansas and Oklahoma (Layher and Maughan 1985, 1987a, 1987b) and in forested mountain (Kozel and Hubert 1989) and warmwater plains streams in Wyoming (Hubert and Rahel 1989). Widely accepted field techniques were used to measure habitat variables and to estimate standing stocks in sample reaches. The Wyoming studies used nontransformed, raw values of habitat variables to derive OLS regression equations to estimate standing stock. The Kansas and Oklahoma studies used values of habitat variables that had been transformed to suitability indices (suitability index range, 0.0–1.0).

Detailed descriptions of data collection techniques are presented in the original publications. In general, standing stock and habitat variables were measured in 30–300-m-long stream reaches that had been isolated with nets to prevent fish movement. Original definitions of the habitat variables for the 35 field data sets are given in Table 1. Most habitat variables were based on direct measurement of cover, substrate, and velocity within a stream reach concurrent with fish sampling. Field data were collected over a wide range of habitat conditions for eight species: channel catfish *Ictalurus punctatus* (Layher and Maughan 1985), slenderhead darter *Percina phoxocephala* (Layher and Maughan 1987a), green sunfish *Lepomis cyanellus* (Layher and Maughan 1987b), brook trout *Salvelinus fontinalis* (Kozel and Hubert 1989), and white sucker *Catostomus commersoni,* creek chub *Semotilus atromaculatus,* longnose dace *Rhinichthys cataractae,* and common shiner *Luxilus cornutus* (Hubert and Rahel 1989).

Simulated data.—The true pattern of variability of standing stock and measured habitat variables

59

TABLE 2.—Slopes and probabilities (null hypothesis, H_0: $\beta_1 = 0$) for ordinary least-squares (OLS) and least-absolute-deviation (LAD) regression and 90th regression quantiles for selected fish habitat models. Probability values for OLS regressions and tests for homoscedasticity were from normal theory F-tests. Probabilities for LAD regressions were from permutation tests with 10,000 random permutations. Homoscedasticity test 1 used absolute values of residuals from LAD regression as the dependent variable in an OLS regression on independent variables. Homoscedasticity test 2 used $\log_e$ of absolute values of residuals (two zero values were deleted) from LAD regression as the dependent variable in an OLS regression on $\log_e$ independent variables. See Table 1 for definitions of independent variable abbreviations.

Independent variable	OLS		LAD		90th quantile slope	Homoscedasticity	
	Slope	*P*	Slope	*P*		Test 1 *P*	Test 2 *P*
Kozel and Hubert (1989: brook trout; *N* = 28)							
OVCOVA	1.386	0.003	1.731	0.002	1.143	0.355	0.081
WDC	−1.490	0.052	−2.405	0.151	−2.398	0.250	0.352
BPROOT	10.558	0.008	14.577	0.017	11.122	0.525	0.352
AQVEGCOV	3.415	0.017	4.610	0.019	2.411	0.333	0.190
PAQVEG	4.288	0.011	7.556	0.038	6.000	0.057	0.201
GLIDE	2.542	0.011	3.963	0.015	1.665	0.181	0.377
LGSED	9.376	0.041	12.539	0.040	11.381	0.821	0.872
SMSED	18.164	0.018	27.486	0.074	19.428	0.099	0.126
EMBEDD	2.029	0.041	2.248	0.065	4.058	0.033	0.044
BPDEBRIS	−5.100	0.060	−5.291	0.067	−10.189	0.011	0.004
ELEV	0.310	<0.001	0.323	<0.001	0.249	0.518	0.625
Hubert and Rahel (1989: white sucker; *N* = 27)							
MCRUN	−0.119	0.008	−0.105	0.014	−0.108	0.218	0.204
SHADE	0.146	0.014	0.163	0.043	0.295	0.147	0.160
TURBID	−0.012	0.041	−0.011	0.118	−0.014	0.177	0.200
WTEMP	−0.341	0.036	−0.334	0.095	−0.725	0.081	0.114
LGWOOD	0.107	0.052	0.132	0.035	0.235	0.508	0.738
Hubert and Rahel (1989: creek chub; *N* = 26)							
CVCURREN	1.237	0.007	1.258	0.004	0.160	0.589	0.849
COVSV	0.033	0.013	0.031	0.048	0.061	0.491	0.854
ACURRENT	−3.314	0.019	−3.209	0.004	−1.551	0.508	0.909
ALLMCP	0.050	0.024	0.016	0.090	0.039	0.275	0.669
Hubert and Rahel (1989: longnose dace; *N* = 28)							
COVSV	0.027	0.001	0.028	0.011	0.041	0.074	0.108
PH	−0.752	0.014	−0.407	0.050	−0.985	0.286	0.470
RMCRUN	1.341	0.002	0.947	<0.001	1.693	0.056	0.003
OVCOV	0.012	0.002	0.008	0.021	0.023	0.050	0.146
ADEPTH	0.029	0.004	0.017	0.110	0.045	0.073	0.987
ALLBP	0.108	0.011	0.029	0.336	0.166	0.049	0.961
MWIDDEP	−0.035	0.021	−0.022	0.110	−0.041	0.160	0.426
WREACH	−0.194	0.033	−0.118	0.112	−0.399	0.119	0.643
WTEMP	−0.074	0.042	−0.059	0.067	−0.121	0.142	0.206
Hubert and Rahel (1989: common shiner; *N* = 8)							
COVSV	0.006	0.011	0.007	0.037	0.011	0.604	0.358
ALLBP	−0.090	0.011	−0.089	0.017	−0.105	0.556	0.512
SILT	0.026	0.042	0.035	0.048	0.029	0.914	0.627
Layher and Maughan (1985: channel catfish; *N* = 22)							
RISI	3200.298	<0.001	798.763	0.005	5434.729	0.005	0.004
Layher and Maughan (1987a: slenderhead darter, *N* = 10)							
CASI	8.167	<0.001	8.252	0.003	7.939	0.598	0.656
Layher and Maughan (1987b: green sunfish; *N* = 29)							
MXWSI	70.420	0.004	36.602	<0.001	189.453	0.032	<0.001

for the statistical populations represented by the field data sets is unknown. Thus, it is impossible to determine whether differences between regressions describing central tendency and upper regression quantiles represent actual trends in the sampled populations or are due to random variation. Therefore, we also analyzed one simulated data set (N = 60, df = 58 for univariate models)

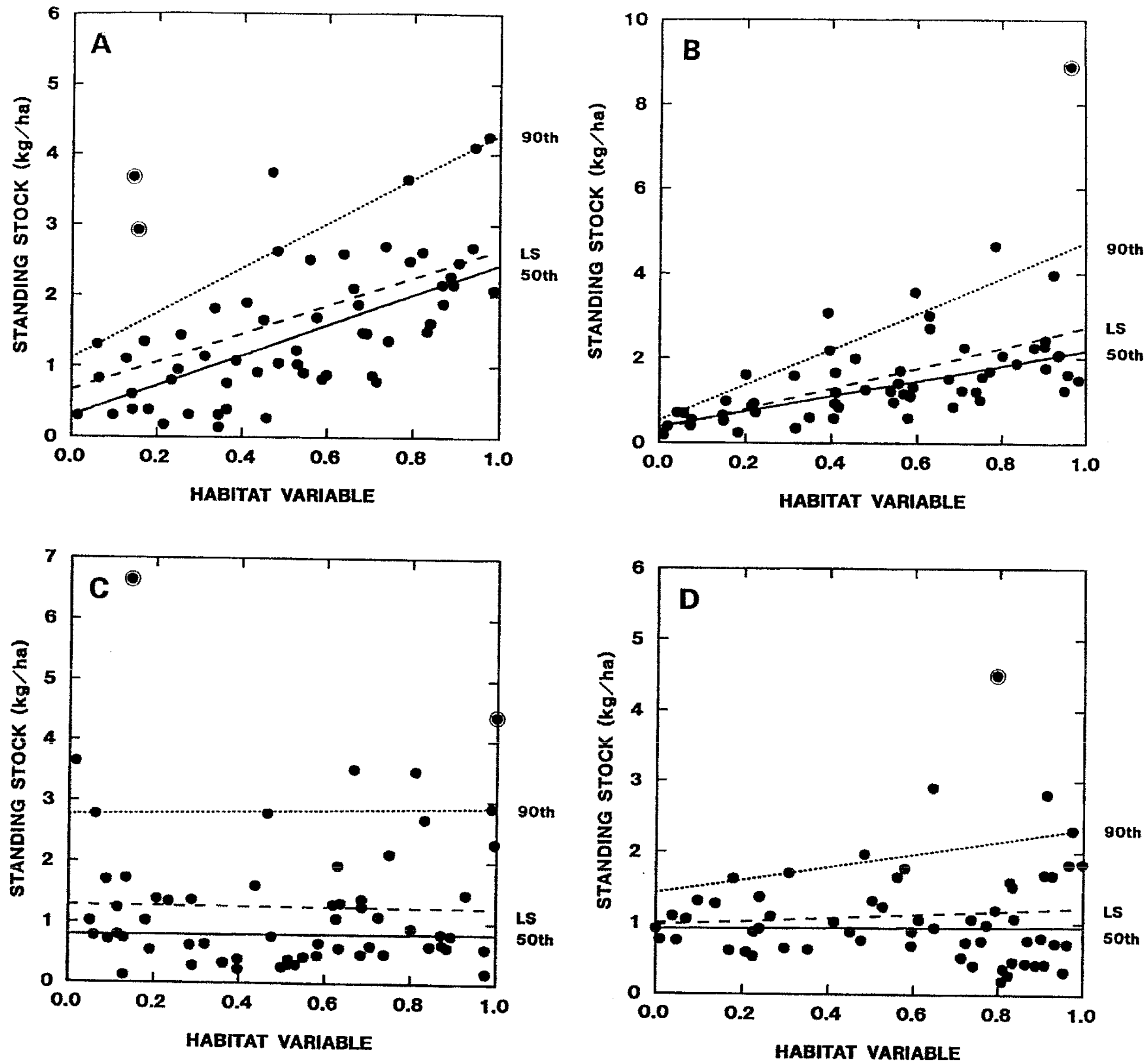

FIGURE 2.—An example of estimates of 50th and 90th regression quantiles and ordinary least-squares (OLS) regression estimates for random samples (N = 60) from identical (A, C) and nonidentical (B, D) error distributions for the linear models $Y = 0.25 + 2.0X + \epsilon$ (A, B) and $Y = 1.0 + 0.0X + \epsilon$ (C, D). Errors are lognormally distributed with median = 0 and constant variance [$\sigma^2 = e^c(e^c - 1)$, $c = 0.75^2$] in A and C and with variance increasing as a function of the independent variable [$\sigma^2 = e^c(e^c - 1)$, $c = (0.2 + 0.8X)^2$] in B and D. Each independent variable (X) is a random uniform variate (0,1). Circled values indicate outliers (studentized residuals > 3) detected by OLS regression.

generated as a random sample from each of four statistical populations designed to mimic four patterns of standing stock variation as a linear function of a habitat variable. In the first population, median standing stock increased as a function of a habitat variable, and error distributions were identical (homogeneous variances) at all values of the habitat variable (Figure 2A). In the second population, median standing stock and error variances both increased as a function of a habitat variable (Figure 2B). In the third population, median standing stock was unrelated (slope = 0) to the habitat variable and error distributions were identical (Figure 2C). In the fourth population, median standing stock was unrelated to the habitat variable, but error variance increased as a function of the habitat variable (Figure 2D). We used lognormal error distributions as a convenient method of simulating patterns of standing stock that have positive values and that are right-skewed toward high values (Dennis and Patil 1988). The simulated data were generated by sampling across the entire range of the independent variable. The sample size of 60 was based on Fausch et al. (1988). They recommended a sample size (N) of at least 20 for field studies and reported an average of 58 df for linear regression models that predicted standing stock from several types of independent (habitat) variables based on sample sizes exceeding 20. Analyses of the simulated data provided realistic examples of results that could be obtained from data collected in a rigorous field study of relations between habitat variables and standing stock.

Definition of regression quantiles.—A quantile is a plane that splits a frequency distribution into parts. For example, the 50th quantile, or median, splits a frequency distribution into two parts containing an equal number of observations. The quantile concept can be extended to regression by minimizing a function of absolute deviations (i.e., regression quantiles) and can estimate any quantile conditional on a linear model (Koenker and Bassett 1978, 1982). The linear model is $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{y}$ is an $N \times 1$ vector of dependent responses (e.g., standing stock), $\boldsymbol{\beta}$ is a $(p + 1) \times 1$ vector of unknown regression parameters, $\mathbf{X}$ is an $N \times (p + 1)$ matrix of predictors (e.g., habitat variables), and $\boldsymbol{\epsilon}$ is an $N \times 1$ vector of random errors that are not necessarily identically distributed. Regression quantile estimates ($\boldsymbol{b}$) for any quantile θ, $0 < \theta < 1$, are solutions to the following minimization problem in simplex linear programming (Koenker and D'Orey 1987):

$$\min\left[\sum_{i=1}^{N} \tau\left|y_i - \sum_{j=0}^{p} b_j x_{ij}\right|\right]; \qquad (1)$$

$$\tau = \theta \qquad \text{for } y_i \geq \sum_{j=0}^{p} b_j x_{ij}, \text{ and}$$

$$\tau = 1 - \theta \quad \text{for } y_i < \sum_{j=0}^{p} b_j x_{ij}.$$

The essence of estimating function (1) is that positive and negative residuals are differentially weighted for regression quantiles other than $\theta = 0.5$. The 50th regression quantile ($\theta = 0.5$) is equivalent to least-absolute-deviation (LAD) regression. The LAD regression estimates conditional medians in a linear model, whereas OLS regression estimates conditional means. Use of LAD regression as an alternative to OLS regression for modeling central tendency is often recommended because it is more resistant to the influence of outlying values and has greater power for both asymmetric (e.g., lognormal) and thick-tailed symmetric (e.g., double exponential), identical error distributions (Dielman and Pfaffenberger 1982; Narula and Wellington 1982; Rao 1988; Birkes and Dodge 1993).

A regression quantile, θ, with p parameters passes through at least p sample observations (there are p residuals equal to zero). There are at least $(N - p)(1.0 - \theta)$ sample observations above and at most $N\theta$ sample observations below a regression quantile (Koenker and Bassett 1982). Regression quantile estimates are median unbiased and remain unbiased for nonlinear (e.g., logarithmic) as well as linear, monotonic transformations, whereas OLS estimates are mean unbiased only for linear, monotonic transformations (Bassett 1992). When error distributions are unimodal, sampling variation of regression quantiles increases for quantiles farther away from the 50th. Regression quantile estimates are insensitive to extreme values of outlying dependent variables. For example, the outlier in Figure 2B could be 10, 100, or 1,000 kg/ha and the 90th and 50th regression quantile estimates would remain unchanged.

All regression quantiles and OLS regression are parallel lines (planes) with slope estimates differing only due to sampling variation (Figure 2A, C) when error distributions are identical (e.g., homoscedastic variance; Koenker and Bassett 1978, 1982). Estimating outer quantiles for identically distributed errors provides a minimum of new information on habitat variables that could be limiting standing stock. Quantile and OLS regression slopes would be similar but intercepts would differ. When error distributions are not identical (e.g., heteroscedastic variances; Figure 2B, D), regression quantiles are no longer all parallel (Koenker and Bassett 1978, 1982). Estimating outer quantiles could provide models of standing stock limits with different slopes than models of central tendency and identify variables that limit standing stocks but are not related to mean (or median) standing stock.

Methods to test hypotheses about slopes of regression quantiles when error distributions are not identical are currently being developed and evaluated in the statistical literature (Koenker 1994; Parzen et al. 1994). The statistical performance of competing methods is incompletely evaluated, and computational routines reside in developmental software. Therefore, we applied an indirect technique to identify linear relations between habitat variables and standing stock upper limits that could differ from similar models of central tendency. We combine results of well-established tests for homogeneity of error distributions with the fact that all regression quantiles cannot be parallel for heteroscedastic error distributions. We estimate only an upper limit, represented by the 90th regression quantile, because the lower limit of standing stock variation is effectively bound by zero.

Tests for homoscedastic error distributions (step 1).—We made two tests of the null hypothesis that standing stock was homoscedastic based on regression coefficients from two of the many available procedures for estimating weighting functions of the independent variables for weighted least-squares regression. Test 1 for homoscedasticity used absolute values of residuals from an initial LAD regression estimate as the dependent variable in an OLS regression on the independent variable(s) (Glejser 1969). Test 2 used $\log_e$(absolute values of residuals) from an initial LAD regression estimate as the dependent variable in an OLS regression on $\log_e$[independent variable(s)] (Harvey 1976). In test 2, we deleted residuals with values of zero before taking natural logarithms (Davidian and Carroll 1987). We chose LAD regression to generate residuals for the two tests because residuals used as dependent variables are greatly influenced by outlying values (Davidian and Carroll 1987) and LAD regression is the simplest available regression estimate that minimizes the influence of outlying values (Rao 1988). Null hypotheses for both homoscedasticity tests are that the OLS regression slopes are equal to zero. Both tests have good power to detect heteroscedasticity for non-normal error distributions (Davidian and Carroll 1987). Rejection of the null hypothesis for either test was considered evidence of heteroscedasticity. We used both normal theory and permutation procedures (which make no assumptions of normal distribution) in program RT (Manly 1991) to test null hypotheses. Differences between probabilities from permutation tests (10,000 random permutations) and normal theory probabilities were less than 0.02 in all but a few OLS regressions (in which both probabilities were greater than 0.40), so we report only normal theory results.

The same LAD regression equations generated for the homoscedasticity tests are an alternative to OLS regression for modeling central tendency. For a more complete evaluation of the evidence that central tendency of standing stock was linearly related to the habitat variables, we tested significance of slopes of the LAD regression models, using a permutation procedure similar to one described for OLS regression by Kennedy (1995). The observed test statistic, T_o, was the proportionate reduction in sums of absolute deviations when the analysis passed from the reduced (no slope parameter) to the full model LAD regression for the observed sample. The vector of residuals from the reduced model was permuted 10,000 times and similar statistics (T) calculated at each permutation for LAD regressions on the independent variable matrix. Probability under the null hypothesis that the slope equals zero was given by the fraction of 10,000 T values that equaled or exceeded T_o. Validity of the permutation tests for LAD and OLS regression is based on an assumption of identical error distributions.

Regression quantiles developed for this study were estimated with the computer program BLOSSOM (Slauson et al. 1991). This software has two published statistical routines: Koenker and D'Orey's (1987) modification of the Barrodale and Roberts (1974) LAD regression algorithm and Knuth's (1981) permutation routine. The OLS regressions were computed with SYSTAT (Wilkinson 1990).

Data screening (step 2).—Data screening is necessary because significance tests for homogeneity are only indirectly related to differences between 50th and 90th quantiles and because they assume the linear model form is correct. If homoscedasticity was rejected by using test 1 or 2 as described above, we visually examined data plots to determine if heteroscedasticity was due to a wedge-shaped pattern of variability that was increasing or decreasing across values of the independent variables, due to the effect of a few outliers or to the inappropriateness of the linear model form. Residual diagnostics from OLS regressions (studentized residuals and leverage values) were used to identify the most severe outliers. A wedge-shaped pattern without outliers or a wedge-shaped pattern that persists when outliers are ignored is evidence that a linear model form is appropriate for modeling standing stock, and we recommend continuing with step 3 of the analysis. If the linear pattern seems inappropriate, continuing to step 3 implies acceptance of a linear model for what may be a nonlinear response. We did not explore other model forms because we wanted to keep our exposition simple.

Comparison of regression model slopes (step 3).—Complete data sets, including any outliers identified in step 2, were used. Knowing that all quantiles could not be parallel in a heteroscedastic model, we estimated, plotted, and visually compared 90th and 50th regression quantiles. We looked for patterns of variation in which 90th regression quantile slopes were of greater magnitude than 50th regression quantile or OLS regression slopes. Such patterns provide the strongest evidence that limiting relations between standing stock and habitat were different than the relationship between central tendency (average) of standing stock and habitat. The 90th quantile is exceeded by approximately 10% of observations; selecting this quantile to model limits to standing stock was somewhat arbitrary. We assume investigators are interested in identifying habitat variables that may limit populations in stream reaches that are longer than a typical sample reach. It is reasonable to expect that temporary aggregations of fish lead to higher short-term densities in reaches (approximately 30–300 m) typically used as sample units than are found in longer stream reaches over extended periods. A 10% exceedence rate is easy to visualize; other quantiles could be selected depending on the objectives of the study.

Results

Standing Stock Relationships in Field Data Sets

Thirteen of the 35 field data sets had nonidentical error distributions (Table 2; homoscedasticity test 1 or test 2, $P < 0.10$). Eleven of these 13 data sets had high leverage or studentized residuals for OLS regressions. Based on data screening (step 2) with the outliers identified, we attributed 4 of the 13 cases of heteroscedasticity to one or two outlying values of standing stock (Figure 3B–E) and another case to an inappropriate linear model (Figure 3A). For these 5 cases, there was no pattern in the dispersion of standing stock as a linear function of the habitat variables. Application of step 3 to these data implies acceptance of a linear model form of standing stock variation as a function of the habitat variables. We display regressions for these data to emphasize the need to screen and display all data to look for a pattern that is represented with a linear model (Figure 3). Simply comparing regression slopes is insufficient.

In the remaining eight data sets, heteroscedasticity was due to a pattern of variation in standing stock that was a linear function of the independent variable (Figure 4). Based on our three-step technique, the 90th regression quantile for these eight data sets is a model of a single habitat variable acting as a limiting factor with a slope that differs from zero and differs from the OLS and LAD regression slopes for models of central tendency (mean and median) for the same variable. In three of the eight data sets (Figure 4F–H), probabilities for homoscedasticity tests were less than probabilities for tests of zero slopes in LAD and OLS regressions: brook trout standing crop as a function of backwater pools with woody debris (BPDEBRIS) and as a function of embeddedness (EMBEDD), and green sunfish standing stock as a function of a suitability index for maximum channel width (MXWSI; Table 2). Variations in standing stock were more significantly related to these three habitat variables than standing stock central tendency and the 90th regression quantile estimates were farther from zero than either the OLS or LAD estimates.

Simulation Data

Plots of the four simulated data sets ($N = 60$ for each data set) generated from lognormal distributions with LAD (50th regression quantile) slope parameters equal to 2.0 and 0.0 are presented in Figure 2. The first step in our technique is to test for homoscedasticity. For the two simulated data sets generated from identical error distributions (Figure 2A, C), both tests for homoscedasticity were not rejected at $P \geq 0.224$ (Table 3). Failure to reject homoscedasticity leads to the conclusion that the 90th regression quantile as a linear model of the upper limit of standing stocks would have a slope similar to models of central tendency (LAD or OLS); therefore, the remaining steps of the technique are not applied. The habitat variable identified through OLS and LAD regression analysis as having a zero slope (Figure 2C) is not considered a potentially limiting factor because there is no evidence that the 90th quantile is not parallel to these estimates.

In both simulated data sets sampled from nonidentical error distributions (Figure 2B, D) one of the two tests for homoscedasticity was rejected at $P < 0.05$ (Table 3). Rejection of homoscedasticity leads to the conclusion that all regression quantiles cannot be parallel and we proceed to step 2 of our technique. For both of the data sets, at least one outlier is detected (studentized residuals ≥ 3), which is consistent with the skewed nature of the lognormal distribution. Visual inspection of the plots indicates that heteroscedasticity is due to a wedge-shaped pattern of variation across the independent variable (known to be true for the sampled population) rather than to a few outliers. A linear model of standing stocks appears appropriate. For both data sets, the 90th regression quantile slope estimate is greater than LAD or OLS estimates (Table 3). In the data set drawn from a pop-

59

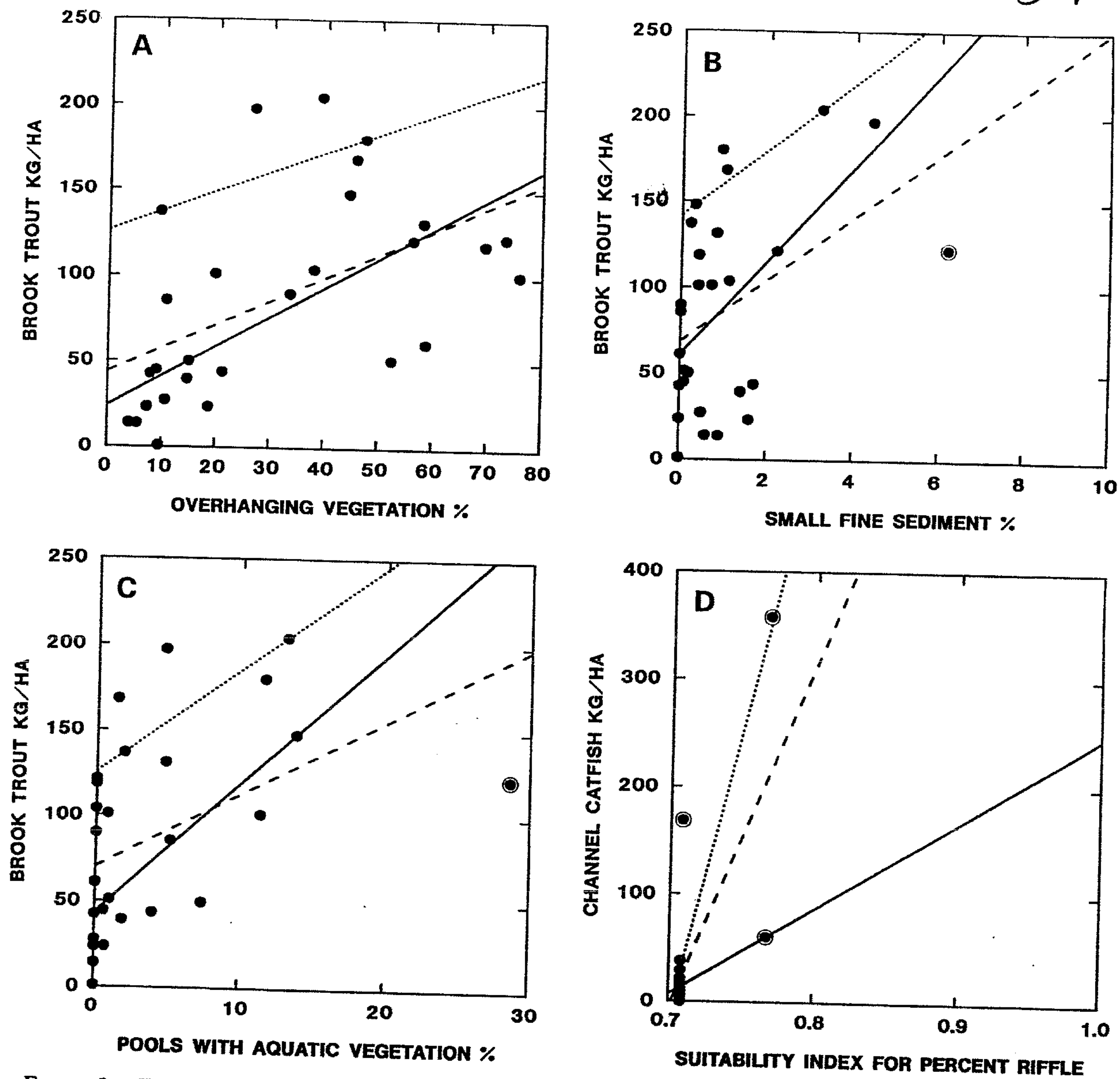

FIGURE 3.—Estimates of ordinary least-squares regression (dashed line), least-absolute-deviation regression (solid line), and 90th regression quantile (dotted line) for fish standing stock (kg/ha) as a function of independent habitat variables. Models had $P < 0.10$ for homoscedasticity test 1 or test 2 in Table 2 but did not show strong evidence of dispersion increasing as a linear function of independent variables. Circled values had outlying studentized residuals or high leverage values from ordinary least-squares regression.

ulation where the LAD slope was not-zero, the significance of the LAD slope estimate clearly implies that the 90th quantile slope is also not zero. In the simulated data set drawn from a population where the slope parameter of the LAD regression was equal to zero (Figure 2D), the tests for homoscedasticity had $P < 0.05$, whereas tests for zero slopes for LAD and OLS regression estimate had $P > 0.40$. The 90th regression quantile slope appears to be nonzero, correctly implying a limiting relationship for the data set in Figure 2D. This data set mimics the wedged-shaped pattern that could be expected when decreasing values of a measured habitat variable (or habitat suitability index) are sufficient to limit standing stock but other factors often further limit standing stock.

Reciprocals of the OLS regression coefficients used to test for homogeneity can be used as weights to estimate a weighted least-squares (WLS) or weighted LAD regression for the data in Figure 2D. Weighted regression produces more precise estimates of the slope parameter, which is zero, and leads to the same correct conclusion that standing stock central tendency is not related to

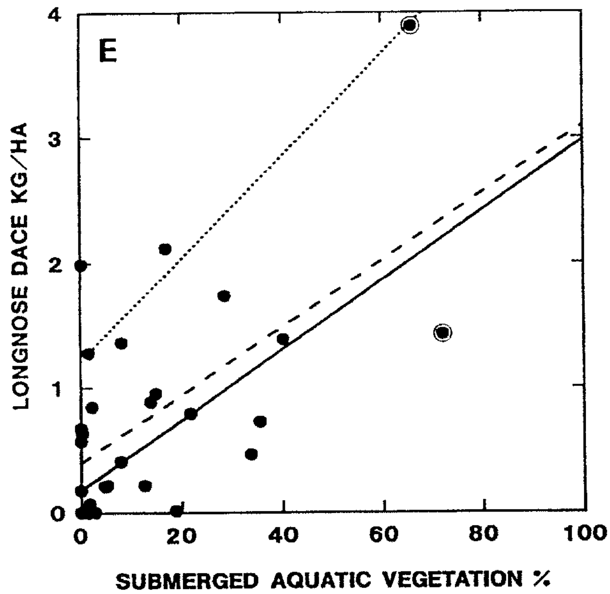

FIGURE 3.—Continued.

the habitat variable. Large field data sets with the pattern of variation shown in Figure 2D would support the hypothesis that habitat is limiting standing stock but would seldom appear in the refereed literature because attempts to produce statistically significant OLS, WLS, or LAD regression models of central tendency would fail. Data sets with a pattern of variation similar to Figure 2B would be more likely to be identified as evidence of habitat acting as a limiting factor because standing stock central tendency does vary with habitat in this scenario.

Discussion

Our technique for modeling limiting relations of standing stock to habitat has some obvious limitations. The lack of thoroughly evaluated procedures to test hypotheses about slopes of 90th regression quantiles (or other quantiles that might be selected) when error distributions are not identical makes it difficult to select a best model based on a standardized statistical comparison. New methods of testing and constructing confidence intervals for regression quantiles are being evaluated in the statistical literature (Koenker 1994), but acceptable procedures for routine use by biologists are still needed. As methods for confidence interval calculation become established, it will be possible to test the significance of regression quantile slopes directly. Confidence intervals for 90th and 50th regression quantiles with minimal or no overlap would be additional evidence that the relation of standing stock limits is different than the relation of standing stock central tendency to a habitat variable. Without directly testing the slope of the 90th regression quantile it must be assumed, based on significance of homoscedasticity tests, that the 90th quantile slope is not parallel to slopes for regression models of central tendency. We speculate that practical permutation tests to compare parallelism of several regression quantiles (e.g., 10th, 25th, 50th, 75th, 90th) based on procedures of Koenker and Bassett (1982) and Welsh et al. (1994) can be developed. These procedures would provide an alternative for testing homogeneity of variances that is more directly related to identifying situations in which the slope of an upper quantile is different from the slope of central tendency.

We limited our technique to a linear model form of a regression quantile. Although this is adequate for wedge-shaped patterns of variability, more complex patterns would require the transformation of data to a wedge-shaped pattern or use of a more complex model form. The use of suitability indices (e.g., Layher and Maughan 1985, 1987a, 1987b) as independent variables is an approach to transforming nonlinear relations of standing stock to habitat into wedge-shaped patterns. Modeling complex relationships with nonlinear forms of different regression quantiles has received theoretical treatment by Welsh et al. (1994), but practical procedures are lacking. Diagnostic procedures are needed to evaluate appropriateness of simple linear versus more complex regression quantile model forms. The Akaike information criterion (AIC) is a general model selection procedure (Burnham and Anderson 1992) used to select among OLS and WLS regression models. Hurvich and Tsai (1990) have extended the use of AIC model selection to LAD (median) regression; extensions to other quantiles may be possible.

As demonstrated with simulated data sets, variation in standing stock can be linearly conditional on a habitat variable and yield nonsignificant regression models for conditional means (OLS regression) or medians (LAD regression). Modeling upper limits of standing stock with these types of data allows otherwise overlooked variables to be identified as having a relation to standing stock. However, quantifying patterns of dispersion in standing stock can be useful even if statistically significant relations between standing stock central tendency and habitat can be described. For example, the OLS regression model of longnose dace standing stock as a function of overhanging cover (OVCOV; Table 2) has a significant P-value and

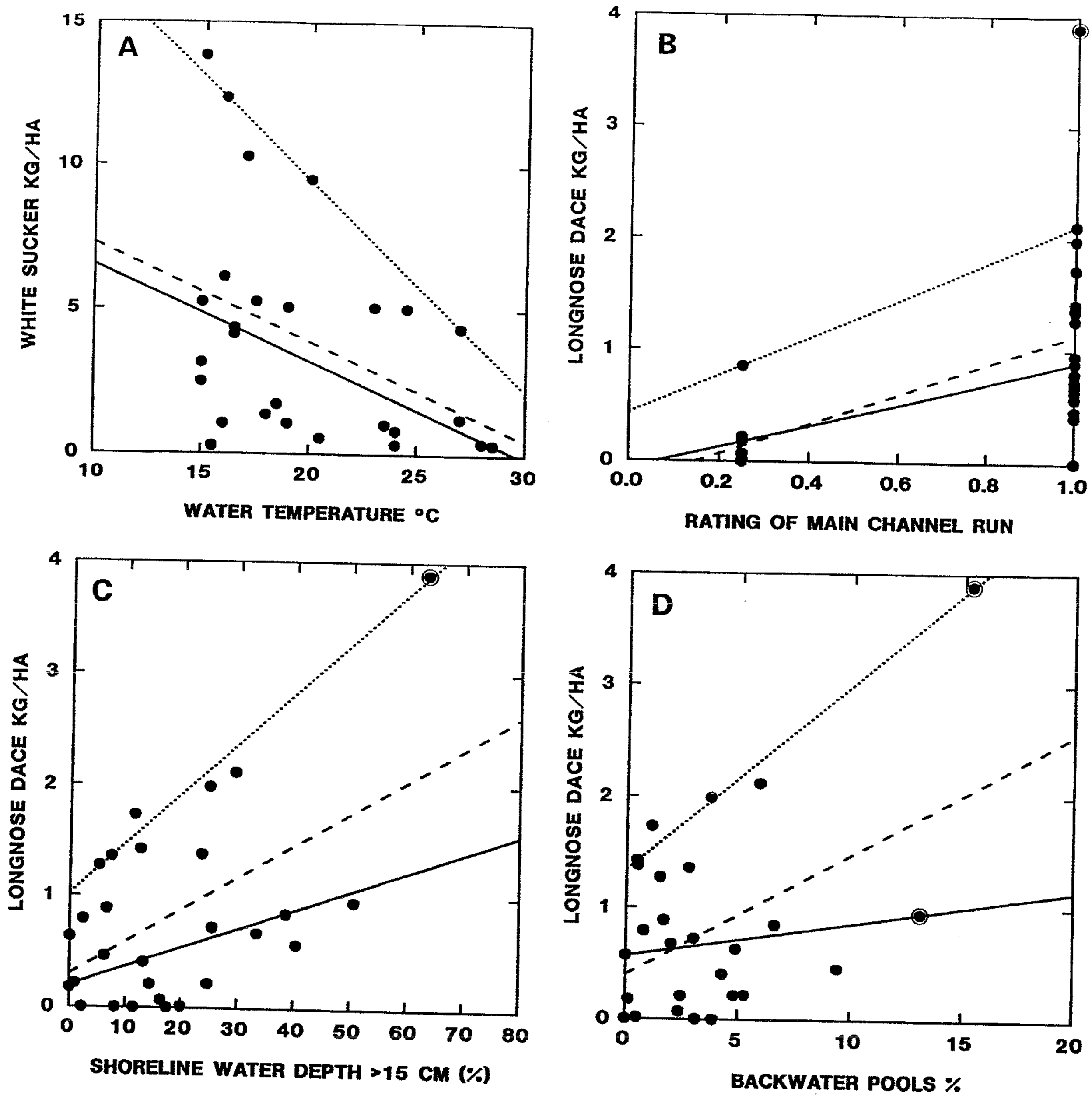

FIGURE 4.—Estimates of ordinary least-squares regression (dashed line), least-absolute-deviation regression (solid line), and 90th regression quantile (dotted line) for fish standing stock (kg/ha) as a function of independent habitat variables. Models had $P < 0.10$ for homoscedasticity test 1 or test 2 in Table 2 and showed strong evidence of dispersion increasing as a linear function of independent variables. Circled values had outlying studentized residuals or high leverage values from ordinary least-squares regression.

an r^2 of 0.312. The coefficient of determination (r^2) measures percent reduction in the variance of standing stock (which is measured in squared units) attributed to the regression model. The value of r^2 must be converted back to the original standing stock units by the relation $1 - (1 - r^2)^{0.5}$ to determine the variation in standing stock attributed to or "explained by" the regression model (Ehrenberg 1975:233). This conversion shows that the OLS regression model has "reduced" variation in standing stock by only 17%, a rather weak relationship that could be difficult to detect, which makes this habitat variable a low priority for testing with independent data. However, examination of the regression quantiles and plotted data (Figure 4E) suggest a strong limiting relationship and the need for more observations at extreme low and extreme high values of the independent variable.

Significant one-variable OLS regression models developed from field data were based on relatively

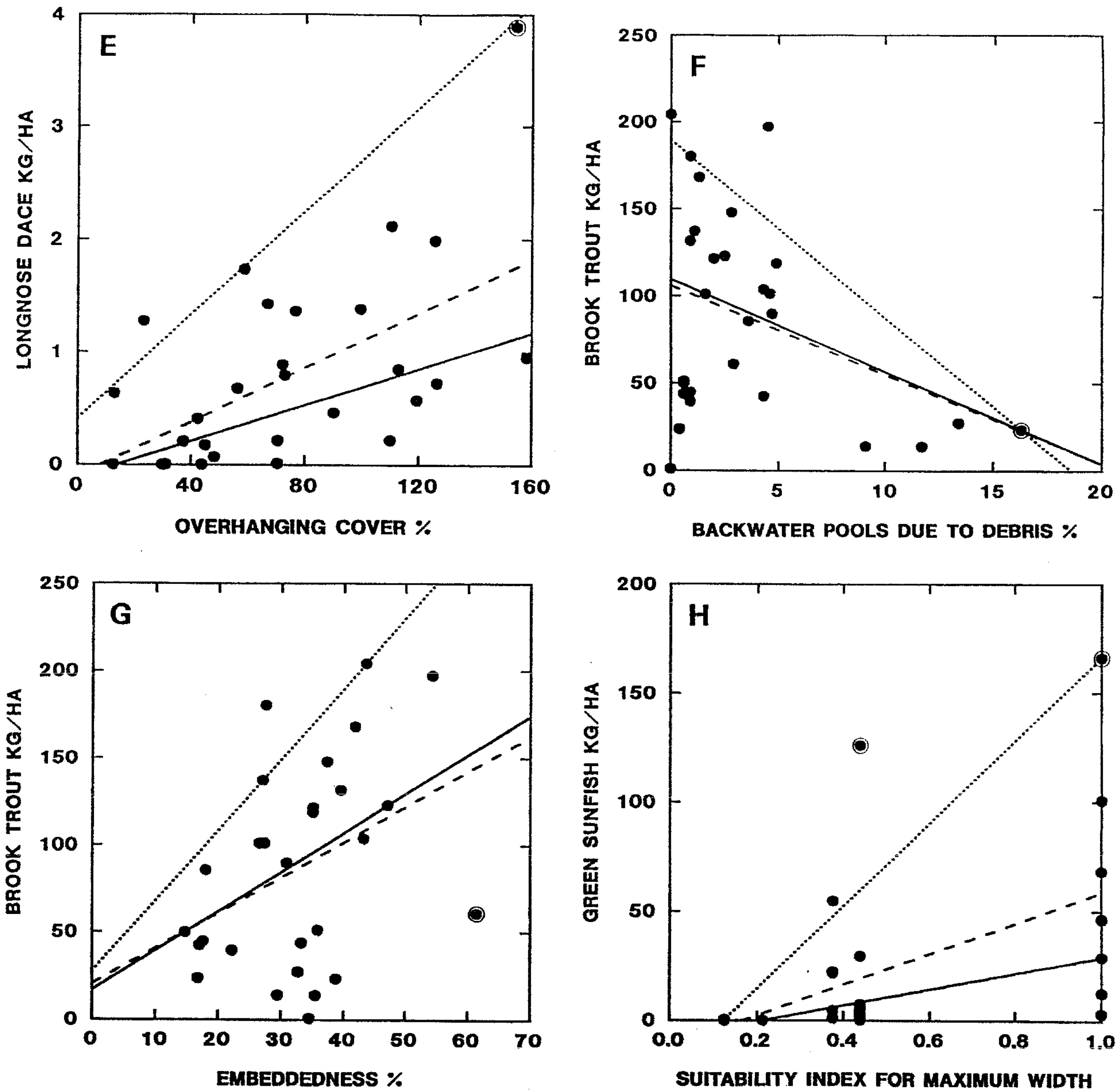

FIGURE 4.—Continued.

TABLE 3.—Slopes and probabilities (H_0: $\beta_1 = 0$) for ordinary least-squares (OLS) and least-absolute-deviation (LAD) regression's and 90th regression quantiles for the four random data sets ($N = 60$) shown in Figure 2. The 50th quantile slope parameters are equal to 2.0 and 0.0; error distributions are identical (I) or nonidentical (N) that increase as a function of the habitat variable. Probability values for OLS regressions and tests for homoscedasticity were from normal theory F-tests. Probabilities for LAD regressions were from permutation tests with 10,000 random permutations. Homoscedasticity test 1 used absolute values of residuals from LAD regression as the dependent variable in an OLS regression on independent variables. Homoscedasticity test 2 used $\log_e$ of absolute values of residuals (two zero values were deleted) from LAD regression as the dependent variable in an OLS regression on $\log_e$ independent variables.

Parameters		OLS estimates		LAD estimates		90th quantile slope	Homoscedasticity	
Slope	Errors	Slope	*P*	Slope	*P*		Test 1, *P*	Test 2, *P*
2.0	I	1.985	<0.001	2.136	<0.001	3.166	0.927	0.224
2.0	N	2.425	<0.001	1.801	<0.001	4.208	0.035	0.077
0.0	I	0.033	0.945	0.028	0.941	0.129	0.932	0.867
0.0	N	0.251	0.422	0.031	0.965	0.907	0.015	0.032

small sample sizes (N = 8–29), and the majority had low r^2. The trend of decreasing model precision with increasing sample size described by Fausch et al. (1988) was clearly evident. For example, Hubert and Rahel (1989) developed three significant OLS regression equations to predict standing stock of common shiner from single habitat variables. These equations were derived from data collected at only eight stream sites, and r^2 ranged from 0.52 to 0.69. Our technique provides minimal additional insight to these data sets, and no new models are developed. In contrast, the nine significant univariate OLS regression equations generated by Hubert and Rahel (1989) to predict standing stock of longnose dace were derived from data collected at 28 stream sites and had r^2 of 0.15–0.35. Our technique yielded three limiting factor models for these data. Sampling a greater number of sites may increase the heterogeneity of responses around a hypothesized mean, but it also has the potential to provide better definition of a pattern. Our technique should be most useful with larger data sets, in which patterns may persist even if outliers are ignored. Investigators (e.g., Layher and Maughan 1985, 1987a, 1987b) routinely attempt to improve OLS regression models of central tendency by adding additional variables to improve the coefficient of determination. Our technique was limited to analysis of a single variable; sampling a greater range of habitat conditions with larger sample sizes should be the best approach to improving univariate regression quantile models of limiting relations.

Regression quantiles can be estimated (Koenker and Bassett 1978, 1982) and homoscedasticity tested (Glejser 1969; Harvey 1976; Davidian and Carroll 1987) for multiple independent variables. However, we limited our application to single independent variables because the lack of established significance tests for slopes of quantiles developed from nonidentical error distributions led to the use of visual data screening procedures. Establishment of formal statistical tests for regression quantile slopes will improve the usefulness of limiting factor models by allowing direct comparisons to slopes of central tendency models. Graphically displaying regression quantiles on multiple independent variables to search for nonparallel regressions and to determine relative contribution of different independent variables will require more creative procedures. Residual diagnostics for OLS regression on multiple independent variables are well established, but similar techniques need to be developed for regression quantiles.

Plots of residuals for the 35 OLS equations derived from field data indicated potential violation of OLS regression assumptions of homoscedastic variances and normality for 32 of the 35 data sets. Variation in standing stocks was linearly related to a habitat variable in 13 of 35 field data sets supporting significant models between conditional means of standing stock and habitat variables (OLS regression). These violations of assumptions weaken the direct comparison of P-values associated with OLS regression and P-values for tests of homoscedasticity. Our main concern, however, is to quantify limiting relationships that are overlooked or misrepresented with conventional analyses. The usefulness of this quantification does not ultimately rest on the P-values for homoscedasticity and regressions but on the ability of the relationships to generate testable predictions and suggest causal mechanisms that can be tested by experimental manipulation. If instances of a single habitat variable limiting standing stock are rare, large numbers of observations will be needed to accurately quantify limiting relationships and to model habitat requirements. Our technique provides an approach to this quantification that takes advantage of readily available statistical tests for homogeneity in OLS regression.

Acknowledgments

We thank W. Hubert and W. Layher for providing their data and reviewing draft versions of this paper, G. Bassett for reviewing drafts, and referees for providing thoughtful comments that improved the final manuscript.

References

Aho, J. M., C. S. Anderson, and J. W. Terrell. 1986. Habitat suitability index models and instream flow suitability curves: redbreast sunfish. U.S. Fish Wildlife Service Biological Report 82(10.119).

Barrodale, I., and F. D. K. Roberts. 1974. Algorithm 478: solution of an overdetermined system of equations in the l_1 norm. ACM (Association for Computing Machinery) Communications 17:319–320.

Bassett, G. 1992. The Gauss Markov property for the median. Pages 23–31 *in* Y. Dodge, editor. L_1 statistical analyses. North-Holland, Amsterdam.

Birkes, D., and Y. Dodge. 1993. Alternative methods of regression. Wiley, New York.

Bowlby, J. N., and J. C. Roff. 1986. Trout biomass and habitat relationships in southern Ontario streams. Transactions of the American Fisheries Society 115: 503–514.

Burnham, K. P., and D. R. Anderson. 1992. Data-based selection of an appropriate biological model: the key to modern data analysis. Pages 16–30 *in* D. R. McCullough and R. H. Barrett, editors. Wildlife 2001: populations. Elsevier Science Publishers, Barking, UK.

Davidian, M., and R. J. Carroll. 1987. Variance function estimation. JASA (Journal of the American Statistical Association) 82:1079–1091.

Dennis, B., and G. P. Patil. 1988. Applications in ecology. Pages 303–330 *in* E. L. Crow and K. Shimizu, editors. Lognormal distributions: theory and applications. Marcel Dekker, New York.

Dielman, T., and R. Pfaffenberger. 1982. LAV (least absolute value) estimation in linear regression: a review. TIMS (The Institute of Management Sciences) Studies in Management Science 19:31–52.

Ehrenberg, A. S. C. 1975. Data reduction: analyzing and interpreting statistical data. Wiley, New York.

Fausch, K. D., C. L. Hawks, and M. G. Parsons. 1988. Models that predict standing crop of stream fish from habitat variables (1950–1985). U.S. Forest Service General Technical Report PNW-GTR-213.

Glejser, H. 1969. A new test for heteroskedasticity. JASA (Journal of the American Statistical Association) 64:316–323.

Harvey, A. C. 1976. Estimating regression models with multiplicative heteroscedasticity. Econometrica 44: 461–465.

Hubert, W. A., and F. J. Rahel. 1989. Relations of physical habitat to abundance of four nongame fishes in high-plains streams: a test of habitat suitability index models. North American Journal of Fisheries Management 9:332–340.

Hurvich, C. M., and C-L. Tsai. 1990. Model selection for least absolute deviations regression in small samples. Statistics and Probability Letters 9:259–265.

Kennedy, P. E. 1995. Randomization tests in econometrics. Journal of Business and Economic Statistics 13:85–94.

Knuth, D. E. 1981. The art of computer programming. Volume 2, seminumerical algorithms. Addison-Wesley, Reading, Massachusetts.

Koenker, R. 1994. Confidence intervals for regression quantiles. Pages 349–360 *in* P. Mandl and M. Huskova, editors. Asymptotic statistics: proceedings of the 5th Prague symposium. Springer-Verlag, New York.

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica 46:33–50.

Koenker, R., and G. Bassett. 1982. Robust tests for heteroscedasticity based on regression quantiles. Econometrica 50:43–61.

Koenker, R. W., and V. D'Orey. 1987. Computing regression quantiles. Applied Statistics 36:383–393.

Kozel, S. J., and W. A. Hubert. 1989. Factors influencing the abundance of brook trout (*Salvelinus fontinalis*) in forested mountain streams. Journal of Freshwater Ecology 5:113–122.

Lanka, R. P., W. A. Hubert, and T. A. Wesche. 1987. Relations of geomorphology to stream habitat and trout standing stock in small Rocky Mountain streams. Transactions of the American Fisheries Society 116:21–28.

Layher, W. G., and O. E. Maughan. 1985. Relati between habitat variables and channel catfish pop ulations in prairie streams. Transactions of the American Fisheries Society 114:771–781.

Layher, W. G., and O. E. Maughan. 1987a. Relations between two darter species and their respective abiotic environments. Proceedings of the Oklahoma Academy of Science 67:11–22.

Layher, W. G., and O. E. Maughan. 1987b. Modeling habitat requirements of a euryhabitat species. Transactions of the Kansas Academy of Science 90:60–70.

Layher, W. G., O. E. Maughan, and W. D. Warde. 1987. Spotted bass habitat suitability related to fish occurrence and biomass and measurements of physicochemical variables. North American Journal of Fisheries Management 7:238–251.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, New York.

McClendon, D. D., and C. F. Rabeni. 1987. Physical and biological variables useful for predicting population characteristics of smallmouth bass and rock bass in an Ozark stream. North American Journal of Fisheries Management 7:46–56.

Narula, S. C., and J. F. Wellington. 1982. The minimum sum of absolute errors regression: a state of the art survey. International Statistical Review 50:317–326.

Neter, J., W. Wasserman, and M. H. Kutner. 1989. plied linear regression models. Irwin, Homewo Illinois.

Pajak, P., and R. J. Neves. 1987. Habitat suitability and fish population: a model evaluation for rock bass in two Virginia streams. Transactions of the American Fisheries Society 116:839–850.

Parzen, M. I., L. J. Wei, and Z. Ying. 1994. A resampling method based on pivotal estimating functions. Biometrika 81:341–350.

Persons, W. R., and R. V. Bulkley. 1984. Evaluation of the riverine cutthroat trout habitat suitability index model. U. S. Fish and Wildlife Service Biological Report 85(6):112–181.

Rao, C. R. 1988. Methodology based on the L_1-norm in statistical inference. Sankhya A 50:289–313.

Slauson, W. L., B. S. Cade, and J. D. Richards. 1991. User manual for BLOSSOM statistical software. U.S. Fish and Wildlife Service, National Ecology Research Center, Fort Collins, Colorado.

Welsh, A. H., R. J. Carroll, and D. Rupert. 1994. Fitting heteroscedastic regression models. JASA (Journal of the American Statistical Association) 89:100–116.

Wesche, T. A., C. M. Goertler, and W. A. Hubert. 1987. Modified habitat suitability index model for brown trout in southeastern Wyoming. North American Journal of Fisheries Management 7:232–237.

Wilkinson, L. 1990. SYSTAT: the system for statistic SYSTAT, Evanston, Illinois.

Received November 4, 19
Accepted June 29, 1995