ANL/EVS/R-10/4

The Effects of Overwinter Flows on the Spring Condition of Rainbow and Brown Trout Size Classes in the Green River Downstream of Flaming Gorge Dam, Utah

Environmental Science Division

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

Availability of This Report
This report is available, at no cost, at http://www.osti.gov/bridge. It is also available on paper to the U.S. Department of Energy and its contractors, for a processing fee, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone (865) 576-8401
fax (865) 576-5728
reports@osti.gov

ANL/EVS/R-10/4

The Effects of Overwinter Flows on the Spring Condition of Rainbow and Brown Trout Size Classes in the Green River Downstream of Flaming Gorge Dam, Utah

prepared by
A.K. Magnusson, K.E. LaGory, and J.W. Hayse
Environmental Science Division, Argonne National Laboratory

for
Western Area Power Administration
Colorado River Storage Project Management Center

May 2010

CONTENTS

CONTENTS (Cont.)

TABLES

FIGURES

FIGURES (Cont.)

NOTATION

ACRONYMS, INITIALISMS, AND ABBREVIATIONS

N	sample size
OOB	out-of-bag
P	probability of statistical significance
PIT	passive integrated transponder
Q25	25th percentile
Q75	75th percentile
r_s	Spearman rank correlation coefficient
Reclamation	Bureau of Reclamation
SD	standard deviation
SE	standard error
UDWR	Utah Division of Wildlife Resources
Western	Western Area Power Administration

VARIABLES USED IN THE ANALYSES

cv_dmean	coefficient of variation of mean daily flows
hgt1000	total hours with flows >1,000 cfs
hgt2000	total hours with flows >2,000 cfs
hgt3000	total hours with flows >3,000 cfs
hgt4000	total hours with flows >4,000 cfs
mean_dcv	mean daily coefficient of variation in flow
mean_ddelta	mean change in flow between days (absolute values)
mean_ddelta (%)	mean change in flow between days (%)
mean_dmax	mean of maximum daily flow
mean_dmean	mean of mean daily flow
mean_dmin	mean of minimum daily flow
mean_drange	mean daily range of flows
mean_dskew	mean daily skewness of flow volumes
mean_hdelta	mean hourly change of flows (absolute values)
mean_hdelta (%)	mean hourly change of flows (%)
mean_mdelta	mean change in flow between months (absolute values)
mean_mdelta (%)	mean change in flow between months (%)
med_dmean	median of mean daily flow
range_dmean	range of mean daily flows
ratioWR	relative change in WR from fall to spring
skew_dmean	skewness of mean daily flows
WR	relative weight

UNITS OF MEASURE

cfs	cubic foot (feet) per second
cm	centimeter(s)
CPUE	catch per unit effort (number of fish caught/h)
g	gram(s)
h	hour(s)
km	kilometers(s)
m	meter(s)
m^2	square meter(s)
m^3	cubic meter(s)
mm	millimeter(s)
s	second(s)

ACKNOWLEDGMENTS

This project was funded by the Western Area Power Administration Colorado River Storage Project Management Center under contract DE-AC02-06CH11357. The trout data used in this report were provided by Ryan Mosley and Lowell Marthe of the Utah Division of Wildlife Resources (UDWR), and the benthic macroinvertebrate data were provided by Mark Vinson at Utah State University (USU). Clay Perschon (UDWR), Roger Schneidervin (UDWR), David Speas (Bureau of Reclamation), and Scott Miller (USU) provided useful comments on an earlier draft of this report.

SUMMARY

Flaming Gorge Dam, a hydroelectric facility operated by the Bureau of Reclamation (Reclamation), is located on the Green River in Daggett County, northeastern Utah. Until recently, and since the early 1990s, single daily peak releases or steady flows have been the operational pattern of the dam during the winter period. However, releases from Flaming Gorge Reservoir followed a double-peak pattern (two daily flow peaks) during the winters of 2006–2007 and 2008–2009. Because there is little recent long-term history of double-peaking at Flaming Gorge Dam, the potential effects of double-peaking operations on trout body condition in the dam's tailwater are not known.

A study plan was developed that identified research activities to evaluate potential effects from winter double-peaking operations (Hayse et al. 2009). Along with other tasks, the study plan identified the need to conduct a statistical analysis of historical trout condition and macroinvertebrate abundance to evaluate the potential effects of hydropower operations. The results from analyses based on the combined size classes of trout (85–630 mm) were presented in Magnusson et al. (2008). The results of this earlier analysis suggested possible relationships between trout condition and flow, but concern that some of the relationships resulted from size-based effects (e.g., apparent changes in condition may have been related to concomitant changes in size distribution, because small trout may have responded differently to flow than large trout) prompted additional analysis of within-size class relationships. This report presents the results of analyses of three different size classes of trout (small: 200–299 mm, medium: 300–399 mm, and large: ≥400 mm body length). We analyzed historical data to (1) describe temporal patterns and relationships among flows, benthic macroinvertebrate abundance, and condition of brown trout (*Salmo trutta*) and rainbow trout (*Oncorhynchus mykiss*) in the tailwaters of Flaming Gorge Dam, and to (2) evaluate the relative importance of the effects of flow (i.e., flow volumes and flow variability), trout abundance (catch per unit effort [CPUE]), and benthic macroinvertebrate abundance on trout condition for different size classes of trout.

Magnusson et al. (2008) reported that flow volume in the overwinter period had negative effects on the size and weight of the combined size classes of rainbow trout at Little Hole and that flow variability had positive effects on the length, weight, and relative weight (WR) of the combined size classes of trout there and at Tailrace. The ratios of spring condition to fall condition over the period of record (ratioWR values) were used as an index to control for the initial condition of fish as they entered the winter period. These values differed from the WR values in their correlations with flows, and we found only one significant relationship with this variable (a positive relationship to hourly change in flow in Tailrace rainbow trout), suggesting that the relationships observed for WR could have been caused by differences in condition before the onset of winter.

The analyses of the combined size classes were based on Spearman rank correlations (Magnusson et al. 2008). Because of the relatively small number of years available for analysis, there was no adjustment of the statistics to the number of analyses, and some of the apparent relationships could have been statistically significant due to chance alone. For the size-class analyses, we use random forest regression analysis to build models of WR and ratioWR for each

trout group based on all of the independent variables, providing a measure of total explained variance and the relative importance of each independent variable. We also present Spearman rank correlations to allow comparison of the results to our previous analyses of the combined size classes.

The Spearman rank correlations of size classes suggested that the relationships between trout condition and overwinter flows varied not only between trout species and between locations, but also among size classes. The positive correlations between WR and flows for combined size classes (Magnusson et al. 2008) were observed only for medium and/or large trout and not for any of the small trout. The random forest models of WR also suggested that the various trout groups responded differently to changes in flows, CPUE, and total macroinvertebrates. The statistical model that best described WR was for large brown trout at Tailrace (43% variance explained due to a negative effect from CPUE and positive effects from flow variability). The remainder of the models explained at most 22% of the variation in WR. None of the models, however, explained any of the variation in ratioWR, which suggests that the relationships found for WR might have been related to the starting condition of the trout rather than to overwinter conditions.

The abundance of benthic macroinvertebrates at Tailrace and Little Hole was generally negatively correlated with increasing flow volume and flow variability (Magnusson et al. 2008). Total benthic macroinvertebrate abundance in January was also negatively correlated with the condition of medium Tailrace rainbow trout but positively correlated with ratioWR of small combined trout and of medium Little Hole trout. Benthic macroinvertebrate abundance did not, however, explain any of the variation in trout condition for any of the trout groups in the random forest analyses, suggesting that the importance to spring trout condition of overwinter benthic macroinvertebrate abundance was minor over the observed ranges in abundances.

The relationships identified in this report do not necessarily indicate cause and effects because several factors may have influenced the relationships, including changes in fishing practices over the study period and complex interactions between benthic macroinvertebrates and trout. The condition of large brown trout and medium rainbow trout at Tailrace and the condition of medium to large brown trout at Little Hole were negatively correlated with CPUE, and CPUE was one of the most important variables in models for describing the variability in WR for several of the medium and large trout groups. Moreover, it is possible that additional and stronger relationships may be identified once additional data have been collected. None of the observed relationships for WR seemed to be important after adjustment of the condition values to the starting condition of the trout and when sample size and the number of analyses were accounted for.

Gathering of additional information is needed to verify that the lack of response between flows and ratioWR in these analyses was not due to low sample size, interactions with unmeasured variables, and/or problems with the ratio values. The following studies or analyses could address these issues: (1) study of feeding behavior of different size classes at Tailrace and Little Hole, comparing macroinvertebrate contents in gut, drift, and benthos over the overwinter period; (2) comparison of ratioWR values calculated from average data (used in this report) with ratioWR calculated from individual trout (using passive induced transponder [PIT]-tagged trout);

(3) hydraulic model analyses to determine how flow affects the availability of different habitat types for the trout at Tailrace and Little Hole.

THE EFFECTS OF OVERWINTER FLOWS ON THE SPRING CONDITION OF RAINBOW AND BROWN TROUT SIZE CLASSES IN THE GREEN RIVER DOWNSTREAM OF FLAMING GORGE DAM, UTAH

Prepared by

A. Katarina Magnusson, Kirk E. LaGory, and John W. Hayse
Environmental Science Division
Argonne National Laboratory

1 INTRODUCTION

Flow volume and flow patterns strongly affect the distribution, assemblage structure, and condition of native and nonnative fishes in regulated rivers (Gore 1996; Marchetti and Moyle 2001; Osmundson et al. 2002). These relationships are especially apparent downstream of some hydroelectric facilities, where natural flow regimes have been altered by changing the seasonal flow volumes, introducing greater daily and seasonal variability in flows, altering seasonal temperature regimes, and changing suspended sediment levels. In some cases, changes in water temperature and turbidity have resulted in shifts from conditions suitable for supporting warmwater fish communities to conditions that favor the establishment of coldwater fish communities.

Flaming Gorge Dam is a hydroelectric facility that is located on the Green River (a tributary of the Colorado River) in Daggett County, northeastern Utah. Construction of the dam, which was completed in 1964, resulted in the formation of Flaming Gorge Reservoir, which has a surface elevation of approximately 1,840 m above mean sea level. Because of the hypolimnetic releases from the dam, the tailwaters of the Green River were transformed from a warmwater ecosystem to a coldwater ecosystem that supports a highly regarded trout fishery (Reclamation 2005). The dam has three turbines with a maximum combined release capacity of approximately 4,600 cubic feet per second (cfs) (130 m^3/s). The maximum combined release capacity is now about 4,200 cfs (119 m^3/s) following recent upgrades to the turbines. Agreement between the Bureau of Reclamation (Reclamation) and the state of Utah has resulted in year-round minimum releases from Flaming Gorge Reservoir of 800 cfs (23 m^3/s) to maintain the trout fishery. In order to improve conditions for trout growth and survival during summer months, a selective water withdrawal structure was installed at the dam in 1978 that allows warmer epilimnetic water to be released through the turbines.

There have been changes in Flaming Gorge Dam operations since its construction in 1962 (Muth et al. 2000). Prior to 1984, Flaming Gorge Dam was operated to closely match electrical power demands. This typically resulted in relatively high discharges during the winter and summer months and high within-day fluctuations in flow. Since 1984, operations have been modified to protect downstream resources. Flow and temperature recommendations were developed to protect endangered fish and their habitats in the middle Green River downstream of the Yampa River confluence (USFWS 1992; Muth et al. 2000; Reclamation 2005, 2006). These changes have resulted in higher spring peak flows, lower base flows from summer through winter, and moderation of within-day fluctuations, especially during summer and fall.

The match of power generation with electrical demand increases the market value of the power produced. In recent years, single daily peak releases or steady flows have typified winter operations at Flaming Gorge Dam. At the request of Western Area Power Administration (Western), a double-peaking release pattern at Flaming Gorge Dam that featured peaks in the morning and evening that continued to meet flow requirements for endangered fish in the middle Green River was implemented during the winters of 2006–2007, 2008–2009, and 2009–2010 by Reclamation. The potential effects of double-peaking operations on trout condition in the dam's tailwaters are not known, and concerns have been expressed that double-peaking operations could have negative effects on the growth, survival, and reproduction of trout. Consequently, a study plan that identified research activities to evaluate the potential effects of winter double-peaking was developed (Hayse et al. 2009).

Along with other tasks, the study plan identified the need to conduct a statistical analysis of existing data on trout condition and benthic macroinvertebrate abundance to evaluate potential effects of hydropower operations. In Magnusson et al. (2008), we analyzed data collected from 1990–2006 to (1) describe temporal patterns and relationships among flows, benthic macroinvertebrate abundance, and condition of brown trout (*Salmo trutta*) and rainbow trout (*Oncorhynchus mykiss*) in the Flaming Gorge Dam tailwaters and (2) evaluate the degree to which flow characteristics (i.e., flow volumes and flow variability) and benthic macroinvertebrate abundance may affect trout condition. Trout data aggregated by year with all size classes combined into a single annual mean value were provided to us by the Utah Division of Wildlife Resources (UDWR).

The two sites for which trout and benthic macroinvertebrate data were available are known as Tailrace and Little Hole; both sites are approximately 1 km in length. The Tailrace site extends about 1 km downstream from the first boat ramp below the dam to the confluence of the Green River and Pipe Creek. This reach consists primarily of deeper runs, although there are some short riffle sections. Macroinvertebrate samples in the Tailrace section were collected just upstream from the Tailrace boat ramp (Vinson et al. 2006). The Little Hole reach extends downstream from the upstream-most Little Hole boat ramp to Grasshopper Island. The Little Hole site consists of riffles, runs, and pools of various depths but has less run habitat than the Tailrace site. Samples of benthic macroinvertebrates for this site were collected just upstream of the Little Hole boat ramp (Vinson et al. 2006).

Because fish in temperate zones may experience a decline in condition over the winter due to colder water temperatures, reduced food supply, and behavioral changes, overwinter declines in condition following a winter with double-peaking operations would not be expected to be solely the result of double-peaking. To compensate, we proposed two approaches: (1) a comparison of spring condition data as measured by relative weight [WR] in years that featured double-peaking to spring condition data from other years and (2) an examination of the ratio of spring condition to fall condition [ratioWR] over the period of record. Hayse et al. (2009) hypothesized that the ratio would normalize condition data to more accurately reflect the effect of winter operations because it would be based on the starting condition of trout as they enter a particular winter. Such a ratio could be used to evaluate whether there is an incremental effect of double-peaking operations on the condition of overwintering trout.

In Magnusson et al. (2008), we summarized the relationships between flow and trout condition during a period of single daily peak winter operations. Our analyses indicated that flow volume was not correlated with the spring condition of Tailrace brown and rainbow trout or with the condition of Little Hole brown trout, but was negatively correlated with the spring length and weight and positively correlated with relative weight of Little Hole rainbow trout (Magnusson et al. 2008). Flow variability was positively correlated with the length, weight, and relative weight of trout at both locations. However, none of these relationships were significant when the condition values were adjusted to the starting condition of the trout (ratioWR). Except for a positive correlation between within-day flow variability and Tailrace rainbow trout ratioWR, no flow variables were correlated with ratioWR. The relationships between benthic macroinvertebrate abundance and trout condition revealed no conclusive patterns, but January abundance of benthic macroinvertebrates was negatively correlated with WR for Tailrace rainbow trout and positively correlated with ratioWR for Little Hole brown and rainbow trout. Since trout of different sizes may respond differently to increased levels of flows and to food availability and there were observed changes in the size structure of the populations during the study period, we continue this investigation in the present report by examining some of these relationships for three different size classes of trout.

In this report, we specifically evaluate the relationships among fall and winter flows, fall and winter macroinvertebrate abundance, and the condition of three size classes of trout (small: 200–299 mm; medium: 300–399 mm; and large: ≥400 mm) in the spring. Due to the limited number of sampled years, however, the relationships that are identified in this report should be regarded only as potentially important relationships that warrant further examination once additional data have been collected. This information also may serve as a baseline to which the effects of potential future double-peaking flows can be compared.

Section 2 of the report presents and evaluates the temporal patterns found in existing data on fall and winter flows (1989–2006), macroinvertebrate abundance (1994–2006), and the body condition of trout in spring (1990–2002 and 2006). Section 3 evaluates statistical relationships among variables derived from these data. Section 4 provides overall conclusions and recommendations. Additional detailed information including the results of statistical analyses and a detailed description of our random forest analysis are presented in appendices.

2 TEMPORAL PATTERNS IN FALL AND WINTER FLOWS, MACROINVERTEBRATE ABUNDANCE, AND TROUT CONDITION

This section provides an overview of the temporal patterns relationships found in data from 1990 through 2006 on condition of trout size classes, macroinvertebrate abundance, and flows in the tailwaters of Flaming Gorge Dam. This time period encompasses the years for which information is available for trout from spring and fall electrofishing surveys conducted by UDWR. Macroinvertebrate data are available from 1994 through 2006. A variety of hydrologic conditions and ecological changes occurred in the study area over this time period. For example, there were (1) consecutive years of severe drought (2000 to 2005) and years with moderately wet conditions (1997 and 1999); (2) annual operations with extended periods of steady base flows and single-daily peak operations; (3) an influx of fine sediment to the river after the Mustang Ridge wildfire in 2002; (4) an invasion of the exotic New Zealand mud snail (*Potamopyrgus antipodarum*) starting in 2002; (5) large increases in wild-spawned brown trout, with a subsequent switch in numerical dominance from rainbow to brown trout; and (6) an increase in the numbers of anglers and fishing pressure.

2.1 METHODS

Annual patterns of brown and rainbow trout size and body condition variables are summarized graphically in this section. Annual patterns of total abundance of benthic macroinvertebrates and flow variables were summarized in Magnusson et al. (2008). The sources and types of data used for analyses of trout, macroinvertebrates, and flows are summarized below.

2.1.1 Trout Data

All trout data evaluated in this report were collected and provided by UDWR. Electrofishing surveys of trout in the Flaming Gorge tailwaters were conducted by UDWR personnel each fall and spring from 1990 through 2006, with the exception of the springs of 2003, 2004, and 2005 when these data were not collected. Data provided by UDWR for the analyses included catch per unit effort (CPUE, defined as the number of trout caught per hour; Figure 1a), lengths, weights, and calculated condition factor (WR) for individual brown and rainbow trout segregated by sample collection location (Tailrace or Little Hole). Data were separated into three different size classes of trout (small: 200–299 mm, medium: 300–399 mm, and large: ≥400 mm; Table 1, Figures 1b, 2a, and 2b). In Magnusson et al. (2008), the trout data that were analyzed consisted of single annual mean values for all captured trout, including trout smaller than 200 mm (>84 mm). Trout smaller than 200 mm were absent for some of the years and are therefore not included in the analyses of size classes in this report. Small (200–299 mm) brown trout at Tailrace and small rainbow trout at Little Hole were also captured in low numbers (<100 fish captured in total over 14 years; Table 1, and Figures 1b and 2b), and therefore all small trout are pooled together (combined species and locations) in the analyses of size classes in this report.

TABLE 1 Number of Trout Caught in Each Size Class during Spring Electrofishing Surveys, 1990–2006

	Brown Trout at Tailrace				Brown Trout at Little Hole				Rainbow Trout at Tailrace				Rainbow Trout at Little Hole			
	<200 mm	200–299 mm	300–399 mm	≥400 mm	<200 mm	200–299 mm	300–399 mm	≥400 mm	<200 mm	200–299 mm	300–399 mm	≥400 mm	<200 mm	200–299 mm	300–399 mm	≥400 mm
1990	6	0	3	22	83	2	22	139	24	27	284	384	4	4	185	285
1991	1	8	23	19	17	19	21	175	20	10	130	136	0	3	76	101
1992	1	0	16	74	18	27	79	237	5	7	260	466	1	0	69	105
1993	0	1	6	70	10	3	39	290	31	4	107	298	5	0	36	101
1994	4	2	7	121	36	6	57	333	31	16	240	415	0	4	167	109
1995	4	11	18	142	6	103	97	285	21	90	433	317	0	12	120	68
1996	0	8	50	154	4	30	306	298	0	124	359	168	1	11	95	35
1997	3	3	37	173	6	22	293	381	2	26	382	246	0	4	46	63
1998	3	6	43	187	13	3	175	444	2	79	364	351	0	15	61	40
1999	0	11	25	122	2	4	84	352	7	170	269	273	0	3	95	49
2000	4	11	63	146	13	6	135	403	4	156	273	110	0	8	132	41
2001	1	10	98	215	0	14	115	361	0	23	486	208	0	4	70	52
2002	1	0	50	149	1	3	48	291	0	18	194	162	0	3	75	44
2003																
2004																
2005																
2006	0	6	293	100	3	18	400	102	1	15	309	30	0	21	85	5
Total fish	28	77	732	1,694	212	260	1,871	4,091	148	765	4,090	3,564	11	92	1,312	1,098
Number of years with fish	10	11	14	14	13	14	14	14	11	14	14	14	4	12	14	14

Length and weight from the combined size-class data were analyzed in Magnusson et al. (2008), but are not included in the analyses here since body length was used to separate the trout into size classes and weight was correlated with the length of the trout.

Two trout condition variables were used in the size-class analyses:

- Relative weight in fall and spring (WR): $W/W_i \times 100$, where W_i = “standard weight” calculated from previously established species-specific length-weight regression equations (Murphy et al. 1991); and

- Ratio of spring to fall relative weight values (ratioWR) calculated by dividing the spring WR value by the fall WR value of each size class.

Relative weight was calculated from the weight of individual trout (summarized in Table A.1), whereas the spring to fall ratio of WR was calculated from mean WR values of all

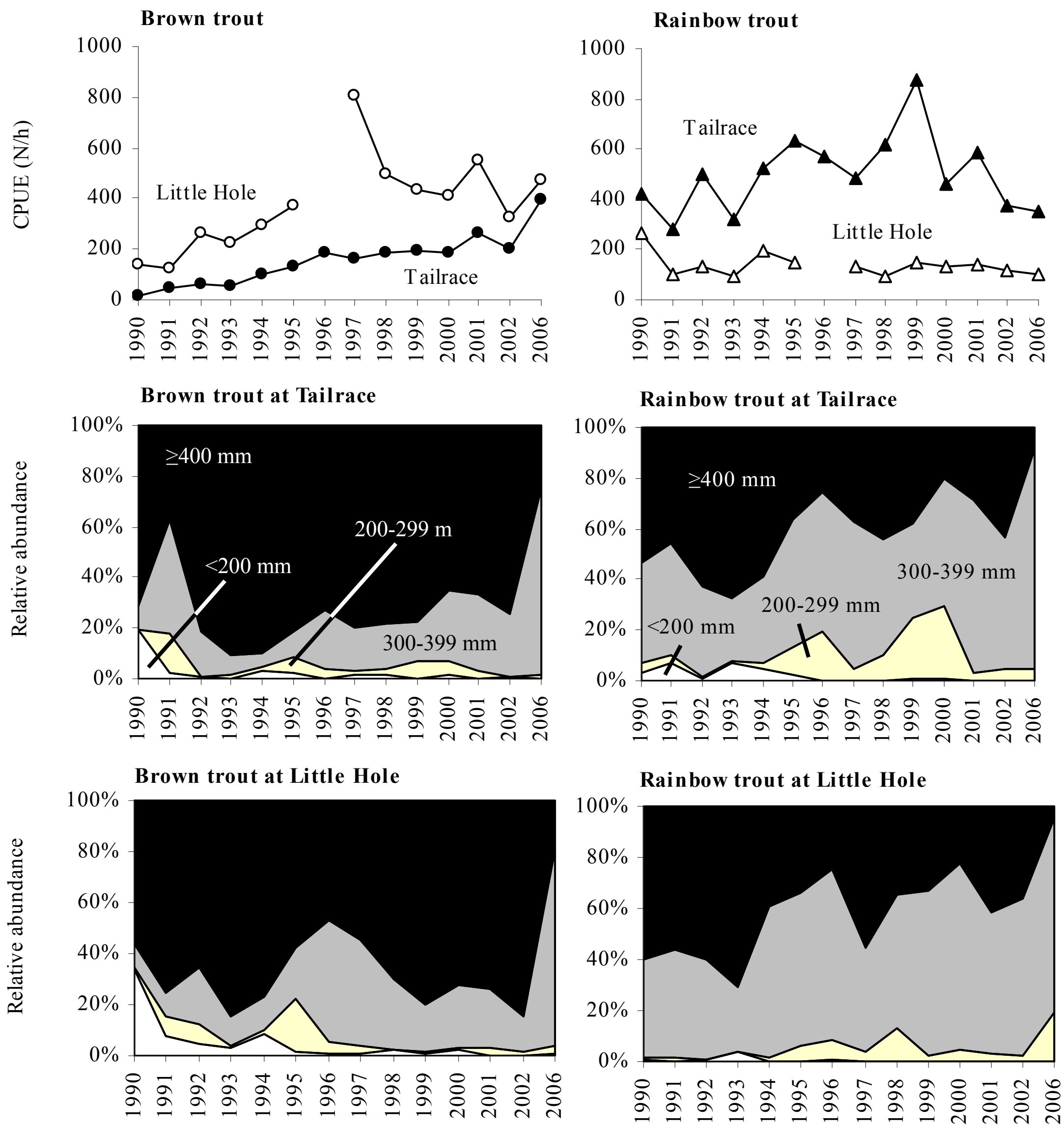


FIGURE 1 Total Abundance Measured as (a) CPUE and (b) Relative Abundance of Trout Size Classes (<200 mm, 200–299 mm, 300–399 mm, and ≥400 mm), 1990–2006 (No data were collected in spring 2003–2005, and no CPUE data were available for the Little Hole site in 1996.) Source: Modified from Magnusson et al. 2008

(a)

(b)

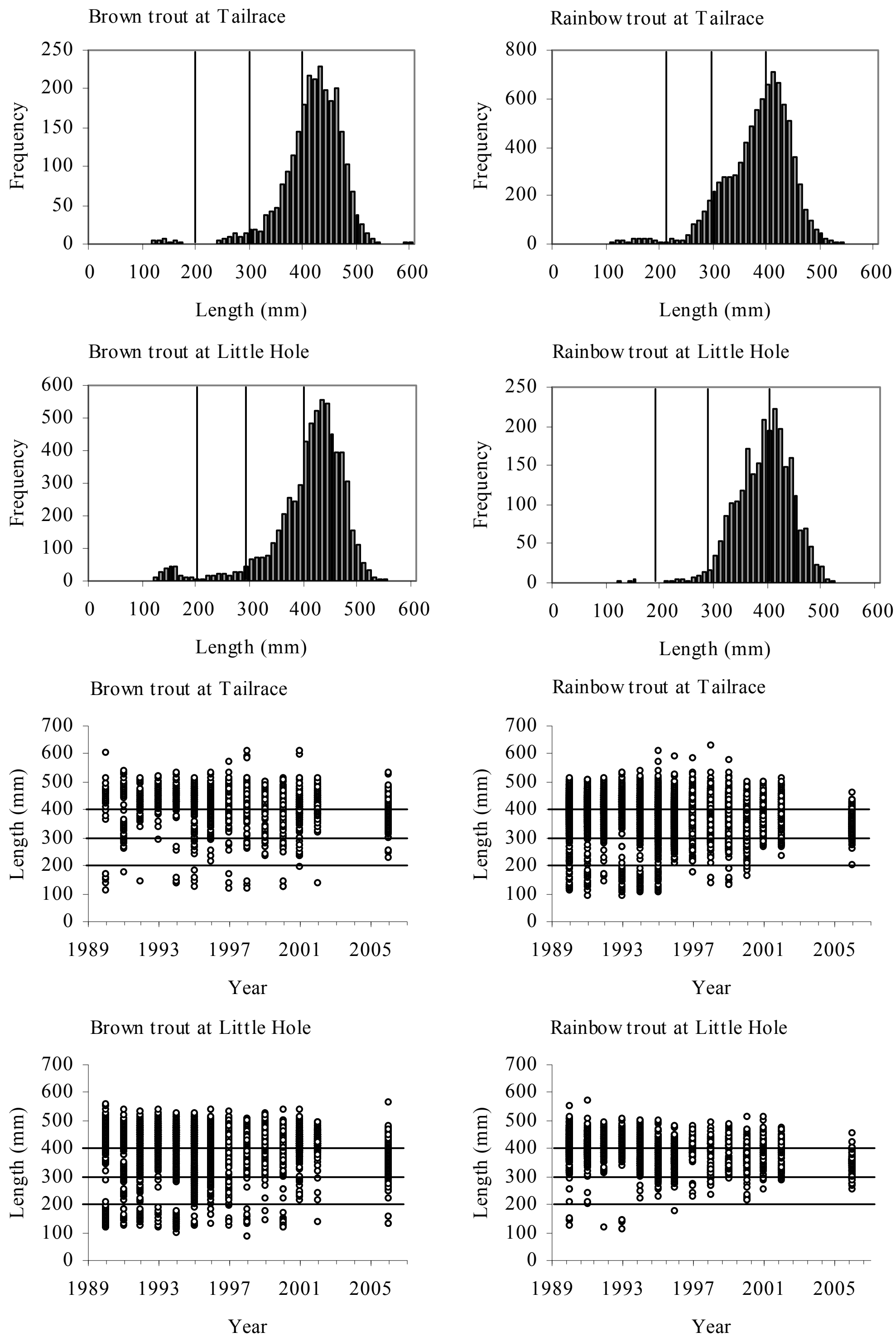


FIGURE 2 (a) Length Frequency Distribution and (b) Body Length of Individual Trout Captured during Spring Electrofishing Surveys, 1990–2006 (Vertical or horizontal lines separate size classes.)

trout within a specific size class. RatioWR could not be calculated for individual trout because fish were not marked in this study.

2.1.2 Macroinvertebrate Data

Benthic macroinvertebrate abundance data from the Flaming Gorge tailwaters were provided by M. Vinson at Utah State University. Macroinvertebrates were collected four times annually using Hess nets in riffle habitats located upstream of the Tailrace and Little Hole electrofishing sites in January, April, July, and September or October[1] from 1994 to 2006 (see Vinson et al. 2006 for a description of the sampling protocol). The data available for our analysis consisted of mean values, reported as estimated densities (i.e., number of individuals/m^2), from eight pooled samples for each site and season. We only used macroinvertebrate data from the overwinter period for the analyses in this report (i.e., total abundances in October, January, and April collections, and the total mean over the three months; Table 2) because it was assumed that the condition of trout in April is most likely affected by food availability during the period immediately preceding April. New Zealand mud snails first appeared at the sites in 2002 but were kept in the analyses since their abundance was negligible compared with the total abundance (they never exceeded 3.2% of the total abundance from 2002 to 2006). Since monitoring of macroinvertebrates did not start until 1994, the number of years for which such data were available (N=10 for January and April samples from 1994 to 2002 and 2006; N=9 for October samples and total macroinvertebrates, from 1995 to 2002 and 2006) is lower than the total number of years for which trout electrofishing data were available (N=14; 1990–2002 and 2006). Since trout tend to feed on the taxa that are most abundant to them (Vinson et al. 2006), we used only the total abundance of benthic macroinvertebrates rather than data for each taxon in our analyses. Seasonal abundance of benthic macroinvertebrate taxa at Tailrace and Little Hole during the study period can be found in our previous report (Magnusson et al. 2008).

2.1.3 Flow Data

Flaming Gorge Dam release data from 1989 to 2006 were provided by Western and were based on electrical generation data and non-power releases for those years. The values for 20 different flow variables were calculated with SAS statistical software (Version 9.1; SAS Institute Inc., 2002) for the fall through winter period (October through April) of each year by using hourly and daily mean flow values (Table 3). Each variable is described in Magnusson et al. (2008).

Flow variables were segregated into categories considered descriptive of various aspects of volume or variability. In total, 8 variables were used to evaluate flow volume and 12 variables to evaluate flow variability (Table 3). Variables describing flow variability were further divided

[1] Samples collected in September or October are referred to as “October” samples in this report because relatively more of the samples in this period were taken in October and samples from both months represent the same approximate time period.

TABLE 2 Benthic Macroinvertebrate Variables Used in the Analysis

Macroinvertebrate Variables	Definition
October macroinvertebrates[a,b]	Mean number/m^2 of all macroinvertebrate taxa in October samples
January macroinvertebrates	Mean number/m^2 of all macroinvertebrates taxa in January samples
April macroinvertebrates	Mean number/m^2 of all macroinvertebrate taxa in April samples
Total macroinvertebrates[b]	Mean number/m^2 of all macroinvertebrates taxa in October,[a] January, and April samples

[a] October macroinvertebrate samples for most years were collected in October, but samples for 2001–2006 were collected in late September.

[b] October and total macroinvertebrate samples had one replicate less than January and April samples because October samples were correlated with trout in the following spring, 1995–2006.

Source: Magnusson et al. 2008

according to whether they represented within-day variability (5 variables), between-day variability (2 variables), within-season variability (3 variables), or between-month variability (2 variables). Because it was anticipated that these variables would not be fully independent measures of flow conditions, we also calculated correlations among the flow variables (Magnusson et al. 2008).

2.2 RESULTS AND DISCUSSION

2.2.1 Trout Abundance and Condition

Rainbow trout was the most abundant trout species at Tailrace, whereas brown trout was the most abundant species at Little Hole (Figure 1a). Brown trout were, on average, larger in size than the rainbow trout (71% of the Tailrace brown trout and 66% of the Little Hole brown trout were ≥400 mm, compared with 41% of the rainbow trout at both locations; Figures 1b, 2a, and 2b). CPUE varied among years within all trout groups (Figure 3). Small trout CPUE ranged from 0 to 207 fish per hour over the study period, and, as averaged over the period, was higher for Tailrace rainbow trout (49 fish/hr) than for Little Hole rainbow trout (5 fish/hr) and brown trout at both locations (Tailrace: 5 fish/hr; Little Hole: 14 fish/hr; Figure 3). Medium trout CPUE ranged from 2 to 396 fish/hr, and, as averaged over the period, was greater for Tailrace rainbow trout (241 fish/hr) than for Little Hole rainbow trout (73 fish/hr) and brown trout (Tailrace: 48 fish/hr; Little Hole: 109 fish/hr). Large trout CPUE ranged from 5 to 440 fish/hr, and, as averaged over the period, was greater for Little Hole brown trout (245 fish/hr) and Tailrace rainbow trout (201 fish/hr) than for Tailrace brown trout (102 fish/hr) and Little Hole rainbow trout (60 fish/hr).

TABLE 3 Variables Used to Describe Flow Volume and Flow Variability from October through April in the Green River below Flaming Gorge Dam

Variable Category	Variable Name	Definition
Flow Volume		
	med_dmean	Median of mean daily flow
	mean_dmean	Mean of mean daily flow
	mean_dmin	Mean of minimum daily flow
	mean_dmax	Mean of maximum daily flow
	hgt1000	Total hours with flows >1,000 cfs
	hgt2000	Total hours with flows >2,000 cfs
	hgt3000	Total hours with flows >3,000 cfs
	hgt4000	Total hours with flows >4,000 cfs
Flow Variability		
Within-day	mean_hdelta	Mean hourly change of flows (absolute values)
	mean_hdelta%	Mean hourly change of flows (%)
	mean_drange	Mean daily range of flows
	mean_dcv	Mean daily coefficient of variation in flow[a]
	mean_dskew	Mean daily skewness of flows[b]
Between-day	mean_ddelta	Mean change in flow between days (absolute values)
	mean_ddelta%	Mean change in flow between days (%)
Within-season	range_dmean	Range of mean daily flows
	cv_dmean	Coefficient of variation of mean daily flows[a]
	skew_dmean	Skewness of mean daily flows[b]
Between-month	mean_mdelta	Mean change in flow between months (absolute values)
	mean_mdelta%	Mean change in flow between months (%)

[a] Coefficient of variation (CV) = ratio of the standard deviation to the mean.

[b] Skewness = lack of symmetry in the data distribution for flows. Zero skewness has the median value equal to the mean. Positive skewness has the median value lower than the mean (i.e., the distribution is skewed to the right and has a longer right tail than expected for a normal distribution); this would suggest that there are more hours or days with low flows. A data distribution with a negative skewness has a median value greater than the mean (i.e., the distribution is skewed to the left and has a longer left tail). A negatively skewed distribution for flow data would indicate there were more hours or days with high flows.

Source: Magnusson et al. 2008

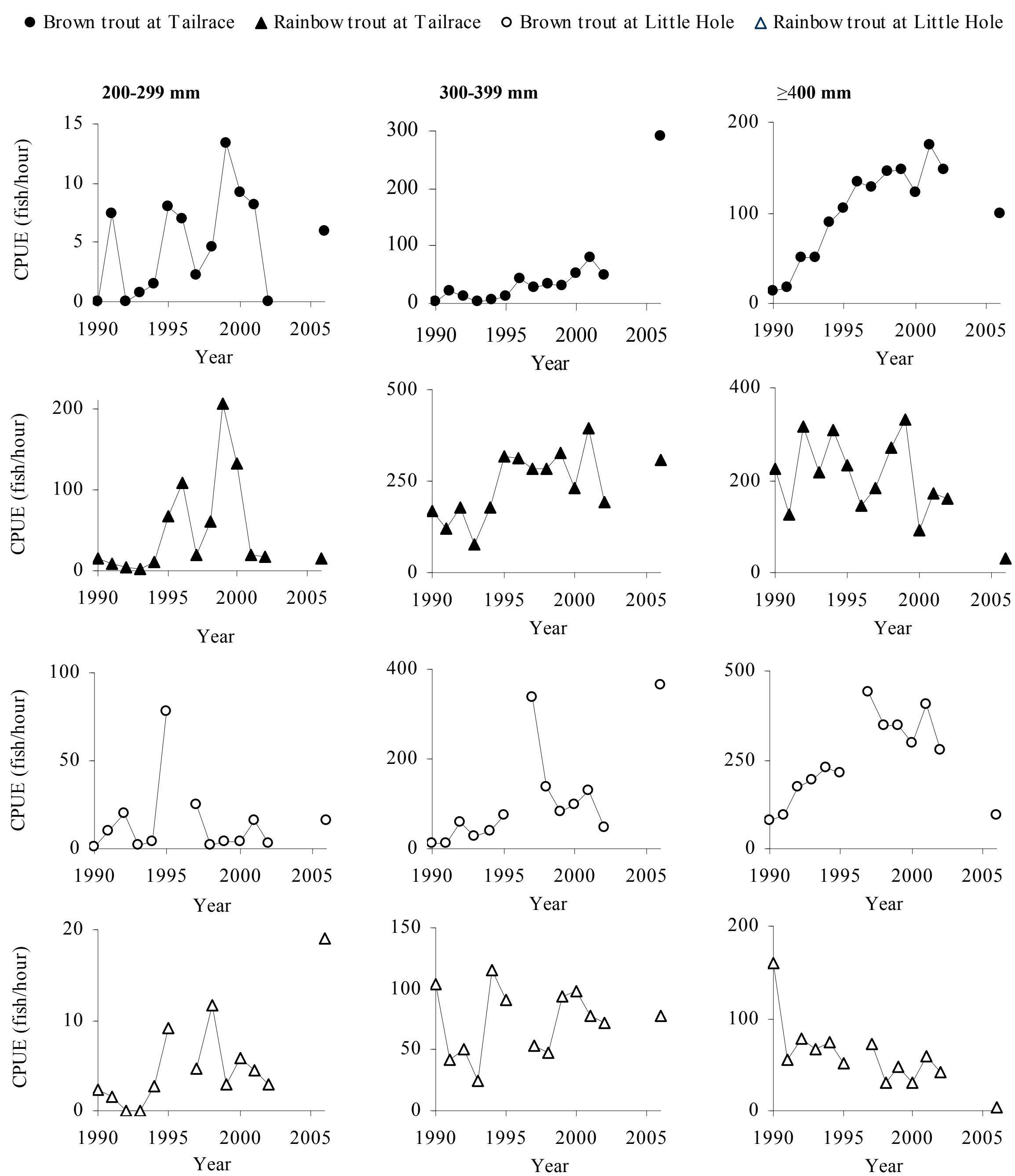


FIGURE 3 Spring Abundance (CPUE) of Trout Size Classes (200–299 mm, 300–399 mm, and ≥400 mm), 1990–2006 (No data were collected in spring 2003–2005, and CPUE was not determined for Little Hole in 1996. Note differences in scale of the *y*-axes.)

Mean WR values varied among years, and the difference between the lowest and highest values during the study period was 15 to 32 units depending on trout group (Table A.1, Figure 4). Mean WR was, as averaged over the study period, 104.2 for rainbow trout and 101.9 for brown trout; 102.0 at Tailrace and 104.0 at Little Hole; 95.8 for small trout, 107.1 for medium trout, and 106.2 for large trout.

RatioWR values also varied among years, and the difference between the lowest and the highest values during the study period was 0.10 to 0.36 units depending on trout group (Figure 5). RatioWR was, as averaged over the study period, 1.01 for rainbow trout and 0.94 for brown trout; 0.97 at Tailrace and 0.98 at Little Hole; 0.97 for small trout and 0.98 for medium to large trout. A ratioWR value above 1.0 (observed for medium and large rainbow trout) suggests that the condition within the population improved over the winter, whereas a ratioWR value below 1.0 (observed for brown trout) suggests that the trout decreased in condition over the winter (presumably because brown trout are winter spawners). The ratioWR values were generally not correlated with the WR values for the different trout groups (r_S=–0.46 to 0.56, P>0.05), with the exception of medium Little Hole rainbow trout (r_S=–0.61, P<0.05).

The WR of large Tailrace brown trout, medium Tailrace rainbow trout, and medium to large Little Hole brown trout was negatively correlated with CPUE (Table 4, Figure 6), suggesting that body condition for these trout could be density-dependent and possibly related to increased competition for resources. CPUE was not correlated with Little Hole rainbow trout WR for any size class, nor was it correlated with ratioWR for any of the trout groups (Table 4).

2.2.2 Benthic Macroinvertebrate Abundance

Total macroinvertebrates in the overwinter period (average over October, January, and April) ranged from 9,200/m^2 in 1998 to 26,000/m^2 in 2004 at Tailrace and from 6,700/m^2 in 1999 to 15,000/m^2 in 2001 at Little Hole. The dominant taxa consisted of amphipods (58% at Tailrace and 70% at Little Hole) and dipterans (28% of the total abundance at Tailrace and 26% at Little Hole). Ephemeropterans, coleopterans, and New Zealand mud snails together made up less than 6% of the total benthic community. The different taxa are not examined in this report. For further information about the benthic macroinvertebrates, see Magnusson et al. (2008) and Vinson et al. (2006).

Flow Volume

Eight different variables were used to measure flow volume (Table 3), all of which were more or less intercorrelated; four variables measured daily minimum, maximum, median, and mean flow, and the remaining four variables measured the total number of hours with flows above certain thresholds (1,000, 2,000, 3,000, and 4,000 cfs). Most of these variables had higher values in the middle of the study period (representing wetter hydrologic conditions) than early or late in the period. Mean daily flow volume (mean_dmean) ranged from 800 cfs in 2002 to 2,800 cfs in 1998.

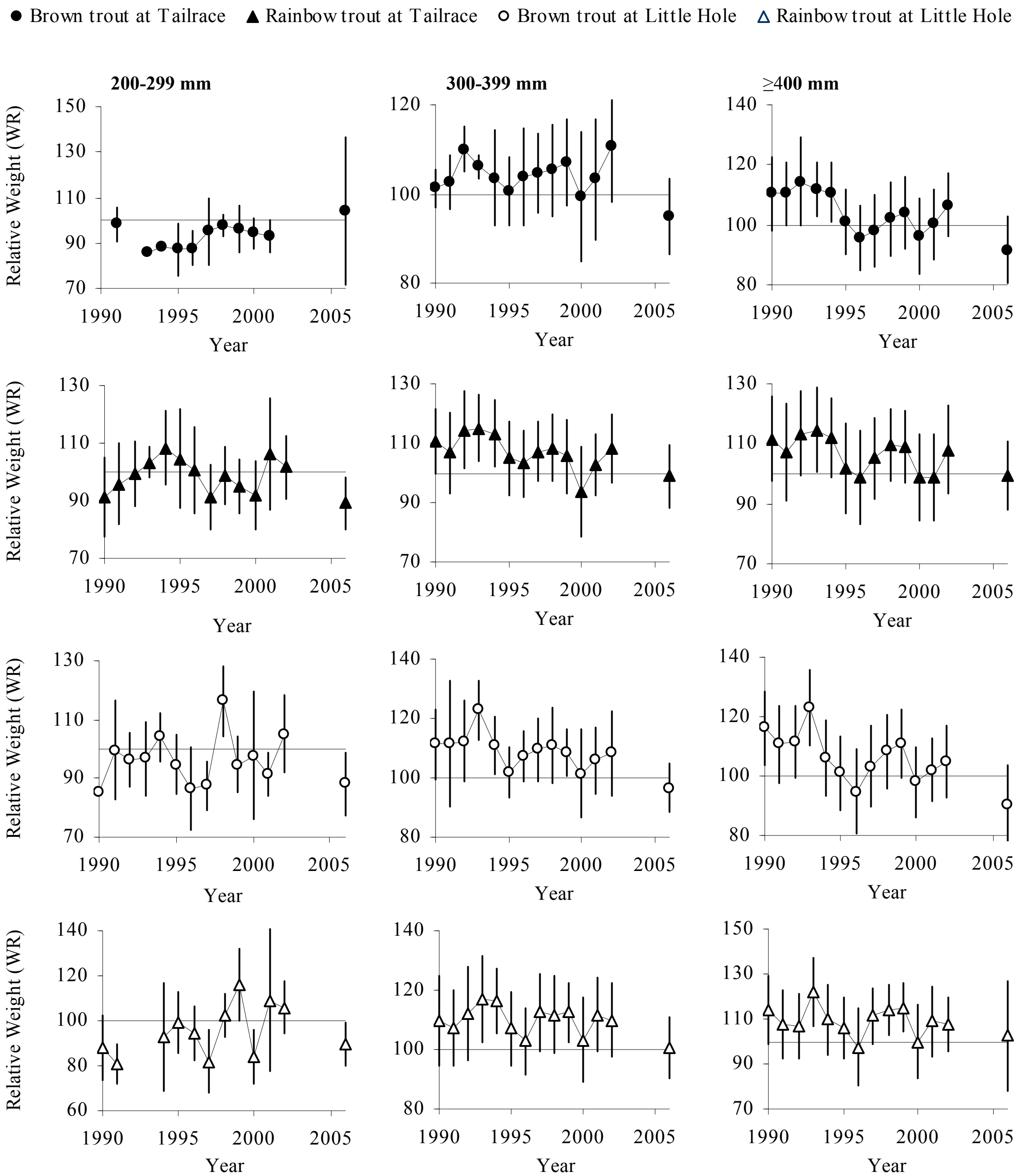


FIGURE 4 Mean Relative Weight (± SD) of Trout Size Classes, 1990–2006 (No data were collected in spring 2003–2005. Note differences in scale of the *y*-axes.)

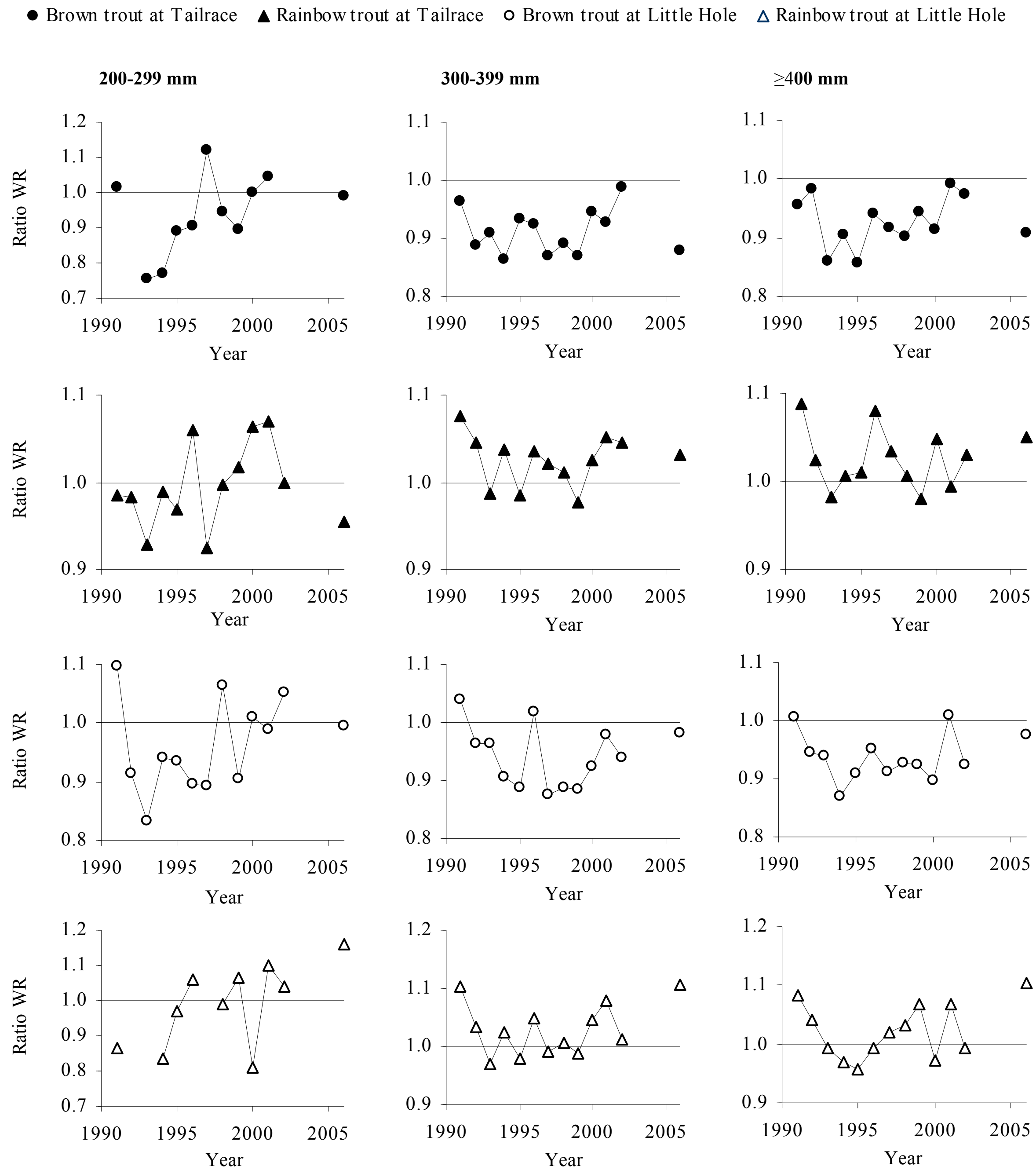


FIGURE 5 RatioWR of Trout Size Classes, 1990–2006 (No data were collected in spring 2003–2005. Note differences in scale of the *y*-axes; standard errors could not be calculated because the ratioWR values were calculated from average WR values in fall and spring.)

TABLE 4 Statistically Significant Correlations between Trout Condition (WR and ratioWR) and CPUE[a]

	Trout WR and RatioWR Correlated with CPUE										
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace			Brown Trout at Little Hole			Rainbow Trout at Little Hole	
Variable	200–299 mm	300–399 mm	≥400 mm	200–299 mm	300–399 mm	≥400 mm	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
$CPUE_{Total}$	none	none	-WR	none	-WR	none	none	-WR	-WR	none	none

[a] The sign (+ or -) preceding a variable indicates whether the Spearman rank correlation is positive or negative. The significance of the relationship is indicated by the number of signs shown (1: $P<0.05$, 2: $P<0.01$, 3: $P<0.001$). None = none of the trout variables were significantly correlated with the CPUE; $CPUE_{Total}$ = CPUE for combined trout species and size classes. $CPUE_{Total}$ is averaged across the two locations in the analyses for combined small trout in the first column. WR = relative weight; ratioWR = ratio of spring and fall WR.

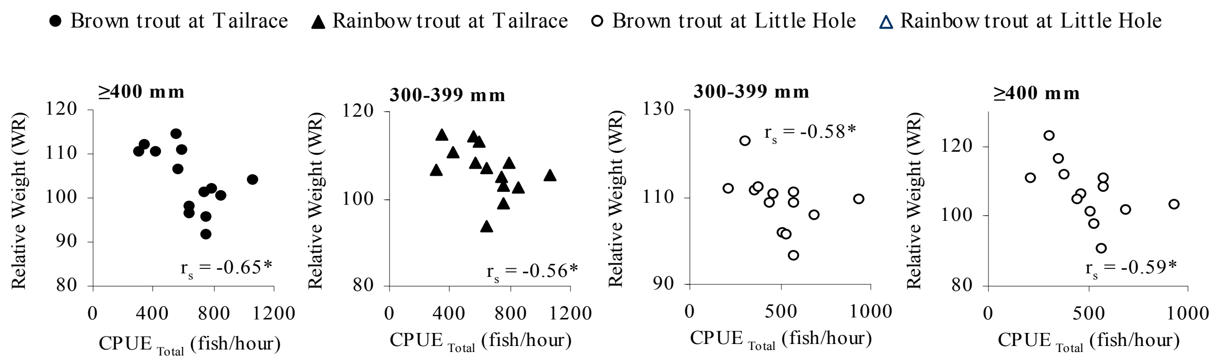


FIGURE 6 Statistically Significant Relationships between WR and CPUE, 1990–2006 (*=*P*<0.05)

Flow Variability

Five variables were used to measure different aspects of within-day flow variability (Table 3); three variables measured the magnitude of the daily change in flow, including hourly change in flow within a day (mean_hdelta and mean_hdelta%) and total range of flow in a day (mean_drange); one variable measured the variability around the daily flow mean (mean_dcv) with values increasing with peak duration and peak magnitude; and one variable measured the asymmetry of the distribution of flow values within a day (mean_dskew) with values increasing with decreasing peak duration regardless of peak magnitude. The first four variables were highly correlated with each other ($r_S \geq 0.87$, $P<0.05$) and had high values in 1992 and 1994 and low values in 1995, 1998, 2001, and 2002. Mean_dskew was not correlated with the other four variables ($P>0.05$) and had high values in 1991 and 1993 and the lowest (negative) values in 1998 and 2006.

Between-day flow variability was measured with two related variables (mean_ddelta and mean_ddelta%; $r_S=0.90$, $P<0.05$), which measure how much daily mean flow changes from one day to the next day (Table 3). These variables had high values in the beginning of the study period (mean_ddelta >100 cfs, 1990–1992 and 1994) and low values in the end of the study period (mean_ddelta <10 cfs, 2000–2002 and 2006).

Three variables were used to measure within-season flow variability, or overall variability of daily mean flows, namely range_dmean, cv_dmean, and skew_dmean (Table 3). Range_dmean and cv_dmean were intercorrelated ($r_S=0.89$, $P<0.05$); range_dmean peaked in 1997 (close to 3,000 cfs) whereas cv_dmean peaked in 1990 (39%), and both variables had low values in 2000–2006 (<500 cfs and <10%, respectively). Skew_dmean measures whether the overall pattern in mean daily flow volume over the winter period was stable or more erratic. Skew_dmean peaked in 1993 (value of 8.1) and had a low negative value in 2006 (value of –2.7).

The two between-month flow variability variables, mean_mdelta and mean_mdelta% ($r_S=0.95$, $P<0.05$), measured the magnitude of change in mean flow from month to month (Table 3). They had high values in 1994, 1996–1998 (mean_mdelta >400 cfs) and low values in 1993 and 2000–2006 (mean_mdelta <100 cfs).

3 RELATIONSHIP OF SPRING CONDITION OF TROUT SIZE CLASSES TO OVERWINTER FLOW AND MACROINVERTEBRATE ABUNDANCE

Our previous analyses of the combined size classes of trout at Tailrace and Little Hole indicated that trout condition (WR) was generally not correlated with flow volume (except for a positive correlation between WR for Little Hole rainbow trout and the number of hours with flows above 3,000 cfs) but positively correlated with some measures of flow variability in the fall and winter (Magnusson et al. 2008). These relationships were, however, not statistically significant after adjustment of the WR values to the starting condition of the trout (ratioWR). With the exception of a positive correlation between Tailrace rainbow trout ratioWR and hourly change in flow, ratioWR was generally not correlated with any flow variables. The length and weight of rainbow trout at Little Hole were, however, negatively correlated with flow volume and the length and/or weight of rainbow trout at both locations and were positively correlated with flow variability for brown trout at Little Hole. Trout variables were also correlated with the abundance of benthic macroinvertebrates: WR of rainbow trout at Tailrace was negatively correlated with January abundances of macroinvertebrates; the length of brown trout was positively correlated with October abundances of macroinvertebrates at both locations but negatively correlated with January abundances at Tailrace; and ratioWR of both trout was positively correlated with January abundance at Little Hole (Magnusson et al. 2008).

These apparent relationships among flow, benthic macroinvertebrate abundance, and trout condition could be size-dependent. The condition of 300–350 mm fish in the Green River has been reported to be more strongly related to variation in food and temperatures compared to small and large trout (Filbert and Hawkins 1995). In addition, fish larger than 305–405 mm tend to consume more algae, a low quality food, than smaller trout, and the larger trout may therefore be more food-limited than the smaller trout (McKinney and Speas 2001).

Analyses of WR and ratioWR for separate size classes would control for the possible effect of size on these correlations. In this section, we present statistical evaluations of the relationships among (1) condition of size classes in the spring and overwinter flows; (2) condition of size classes and benthic macroinvertebrate abundance; and (3) condition and abundance.

3.1 METHODS

The effects of flows and benthic macroinvertebrate abundance on trout condition were explored by examining bivariate relationships using Spearman rank correlations and by building nonparametric regression models (random forest models) of trout condition based on all of our independent variables. The effects of flows on the abundance of benthic macroinvertebrates were examined in Magnusson et al. (2008). The years for which data were available for these evaluations are shown in Table 5.

TABLE 5 Years for Which Data Were Available for Statistical Evaluations

	Variables			Statistical Evaluations		
Year	Trout WR, RatioWR, and CPUE in April	Overwinter Flow	Macro-invertebrates in October,[a] January, and April	Flows × Trout Random Forest Model 1	Macro-invertebrates × Trout Random Forest Model 2	Flows × Macro-invertebrates[c]
1990	×[a]	×		×[a]		
1991	×	×		×		
1992	×	×		×		
1993	×	×		×		
1994	×	×		×		
1995	×	×	×	×	×	×
1996	×	×	×	×	×	×
1997	×	×	×	×	×	×
1998	×	×	×	×	×	×
1999	×	×	×	×	×	×
2000	×	×	×	×	×	×
2001	×	×	×	×	×	×
2002	×	×	×	×	×	×
2003		×	×			×
2004		×	×			×
2005		×	×			×
2006	×	×	×	×	×	×
Number of years	13–14	17	12	14	9	12

[a] RatioWR could not be calculated for 1990 because the 1989 fall WR was not determined.

[b] October values from the previous year are used in the analyses because they represent conditions at the beginning of the winter period.

[c] Analyzed in Magnusson et al. (2008).

3.1.1 Spearman Rank Correlations

The statistical significance of relationships between trout condition and flow and macroinvertebrate abundance was examined with nonparametric rank correlation analyses (Spearman rank correlation as measured by the correlation coefficient, r_S ; Table B.1–B.3) using SAS statistical software (Version 9.1; SAS Institute Inc. 2002). As in our previous report, and because of the exploratory nature of these analyses, we did not adjust the critical alpha level (0.05) for statistical significance for the number of analyses. The significance level of each relationship is indicated in each figure (* for $P<0.05$, ** for $P<0.01$, and *** for $P<0.001$). Because of the relatively low number of years available for our analysis (N=9–14 years, Table 5), the relationships that are identified in this report should be regarded only as potentially important relationships that warrant further examination once additional data are collected or experiments are conducted.

3.1.2 Random Forest Models

A random forest regression is a nonparametric multivariate tool, developed by Breiman (2001), that is used for building descriptive and predictive models between a dependent variable and one or more independent variables (Quinn and Keough 2002). It can be used as a nonparametric alternative to multiple and stepwise linear regression analysis because, unlike multiple linear regression techniques, the random forest has no assumption of linearity and no requirement that the number of observations should, by far, outnumber the number of variables in the model. Moreover, the random forest can handle correlated independent variables and missing data (missing data are estimated from a proximity matrix). These characteristics make random forest regressions more suitable than stepwise regressions for our analyses (the data consist of 9–14 observations [years] and 21–25 more or less correlated independent variables).

The random forest is an improvement of Classification and Regression Trees (CART, Breiman et al. 1984) because it takes the average of many regression trees wherein each tree is built on a bootstrapped subsample of the data (about two-thirds of the data) using the best of a randomly selected sample of the independent variables. In each regression tree, the subsampled observations are divided into two groups based on the best independent variable that minimizes the sum of squares in the two groups. This procedure is repeated for each group until a prespecified minimum group size is reached. The remaining one-third of the data that was not used to build the model is used to test the model, providing a measure of the total variance explained (pseudo-R^2, which is not intended to be compared across different data sets; negative values should be interpreted as no variance explained) and a measure of variable importance (%IncMSE = percent increase in mean square error of the full model when the variable is permuted).

We used the random forest approach to build models that describe the relationships among trout condition (WR and ratioWR) and trout abundance (total CPUE), flow variables, and macroinvertebrate variables. We built separate models for WR (measured each spring in 1990–2002 and 2006, N=14) and ratioWR (calculated for 1991–2002 and 2006, N=13) for each size class (small, medium, and large), species (brown trout and rainbow trout), and location (Tailrace and Little Hole). For each group, we built two different models: Model 1 with CPUE + 20 flow variables (1990–2006, N=13–14); and Model 2 with CPUE + 20 flow variables + 4 macroinvertebrate variables (1995–2006, N=9). All small trout were pooled in a combined analysis and, in total, we built 36 random forest models. It should be noted that the statistical property of the pseudo-R^2 value generated for each model prevents comparisons of the total variance explained by the models across trout groups.

The random forest analysis was run using the RandomForest™ package (maintained by Liaw and Wiener 2002) by Leo Breiman and Adele Cutler, Version 4.5-30 (http://cran.r-project.org/web/packages/randomForest/index.html) in the statistical program R, Version 2.9.2 (free download http://www.r-project.org/). Technical details about the random forest regression approach and the R commands used to prepare and run our random forest models are provided in Appendix C. Spearman rank correlation coefficients are reported along with the measure of importance for each variable.

3.2 RESULTS

3.2.1 Trout Condition and Flow

Flow Volume

With the exception of a positive correlation between large Little Hole rainbow trout WR and the number of hours of flows above 3,000 cfs (Figure 7, Table 6), WR was not correlated with any flow volume variable.

RatioWR, on the other hand, was negatively correlated with certain flow volume variables for medium brown trout at both locations (Tailrace: mean_dmax, hgt3000, and hgt4000, and Little Hole: hgt4000; Figure 7, Table 6). RatioWR was not correlated with flow volume for small trout (species and locations combined), large brown trout, or medium to large rainbow trout (Table 6).

Flow Variability

The condition of small trout was not correlated with any measure of flow variability (Table 6). In contrast, WR for medium and large trout was positively correlated with various measures of flow variability at both locations: large Tailrace brown trout WR with within-day and between-day flow variability (mean_dskew, mean_ddelta, and mean_ddelta%, Table 6, Figures 8 and 9); medium to large Tailrace rainbow trout WR with between-day flow variability (mean_ddelta and mean_ddelta%, Table 6, Figure 9); medium to large Little Hole brown trout WR with between-day and within-season flow variability (mean_ddelta, mean_ddelta%, cv_dmean and skew_dmean, Table 6, Figures 9 and 10); and medium to large Little Hole rainbow trout WR with within-season flow variability (skew_dmean, Table 6, Figure 10).

With the exception of a negative correlation between ratioWR of medium Tailrace brown trout and within-season flow variability (range_dmean, Figure 10), ratioWR was not correlated with flow variability (Table 6). Examples of years with high and low within-season skewness are illustrated in Figure 11.

3.2.2 Flow, Macroinvertebrate Abundance, and Trout Condition

Macroinvertebrate abundance in January at Tailrace and in April at Little Hole were negatively correlated with both overwinter flow volume and variability (Magnusson et al. 2008). Relationships between flows and the abundance of specific taxa are presented in Magnusson et al. (2008). January macroinvertebrate abundance was negatively correlated with WR for medium Tailrace rainbow trout and positively correlated with ratioWR for small combined trout, medium Little Hole brown trout, and medium Little Hole rainbow trout

TABLE 6 Statistically Significant Correlations between Trout Condition and Flows for Trout Size Classes at Tailrace and Little Hole

	Trout WR and RatioWR Correlated with Flow Variables								
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace		Brown Trout at Little Hole		Rainbow Trout at Little Hole	
Flow Variable	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
Flow Volume									
mean_dmean	None	None	None	None	None	None	None	None	None
med_dmean	None	None	None	None	None	None	None	None	None
mean_dmin	None	None	None	None	None	None	None	None	None
mean_dmax	None	--ratioWR	None	None	None	None	None	None	None
hgt1000	None	None	None	None	None	None	None	None	None
hgt2000	None	None	None	None	None	None	None	None	None
hgt3000	None	-ratioWR	None	None	None	None	None	None	+WR
hgt4000	None	--ratioWR	None	None	None	-ratioWR	None	None	None
Within-Day Flow Variability									
mean_hdelta	None	None	None	None	None	None	None	None	None
mean_hdelta%	None	None	None	None	None	None	None	None	None
mean_drange	None	None	None	None	None	None	None	None	None
mean_dcv	None	None	None	None	None	None	None	None	None
mean_dskew	None	None	+WR	None	None	None	None	None	None
Between-Day Flow Variability									
mean_ddelta	None	None	+WR	+WR	+WR	++WR	+WR	None	None
mean_ddelta%	None	None	+WR	+WR	+WR	++WR	+WR	None	None
Within-Season Flow Variability									
range_dmean	None	-ratioWR	None	None	None	None	None	None	None
cv_dmean	None	None	None	None	None	+WR	None	None	None
skew_dmean	None	None	None	None	None	+WR	++WR	++WR	++WR
Between-Month Flow Variability									
mean_mdelta	None	None	None	None	None	None	None	None	None
mean_mdelta%	None	None	None	None	None	None	None	None	None

[a] The sign (+ or -) preceding a variable indicates whether the correlation is positive or negative. The significance of the relationship is indicated by the number of signs shown (1: P<0.05, 2: P<0.01, 3: P<0.001). None = neither of the trout variables were significantly correlated; WR = relative weight; ratioWR=ratio of spring to fall WR. All small trout were combined in the first column because of low number of captured fish.

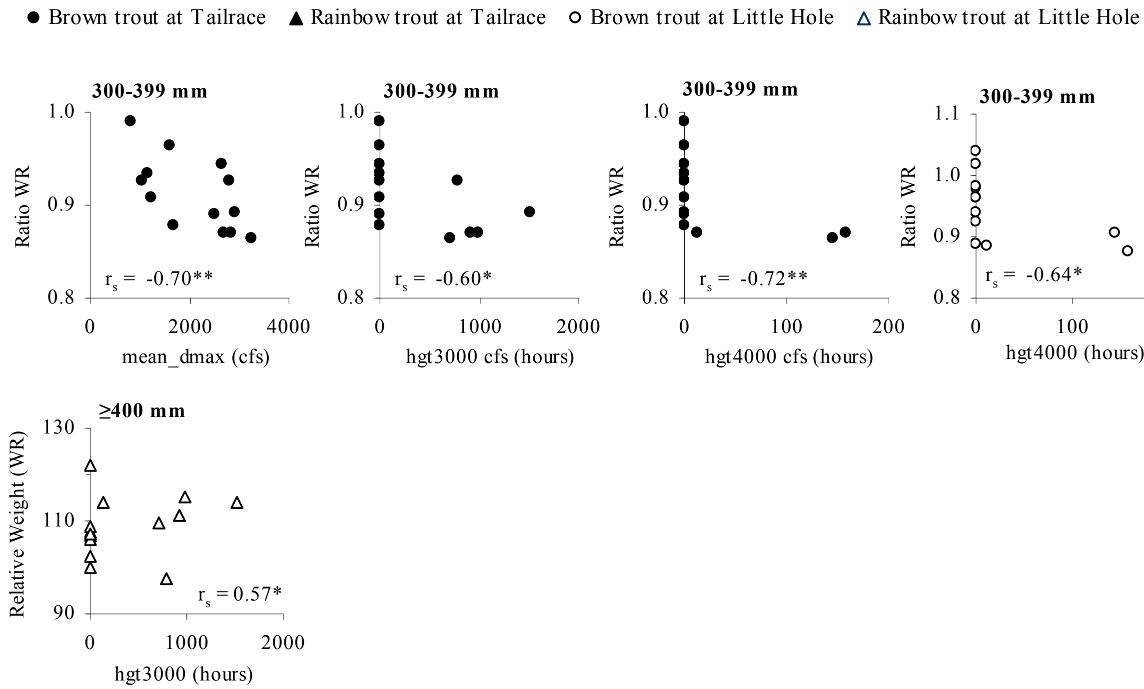


FIGURE 7 Statistically Significant Relationships between Trout Condition and Flow Volume, 1990–2006 (Plots include rank correlation coefficients; asterisks denote level of significance: *=*P*<0.05, **=*P*<0.01, and *=*P*<0.001.)**

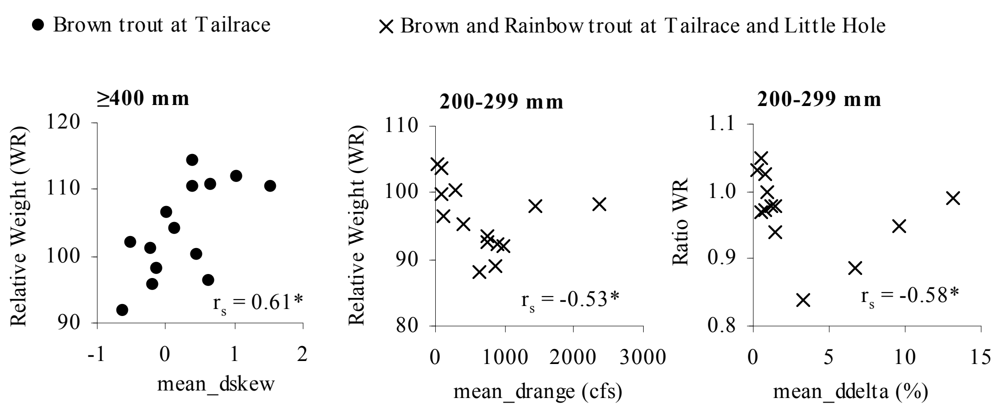


FIGURE 8 Statistically Significant Relationships between Trout Condition and Within-Day Flow Variability, 1990–2006 (Plots include rank correlation coefficients; asterisks denote level of significance: *=*P*<0.05, **=*P*<0.01, and *=*P*<0.001.)**

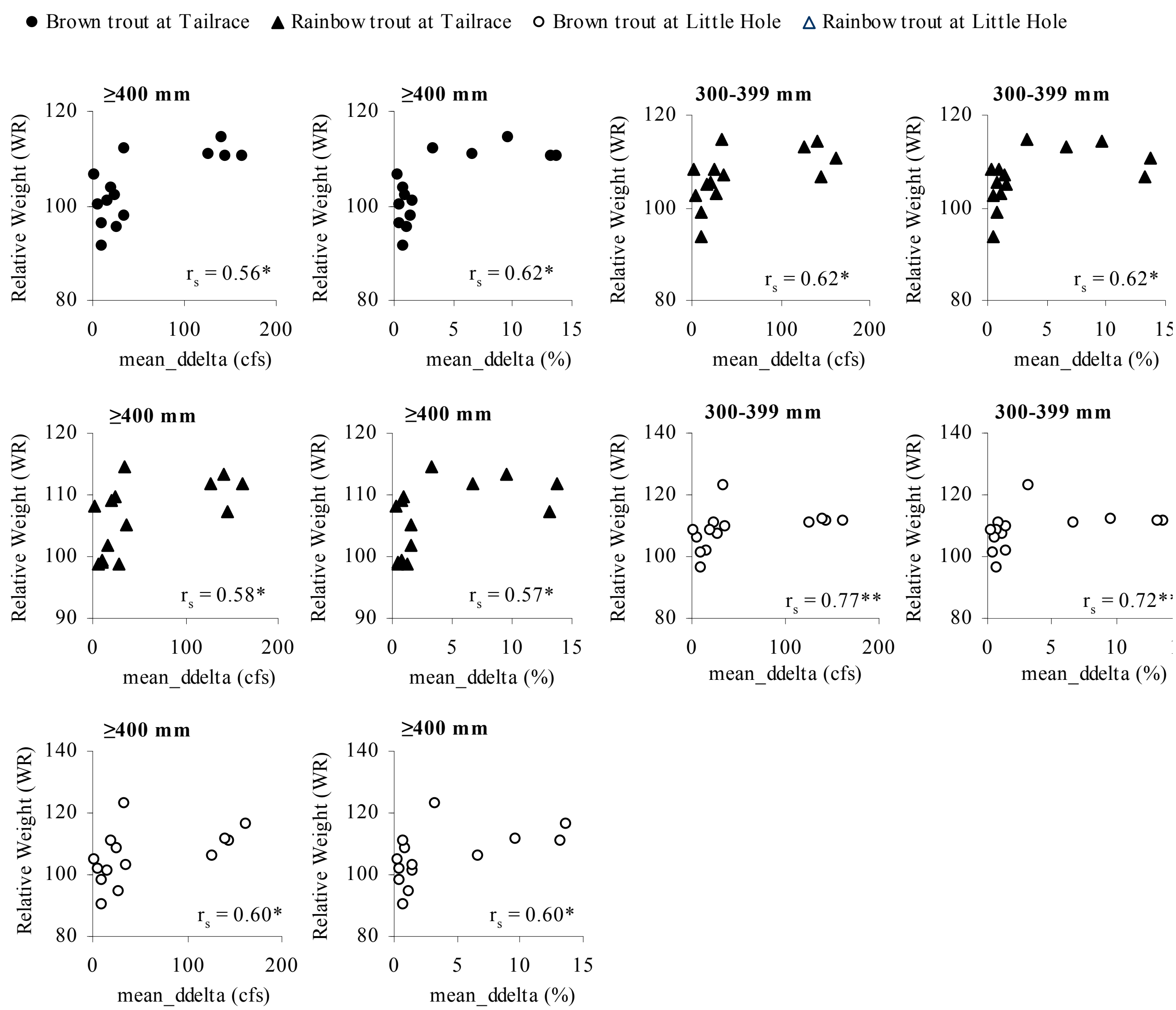


FIGURE 9 Statistically Significant Relationships between Trout Condition and Between-Day Flow Variability, 1990–2006 (Plots include rank correlation coefficients; asterisks denote level of significance: *=*P*<0.05, **=*P*<0.01, and *=*P*<0.001.)**

(Table 7, Figure 12). No correlations were found between October, April, or total macroinvertebrate abundance and WR or ratioWR for any of the trout groups (Table 7).

3.2.3 Random Forest Regression Models

We used random forest regression to build models of WR and ratioWR for each trout group using a total of 21–25 independent variables (Model 1 using measures of CPUE and flow and Model 2 using measures of CPUE, flow, and macroinvertebrate abundance). Only six of the 36 models we built explained any of the variation in the dependent variable, indicating a generally poor ability to account for the variation in condition with the variables we evaluated.

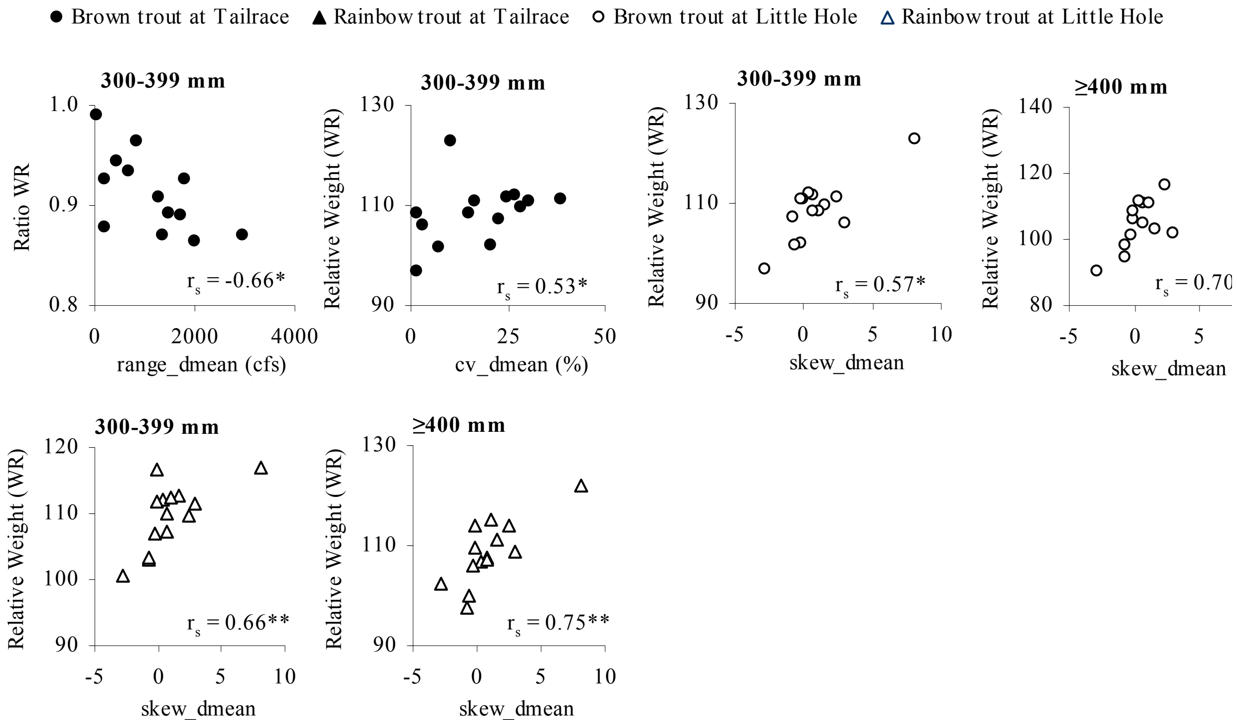


FIGURE 10 Statistically Significant Relationships between Trout Condition and Within-Season Flow Variability, 1990–2006 (Plots include rank correlation coefficients; asterisks denote level of significance: *=*P*<0.05, **=*P*<0.01, and *=*P*<0.001.)**

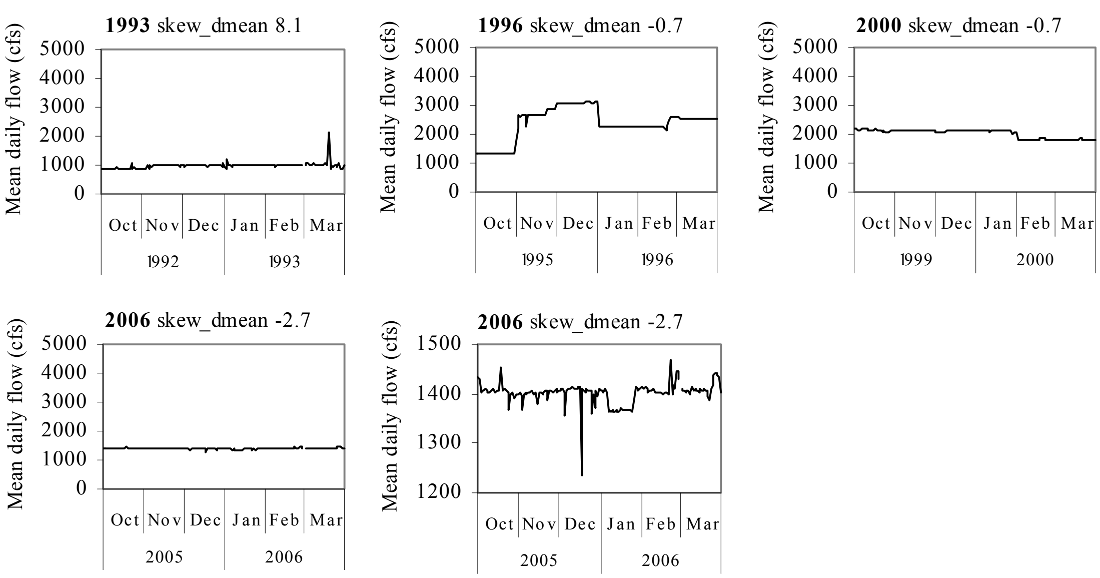


FIGURE 11 Daily Flow Volume in Years with the Highest and Three Lowest Values of Skew_dmean (2006 is shown in two figures with different scales on the *y*-axis.)

TABLE 7 Statistically Significant Correlations between Trout Condition and Macroinvertebrate Variables for Trout Size Classes at Tailrace and Little Hole[a]

	Trout WR and RatioWR Correlated with Macroinvertebrate Variables								
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace		Brown Trout at Little Hole		Rainbow Trout at Little Hole	
Variable	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
October macroinvertebrates	none	none	none	none	none	none	none	none	none
January macroinvertebrates	+ratioWR	none	none	-WR	none	+ratioWR	none	+++ratioWR	none
April macroinvertebrates	none	none	none	none	none	none	none	none	none
Total macroinvertebrates	none	none	none	none	none	none	none	none	none

[a] The sign (+ or -) preceding a variable indicates whether the correlation is positive or negative. The significance of the relationship is indicated by the number of signs shown (1: $P<0.05$, 2: $P<0.01$, 3: $P<0.001$). None = neither WR nor ratioWR were significantly correlated; WR = relative weight; ratioWR = ratio of spring to fall WR. All small trout were combined in the first column because of low number of captured fish. Macroinvertebrate variables were averaged across the two locations in the analyses for combined small trout in the first column.

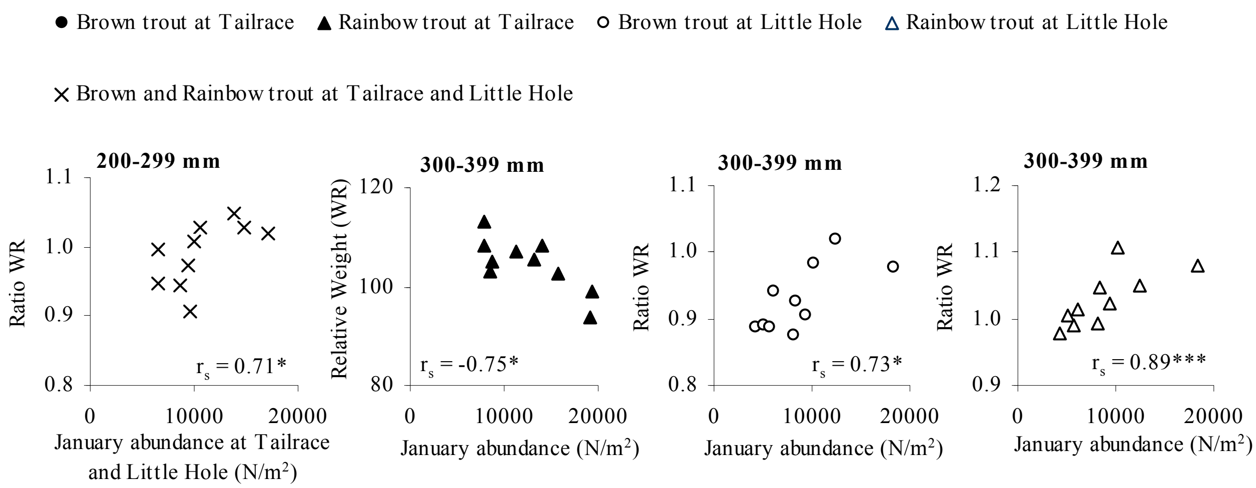


FIGURE 12 Statistically Significant Relationships between Trout Condition and January Macroinvertebrate Abundance, 1994–2006 (Plots include rank correlation coefficients; asterisks denote level of significance: *=*P*<0.05, **=*P*<0.01, and *=*P*<0.001.)**

These six models all described the variation in WR using Model 1 (pseudo-R^2 = 4.6–43.3%). No models described any of the variation in WR when macroinvertebrates were included in the model (Model 2, negative pseudo-R^2 values or nonsignificant %IncMSE values for macroinvertebrate variables, N=397–3,069) despite the fact that benthic invertebrate abundance ranged considerably over the studied years (1994–2002 and 2006, see Magnusson et al. 2008). Negative pseudo-R^2 values imply a very poor model, and a nonsignificant %IncMSE value implies that the variable does not contribute to the model.

All flow variables that were correlated with WR in the Spearman rank correlations were also important in the random forest models (i.e., improved the model with more than 5%). The random forest models also identified additional variables as important for the trout WR (Table 8). Although these variables were not correlated in the Spearman rank correlations (P>0.05), they were useful for splitting the observations in the regression trees.

No models described any of the variation in ratioWR, regardless of species, size group, or location (Model 1 and Model 2, negative pseudo-R^2 values or nonsignificant %IncMSE values for the macroinvertebrate variables, N=397–3,806).

Models of Small Trout

None of the random forest models explained any of the variation in WR or ratioWR for small trout (combined species and locations, N=1,051–1,194).

TABLE 8 Total Explained Variance and Variable Importance in Random Forest Models of Trout Relative Weight[a]

	Variable Importance for WR in Model 1: CPUE + Flows (1990–2002 and 2006)																	
	Brown Trout at Tailrace			Rainbow Trout at Tailrace						Brown Trout at Little Hole						Rainbow Trout at Little Hole		
	≥400 mm			300–399 mm			≥400 mm			300–399 mm			≥400 mm			≥400 mm		
Pseudo-R^2	43.3%			11.1%			4.7%			8.2%			21.7%			4.6%		
Number of fish	1,694			4,090			3,564			1,871			4,091			1,098		
Variable	Rank	%IncMSE	r_s	Rank	%IncMSE	r_s	Rank	%IncMSE	r_s	Rank	%IncMSE	r_s	Rank	%IncMSE	r_s	Rank	%IncMSE	r_s
CPUE total	1	19.7 (-)	-0.65*	4	9.0 (-)	-0.56*	1	10.0 (-)	n.s.	3	7.0 (-)	-0.59*	1	13.7 (-)	-0.55*	-	-1.7	n.s.
mean_dmean	-	0.2	n.s.	-	0.7	n.s.	-	3.0	n.s.	6	3.9 (-)	n.s.	10	2.3 (-)	n.s.	-	-0.9	n.s.
med_dmean	10	3.0 (-)	n.s.	-	1.8	n.s.	-	0.1	n.s.	10	2.4 (-)	n.s.	-	0.9	n.s.	-	-2.5	n.s.
mean_dmin	5	6.6 (-)	n.s.	-	2.3	n.s.	-	4.1	n.s.	11	2.1 (-)	n.s.	6	4.9 (-/+)	n.s.	-	-0.6	n.s.
mean_dmax	-	-0.7	n.s.	-	1.2	n.s.	-	0.1	n.s.	12	1.7 (-/+)	n.s.	-	-0.0	n.s.	-	-0.8	n.s.
hgt1000	11	2.2 (-)	n.s.	8	2.7 (-)	n.s.	-	1.0	n.s.	-	1.3	n.s.	8	2.6 (-)	n.s.	4	5.5 (-)	n.s.
hgt2000	-	0.8	n.s.	-	0.6	n.s.	-	0.8	n.s.	-	-0.3	n.s.	-	0.8	n.s.	-	-0.7	n.s.
hgt3000	-	-0.2	n.s.	-	-0.4	n.s.	-	2.7	n.s.	7	3.9 (+)	n.s.	11	1.6 (+)	n.s.	6	4.2 (+)	0.57*
hgt4000	-	-0.7	n.s.	-	-0.7	n.s.	-	1.1	n.s.	-	-0.4	n.s.	-	-0.7	n.s.	-	-0.0	n.s.
mean_hdelta	6	5.0 (-/+)	n.s.	-	2.4	n.s.	-	-1.4	n.s.	-	1.2	n.s.	7	3.4 (-)	n.s.	5	4.2 (-)	n.s.
mean_hdelta%	9	3.9 (+)	n.s.	7	4.9 (+)	n.s.	-	1.0	n.s.	-	1.1	n.s.	14	1.1 (-/+)	n.s.	-	2.5	n.s.
mean_drange	7	4.9 (-)	n.s.	-	-0.2	n.s.	-	-2.8	n.s.	-	-0.3	n.s.	15	1.0 (-)	n.s.	2	7.6 (-)	n.s.
mean_dcv	-	0.2	n.s.	9	2.7 (-/+)	n.s.	-	-1.8	n.s.	-	0.0	n.s.	-	0.4	n.s.	-	0.7	n.s.
mean_dskew	8	4.2 (+)	0.61*	-	-2.6	n.s.	-	-2.4	n.s.	9	2.9 (+)	n.s.	13	1.2	n.s.	-	-1.0	n.s.
mean_ddelta	3	9.7 (+)	0.56*	3	9.5 (+)	0.62*	4	7.4 (+)	0.58*	1	12.4 (+)	0.77**	3	8.0 (+)	0.60*	7	3.5 (+)	n.s.
mean_ddelta%	2	13.2 (+)	0.62*	1	10.8 (+)	0.62*	2	8.3 (+)	0.57*	4	6.5 (+)	0.72**	4	7.7 (+)	0.60*	-	1.3	n.s.
range_dmean	-	-0.4	n.s.	5	7.1 (+)	n.s.	3	8.2 (+)	n.s.	2	11.6 (+)	n.s.	5	6.2 (+)	n.s.	3	7.3 (+)	n.s.
cv_dmean	-	-2.0	n.s.	6	5.1 (+)	n.s.	-	-0.1	n.s.	8	3.6 (+)	0.53*	9	2.4 (+)	n.s.	-	2.5	n.s.
skew_dmean	4	7.0 (+)	n.s.	2	9.6 (+)	n.s.	-	2.7	n.s.	5	5.3 (+)	0.57*	2	11.0 (+)	0.70**	1	13.4 (+)	0.75**
mean_mdelta	-	-1.0	n.s.	-	-0.8	n.s.	-	-3.1	n.s.	-	-1.4	n.s.	-	-0.9	n.s.	-	-0.7	n.s.
mean_mdelta%	-	-0.8	n.s.	-	-0.8	n.s.	-	-4.6	n.s.	-	1.2	n.s.	12	1.5 (+/-)	n.s.	-	1.8	n.s.
Cut-off value for importance[b]		2.0			2.6			4.6			1.4			0.9			2.5	

TABLE 8 (Cont.)

[a] Total explained variance in WR of each random forest model with positive pseudo-R^2 values (models with negative pseudo-R^2 values are not shown because these explain no variance in WR) followed by variable importance (%IncMSE, with the direction of the relationship following in parentheses [based on partial dependence plots], -/+ denote a positive effect following a negative effect; >5% contributions are denoted in bold); each variable is ranked (Rank: from 1 for highest importance with increasing number for progressively lower importance, “-” denotes no importance when %IncMSE is below the cut-off value for importance); r_s lists the regression coefficients of statistically significant Spearman rank correlations between each variable and WR (asterisk(s) denote significance: *=P<0.05, **=P<0.01, ***=P<0.001, and n.s. denotes P>0.05).

[b] The cut-off value for determining whether a specific variable is influential in a model is determined by the lowest negative value of %IncMSE. The negative values show the range around zero that is caused by randomness, and randomness occurs in both directions. A %IncMSE value above the cut-off value suggests that the variable is important for trout WR.

Models of Medium and Large Trout at Tailrace

The random forest models explained 0–44% of the variation in WR for the medium and large trout at Tailrace (Model 1 for medium brown trout negative pseudo-R^2 value, large brown trout pseudo-R^2 = 43.3%, medium rainbow trout pseudo-R^2 = 11.1%, and large rainbow trout pseudo-R^2 = 4.7%) but no variation in ratioWR (Model 1, negative pseudo-R^2 values, *N*=677–3,806).

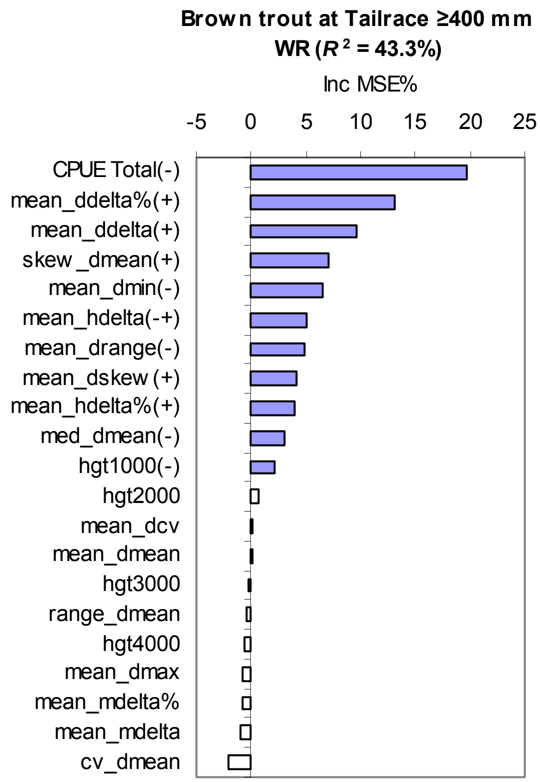


FIGURE 13 Variable Importance in the Random Forest Model of Large Brown Trout WR at Tailrace. Important variables are indicated with filled bars; the direction of the effect for each important variable is indicated in parentheses after the variable name. Plotted values are from Table 8.

Model 1 for large brown trout at Tailrace explained 43.3% of the variation in WR (Table 8, Figure 13) mainly due to the negative effects of trout abundance (CPUE) and flow volume (mean_dmin), the positive effects of between-day flow variability (mean_ddelta% and mean_ddelta), and the positive effects of within-day flow variability (skew_dmean). Model 1 did not explain any of the variation in WR for medium brown trout at Tailrace (negative pseudo-R^2 value, *N*=732).

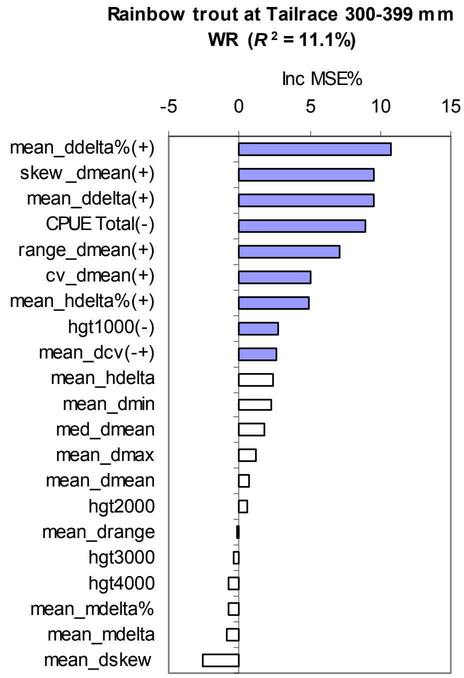
Model 1 for medium rainbow trout at Tailrace explained 11.1% of the variation in WR (Table 8, Figure 14a) mainly due to the positive effects of between-day flow variability (mean_ddelta% and mean_ddelta), the positive effects of within-season flow variability (skew_dmean, range_dmean, and cv_dmean), and the negative effects of CPUE. Model 1 for large rainbow trout at Tailrace explained 4.7% of the variation in WR (Table 8, Figure 14b) mainly due to the positive effects of between-day and within-season flow variability (mean_ddelta%, range_dmean, and mean_ddelta), and the negative effects of CPUE.

Model 2 did not explain any of the variation in WR or ratioWR for any of the medium and large trout at Tailrace (negative pseudo-R^2 values or nonsignificant %IncMSE values for the macroinvertebrate variables, *N*=677–3,069).

Models of Medium and Large Trout at Little Hole

The random forest models explained up to 22% of the variation in WR for the medium and large trout at Little Hole (Model 1 for medium brown trout pseudo-R^2 = 8.2%, large brown trout pseudo-R^2 = 21.7%, medium rainbow trout negative pseudo-R^2 value, and large rainbow

(a)



(b)

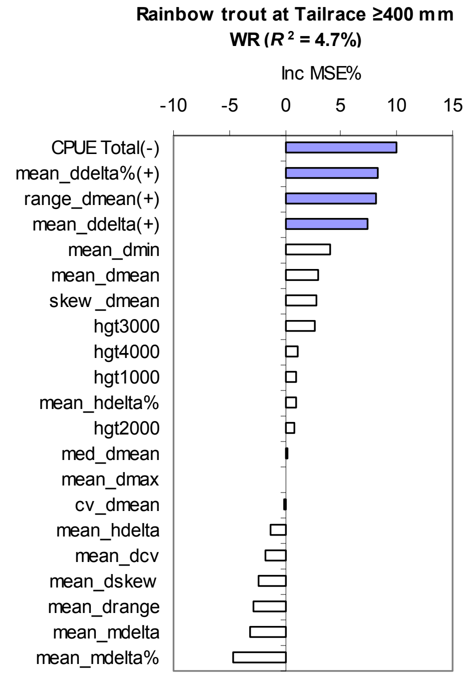


FIGURE 14 Variable Importance in the Random Forest Models of (a) Medium and (b) Large Rainbow Trout WR at Tailrace. Important variables are indicated with filled bars; the direction of the effect for each important variable is indicated in parentheses after the variable name. Plotted values are from Table 8.

trout pseudo-R^2 = 4.6%; *N*=397–2,917) but no variation in ratioWR (Model 1, negative pseudo-R^2 values, *N*=397–3,952).

Model 1 for medium brown trout at Little Hole explained 8.2% of the variation in WR (Table 8, Figure 15a) mainly due to the positive effects of between-day flow variability (mean_ddelta and mean_ddelta%), the positive effects of within-season flow variability (range_dmean and skew_dmean), and the negative effects of CPUE.

Model 1 for large brown trout at Little Hole explained 21.7% of the variation in WR (Table 8, Figure 15b) mainly due to the positive effects of between-day flow variability (mean_ddelta and mean_ddelta%), the positive effects of within-season flow variability (skew_dmean and range_dmean), and the negative effects of CPUE.

Model 1 for large rainbow trout at Little Hole explained 4.6% of the variation in WR (Table 8 and Figure 16) mainly due to the positive effects of within-season flow variability (skew_dmean and range_dmean), the negative effects of within-day flow variability (mean_drange), and the negative effects of flow volume (hgt1000).

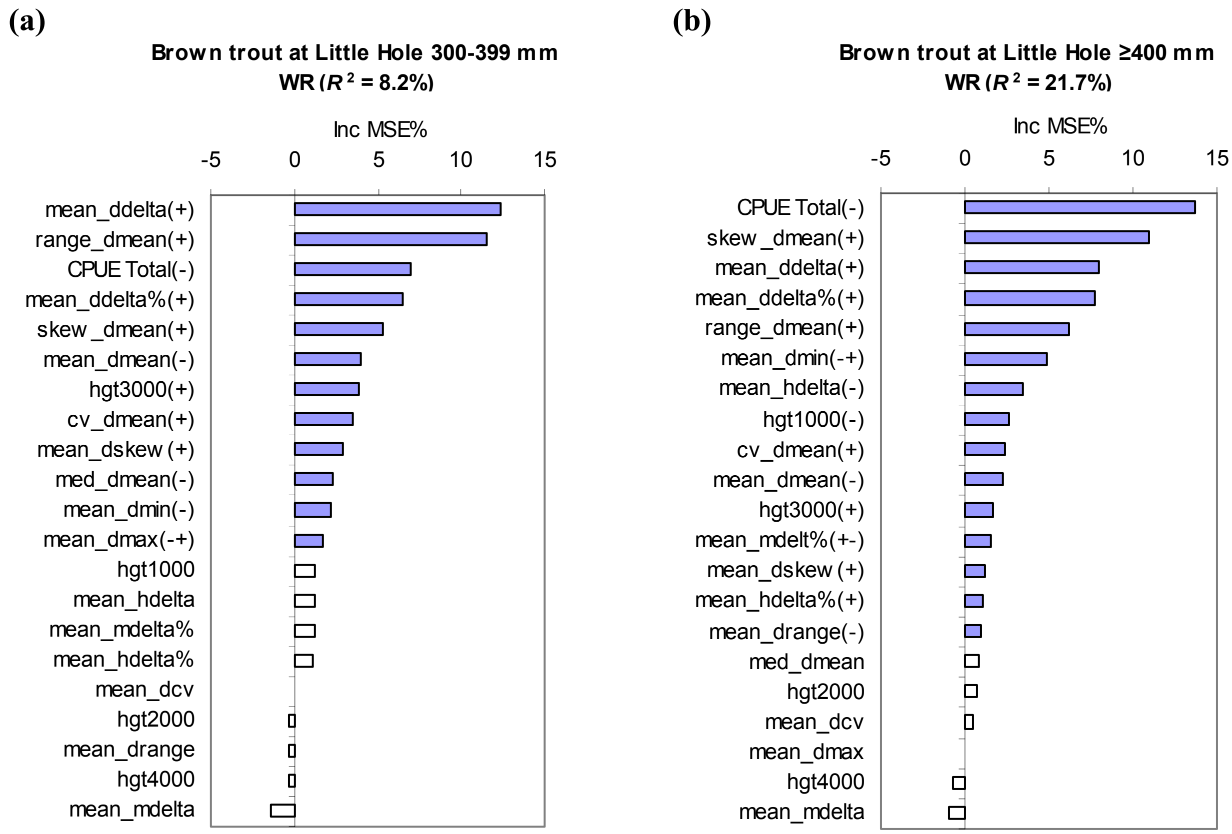


FIGURE 15 Variable Importance in the Random Forest Models of (a) Medium and (b) Large Brown Trout WR at Little Hole. Important variables are indicated with filled bars; the direction of the effect for each important variable is indicated in parentheses after the variable name. Plotted values are from Table 8.

Model 2 did not explain any of the variance in WR or ratioWR values (negative pseudo-R^2 values or nonsignificant %IncMSE values for the macroinvertebrate variables, N=397–2,917).

3.3 DISCUSSION

The results from the analyses of trout size classes (200–299 mm, 300–399 mm, and ≥400 mm) suggest that overwinter flow volume, flow variability, trout abundance, and benthic macroinvertebrate abundance explained relatively little of the variance in spring condition of brown and rainbow trout at Tailrace and Little Hole, particularly when the condition values (WR) had been adjusted to the starting condition values (ratioWR). Less than half of the variation in WR (large brown trout: 43% at Tailrace and 22% at Little Hole) and none of the variation in ratioWR (0% for all trout groups) could be explained by the random forest models. Variance in WR for large Tailrace brown trout was best explained by CPUE (negative effect), flow volume (negative effect), and flow variability (positive effect), whereas variance in WR for large Little Hole brown trout was best explained by CPUE (negative effect) and between-day and within-season flow variability (positive effect). For the remaining medium and large trout, the random forest models explained at most 11% of the variation in WR (medium brown trout: 0% at Tailrace and 8% at Little Hole; medium rainbow trout: 11% at Tailrace and 0% at Little Hole; large rainbow trout: 5% at Tailrace and 5% at Little Hole). Random forest models did not explain any of the variance in WR or ratioWR of small trout.

It is important to note that none of the relationships identified for WR in the Spearman rank correlations and the random forest analyses (discussed below) were found for the ratioWR values, suggesting that the effects on WR could be spurious and related to differences in starting condition of the trout.

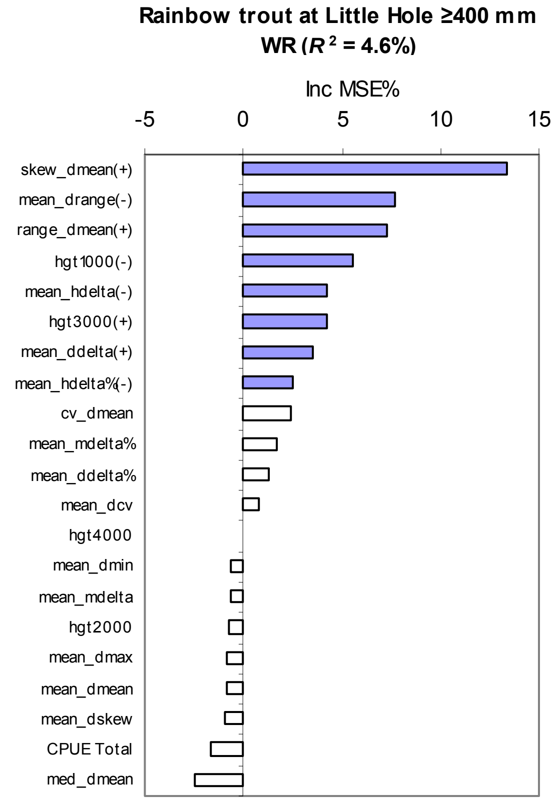


FIGURE 16 Variable Importance in the Random Forest Model of Large Rainbow Trout WR at Little Hole. Important variables are indicated with filled bars; the direction of the effect for each important variable is indicated in parentheses after the variable name. Plotted values are from Table 8.

Flow Volume

The observed inverse relationships between flow volume and trout condition were relatively weak, and none of the flow volume variables explained more than 7% of the variation in WR in the random forest models. Furthermore, flow volume did not appear to be important for ratioWR in any of the random forest models. Habitat simulation studies suggest that increasing flow in winter may limit habitat availability of fingerlings and increase their migration due to a reduction in the availability of low-flow habitats, but increase habitat availability for adults (>250 mm, Johnson et al. 1987). Habitat simulation studies for Flaming Gorge tailwater suggest that habitat availability for rainbow trout may increase at Tailrace but decrease at the Little Hole site with increasing flow (Hann et al. 1991), and individual-based models suggest that trout production at both sites may increase with higher baseflows (Railsback et al. 2006).

The results in Magnusson et al. (2008) suggested that overwinter flow volume may have influenced the size distribution of Little Hole rainbow trout because flow volume was negatively correlated with the length and weight of the combined size classes of Little Hole rainbow trout in spring. This negative correlation appears to be driven by a shift from dominance of large trout early in the study period when flow volume was low toward a dominance of medium-sized fish in the middle of the study period when flow volume was high. This relationship may also have been coupled with trout production because high-flow years may produce more young trout with a larger cohort of 200-299 mm fish appearing the year after.

In the present study, we also found that flow volume tended to have a slight negative effect on WR of large Tailrace brown trout and large Little Hole rainbow trout, and flow volume was negatively correlated with ratioWR of medium-sized brown trout at both locations. Increased flow volume may have caused increased energy expenditure for the fish to maintain their position in the channel, thereby affecting the condition of the trout. In contrast, there was a positive effect on WR from the amount of time flows were above 3,000 cfs for large and combined size classes of Little Hole rainbow trout, possibly due to an increase in habitat availability (present study and Magnusson et al. 2008). Therefore, the observed inverse relationships between condition and flow volume are difficult to explain.

Within-Day Flow Variability

Within-day flow variability appeared to have some effect on WR for large brown trout at Tailrace, medium rainbow trout at Tailrace, medium-to-large brown trout at Little Hole, and large rainbow trout at Little Hole; a majority of these relationships were of limited importance in the random forest models (<10%IncMSE and/or explained variance pseudo-R^2 <5%) and were generally not statistically significant in the Spearman rank correlations. Within-day flow variability was not important for ratioWR for any of the trout groups.

Daily skewness has been reported to be one of the most suitable measures of erratic water releases in different river systems (Olden and Poff 2003), particularly with respect to the effects of daily flow regimes on fish assemblages below hydroelectric dams (Kinsolving and Bain 1993). A day with few high flow values (regardless of magnitude) has a high positive skewness value, whereas a day with predominantly high values (regardless of magnitude) has a negative skewness value. Within-day skewness (mean_dskew) was positively correlated with the length and/or weight of combined size classes of brown and rainbow trout at Little Hole and positively correlated with WR for combined size classes of Tailrace brown trout (Magnusson et al. 2008) and for large Tailrace brown trout (present study), suggesting that the trout benefited from increased within-day flow variability. According to the random forest models, however, mean_dskew described only 4.2% of the variation of WR for large brown trout at Tailrace, and was not an important variable for any of the other trout groups.

Between-Day Flow Variability

Of all flow variables, the two between-day flow variability variables, mean_ddelta and mean_ddelta%, showed the most consistent pattern across species and locations (both trout species at Tailrace and brown trout at Little Hole). Between-day flow variability was positively correlated with the weight of combined size classes of rainbow trout at Tailrace and with the length and weight of combined size classes of rainbow trout at Little Hole (Magnusson et al. 2008). Between-day flow variability was also positively correlated with WR of large Tailrace brown trout, WR of medium and large Tailrace rainbow trout, WR of combined size classes of Tailrace rainbow trout, and WR of medium and large Little Hole brown trout, and was also considered important in the corresponding random forest models. Since correlations were only observed for medium and large trout, it appears that the effects of between-day flow variability on trout condition depended on the size of the trout. None of the observed relationships for WR were, however, significant when the WR values were adjusted to the starting condition of the trout, suggesting that the observed effects for WR were caused by differences in starting condition of the trout.

Within-Season Flow Variability

Within-season flow variability was positively correlated with WR for medium Little Hole brown trout (cv_dmean) but negatively correlated with ratioWR for medium Tailrace brown trout (range_dmean). These two variables were not correlated with WR or ratioWR for any other trout group (present study) or the combined trout (Magnusson et al. 2008) but explained some of the variance in WR in some of the random forest models for medium and large trout (positive effects). No effects of within-season flow variability were observed in small trout WR. It is possible that the observed effects for WR were caused by differences in the starting condition of the trout because no effects were observed for the ratioWR values in the random forest models.

Skewness in daily mean values (skew_dmean) was positively correlated with the length and/or weight of the combined size classes of rainbow trout at both locations and brown trout at Little Hole (Magnusson et al. 2008) and positively correlated with WR for medium, large, and combined size classes of Little Hole brown and rainbow trout. This variable was also one of the more important variables in the random forest models for large Tailrace brown trout (rank 4; %IncMSE=7.0), medium Tailrace rainbow trout (rank 2; %IncMSE=9.6), medium Little Hole brown trout (rank 5; %IncMSE=5.3), large Little Hole brown trout (rank 2; %IncMSE=11.0), and large Little Hole rainbow trout (rank 1; %IncMSE=13.4). The biological relevance of this relationship is, however, not clear. Skew_dmean was used to measure the overall stability in mean daily flow volume over the overwinter period, but closer examination of the flow data show that the year with highest skewness (1993, skew_dmean=8.1, Figure 11) had a single day peak in mid-March (flows increased from 1,050 cfs to 2,150 cfs) but otherwise stable flows. It is unlikely this flow pattern would be more favorable to the trout than in years with similarly stable flows without the short peak (e.g., 2000–2002 and 2006). None of the within-season flow variables were, however, important in the random forest analyses after adjustment of the condition values to the starting condition of the trout (ratioWR).

Between-Month Flow Variability

Neither of the between-month flow variability variables (mean_mdelta and mean_mdelta%) showed any significant correlation with WR or ratioWR of the trout, and these two variables were not important in any of the random forest models.

Food Availability

Filbert and Hawkins (1995) suggested that rainbow trout in the Green River could be food-limited because they found that condition and gut fullness increased with increasing food availability and temperature. The only measure of food availability in our study was the abundance of benthic macroinvertebrates from 1994 to 2006. From these data, we observed that the abundance of benthic macroinvertebrates usually varied five-fold or less during an overwinter period (October, January, and April), and that the total average in the overwinter period varied less than three-fold over the study period (1995–2002 and 2006). It is not known to what extent benthic macroinvertebrate abundance reflected food availability for the trout at Tailrace and Little Hole; trout feed both on drift and benthos (Angradi and Griffith 1990), depending on which is most available (Filbert 1991). Our analyses show that WR of medium Tailrace rainbow trout was negatively correlated with January macroinvertebrate abundance, whereas ratioWR of medium Little Hole trout was positively correlated with January macroinvertebrate abundance. Several taxa were also positively correlated with the length and weight of combined size classes of trout (Magnusson et al. 2008).

However, benthic macroinvertebrate abundance did not explain any variation in WR or ratioWR in the random forest models for any of the trout groups. Instead, the explanatory power of the random forest models generally decreased substantially when invertebrate variables were included in the models, possibly due to the exclusion of observations in 1990–2004 when no macroinvertebrate data were available and because the link between the total abundance of benthic macroinvertebrates and trout condition was weak. It is possible that body condition of trout may be more related to the biomass of specific taxa rather than to total abundance of drifting macroinvertebrates (Annear et al. 2002).

Trout Abundance

According to the random forest models, CPUE was often more important than flows and benthic macroinvertebrate abundance for describing WR. CPUE was particularly important for large Tailrace brown trout (%IncMSE=19.7), large Tailrace rainbow trout (%IncMSE=10.0), and large Little Hole brown trout (%IncMSE=13.7). CPUE also contributed to the variation in WR for medium Tailrace rainbow trout (%IncMSE=9.0) and medium Little Hole brown trout (%IncMSE=7.0). CPUE did not contribute to any of the variation in WR for small combined trout and large Little Hole rainbow trout. The negative effect of CPUE on trout condition could be due to increasing competition for food and habitat. CPUE did not, however, have any effect on ratioWR for any of the trout groups, suggesting that the observed effects on WR could have been an artifact due to differences in starting condition of the trout.

Relative Importance

None of the tested fish abundance, flow, and invertebrate variables appeared to influence ratioWR for any of the trout groups, according to the random forest analyses. However, these three groups of variables were important in explaining variation in WR, with total fish abundance being most important, flows being next in importance, and benthic macroinvertebrate abundance being least important. As stated previously, however, the biological meaning of the WR values is not clear because the observed relationships could have been influenced by differences in starting condition of the trout. Flow variability and the size and condition of trout decreased over the study period, whereas the abundance of trout increased over the same period (Magnusson et al. 2008). It is possible that the trend toward smaller fish over the years could be attributed to increasing mortality rates in rainbow trout and brown trout (including those caused by increased fishing pressure) of the larger fish or increased production of brown trout. There is, however, no quantitative data on mortality rates available for the study period.

Condition versus Condition Ratio

As mentioned throughout the discussion, ratioWR did not appear to have the same relationships to flow and macroinvertebrate variables as the WR values. The random forest models explained none to 43% of the variation in WR compared with none of the variation in ratioWR. In addition, the WR values were either not correlated or positively correlated with flow, whereas the ratioWR were either not correlated or negatively correlated with flow. Also, the WR values were either not correlated or negatively correlated with macroinvertebrate abundance, whereas ratioWR values were either not correlated or positively correlated with macroinvertebrate abundance.

Mean trout WR values ranged from 89 to 123 during the study period, and the spring to fall ratioWR values ranged from 0.85 to 1.11. These values are comparable to condition values reported for subadult rainbow trout (150–300 mm body length) during the winter months in two Wyoming tailwaters (WR=92–110; ratioWR=0.86–1.05, our calculation from October and February data in Hebdon and Hubert 2001). Hebdon and Hubert (2001) also report that small rainbow trout with a WR value above 75 do not die from starvation and that the body condition of trout may remain high over the winter when food availability is sufficiently high. Thus, the average trout in our study was in relatively good condition and medium and large rainbow trout increased in condition over the winter (mean ratioWR=1.03). Brown trout decreased slightly in condition over winter (mean ratioWR=0.93), presumably because they spawn in the fall.

We believe that spring-to-fall ratios more accurately measure the effect of overwinter variables than the WR values and that the relationships found for WR could have been due to initial differences in the starting condition of the trout before the onset of winter. We found no reason to suspect that the ratioWR values should be less reliable than the WR values.

It should be recognized that the analyses were derived from relatively few data points (*N*=9 to 14 years), and the data include a gap between 2003 and 2005 when no trout were collected in the spring. The small sample size limits confidence in our results. Because of the exploratory nature of the analyses in this study, the critical alpha value for rejecting hypotheses was not adjusted to the number of tests made (e.g., no Bonferroni corrections). Had Bonferroni corrections been applied for each trout group (i.e., 0.05 divided by 20 tested flow variables = critical alpha value of 0.0025), only two correlations would be statistically significant in the Spearman rank correlations (medium-sized Little Hole brown trout WR versus mean_ddelta, and large Little Hole rainbow trout WR versus skew_dmean). Therefore, the relationships that are identified from the correlation analyses in this report should be regarded only as potentially important relationships that warrant further examination once additional data have been collected. The significant relationships detected suggest future experimentation and hypothesis testing should focus on the effects of flow variability on macroinvertebrate standing crops, drift rates, and the feeding behavior of trout. Moreover, this analysis only examined overwinter effects. Because the survival and condition of trout in the spring also depend on their condition at the beginning of winter, we suggest that the effects of flows on macroinvertebrate abundance and trout condition for other seasons of the year be examined as well.

4 CONCLUSIONS AND RECOMMENDATIONS

Our analyses suggest that overwinter flows, trout abundance (measured as CPUE), and total abundance of benthic macroinvertebrates had limited effects on trout condition in spring during the study period after adjustment of the trout condition values to the starting condition values (ratioWR). Although ratioWR of medium brown trout was negatively correlated with flow volume at both locations and with within-season flow variability at Tailrace, and ratioWR of medium brown and rainbow trout at Little Hole were positively correlated with benthic macroinvertebrate abundance, none of the flow or macroinvertebrate variables were important in the random forest models of ratioWR for any of the trout groups.

Without adjustment of the condition values, the random forest models described up to 43% (but usually considerably less) of the variation in WR. Trout abundance, measured here as CPUE, was the most important descriptor of WR for large brown trout at both locations and for large rainbow trout at Tailrace. CPUE was, in most cases, a more important variable than any of the flow variables. However, these relationships appeared to be influenced by the starting conditions of the trout, since none of these effects were identified for the ratio values. The effects of flow in other time periods on the trout condition in fall and spring warrant further examination.

Without adjustment of the condition values, the random forest models suggested that flow volume had a negative effect on WR for large Tailrace brown trout and that flow variability had a positive effect on WR for large Tailrace brown trout, medium to large Tailrace rainbow trout, medium to large Little Hole brown trout, and large Little Hole rainbow trout. Since these effects were not observed after adjustment of the condition values to the starting values, these results may have been an artifact due to differences in condition before the onset of winter.

January macroinvertebrate abundance was positively correlated with ratioWR for small combined trout and for medium Little Hole brown and rainbow trout but negatively correlated with WR for medium Tailrace rainbow trout. None of the macroinvertebrate variables were, however, important in the random forest models, suggesting that the relationships between benthos and trout condition were weak or that the sample size was too small to build robust models of the data (N=9–10 years).

To further explore the observed correlations among spring trout condition, flows, and macroinvertebrates, and to understand the mechanisms behind their relationships, we recommend gathering the following information:

- Once additional data have been collected, reanalyze the new data to verify that the lack of significant effects was not caused by the low sample size.

- Collect gut contents from different size classes during the winter period at Tailrace and Little Hole in combination with sampling of drift and benthic macroinvertebrates to determine whether benthic abundance is a good estimate of food availability for the trout.

- Determine and analyze individual ratioWR values from passive integrated transponder (PIT)-tagged trout and condition data collected since 2007. This information may help to determine whether the ratioWR values used in this report could be biased.

- Run the recently updated individual-based model developed by Railsback et al. (2005, 2006) to test how flow volume and variability may affect the availability of different habitat types. This information could help explain some of the relationships between flows and trout condition.

- Analyze the effects of flows (including peak flow duration and magnitude) during other periods of the year on the condition of trout in fall and spring. Such analyses would help to determine whether flows during other seasons are more or less important for the trout than flows during the overwinter periods.

5 LITERATURE CITED

Angradi, T.R., and J.S. Griffith, 1990, "Diel Feeding Chronology and Diet Selection of Rainbow Trout in the Henry's Fork of the Snake River, Idaho," *Canadian Journal of Fisheries and Aquatic Sciences* 47:199–209.

Annear, T.C., W. Hubert, D. Simpkins, and L. Hebdon, 2002, "Behavioural and Physiological Response of Trout to Winter Habitat in Tailwaters in Wyoming, USA," *Hydrological Processes* 16: 915–925.

Breiman, L., 2001, "Random Forests," *Machine Learning* 45:5–32.

Breiman, L, J.H. Friedman, R.A. Olshen, and C.J. Stone, 1984, *Classification and Regression Trees*, Wadsworth, Belmont, Calif.

Filbert, R., 1991, "Food Habits and Condition of Rainbow Trout," in: T. Modde et al., *Flaming Gorge Tailwater Fisheries Investigations—Factors Influencing Trout Populations and Habitat Utilization 1987–1989*, Publication No. 91–10, Utah Division of Wildlife Resources, Salt Lake City, Utah.

Filbert, R., and C.P. Hawkins, 1995, "Variation in Condition of Rainbow Trout in Relation to Food, Temperature, and Individual Length in the Green River, Utah," *Transaction of the American Fisheries Society* 124:824–835.

Gore, J.A., 1996, "Responses of Aquatic Biota to Hydrological Change," in: G. Petts and P. Calow (eds.), *River Biota—Diversity and Dynamics*, Blackwell Science Ltd., Oxford, England.

Hann, D., K. Rose, and M. Holden, 1991, "Sub-project 5: Simulated Habitat Availability," in: T. Modde, D.A. Young, and D.L. Archer (eds.), *Evaluation of Factors Influencing Population Characteristics and Habitat Utilization of Trout in the Flaming Gorge Tailwater 1987-1989*, Publication 91-10, Utah Division of Wildlife Resources, Salt Lake City, Utah.

Hayse, J.W., K.E. LaGory, and W.S. Vinikour, 2009, *Research to Evaluate the Effects on Trout of Overwinter Double-Peaking Operations at Flaming Gorge Dam*, Environmental Science Division, Argonne National Laboratory, Argonne, Ill. October.

Hebdon, J.L., and W.A. Hubert, 2001, "Body Condition and Stomach Contents of Subadult Trout during Fall and Winter in Three Wyoming Tailwaters," *North American Journal of Fisheries Management* 21:897–903.

Johnson, J.E., R.P. Kramer, E. Larson, and B.L. Bonebrake, 1987, *Trout Growth, Harvest, Survival, and Microhabitat Selection in the Green River, Utah, 1978–82*, Publication 87–13, Utah Division of Wildlife Resources, Salt Lake City, Utah.

Kinsolving, A.D., and M.B. Bain, 1993, "Fish Assemblage Recovery along a Riverine Disturbance Gradient," *Ecological Applications* 3:531–544.

Liaw, A., and M. Wiener, 2002, "Classification and Regression by RandomForest," *R News* 2(3):18–22.

Magnusson, A.K., K.E. LaGory, and J.W. Hayse, 2008, *Statistical Evaluation of the Effects of Fall and Winter Flows on the Spring Condition of Rainbow and Brown Trout in the Green River Downstream of Flaming Gorge Dam*, Environmental Science Division, Argonne National Laboratory, final report to Western Area Power Administration, Colorado River Storage Project Management Center.

Marchetti, M.P., and P.B. Moyle, 2001, "Effects of Flow Regime on Fish Assemblages in a Regulated California Stream," *Ecological Applications* 11:530–539.

McKinney, T., and D.W. Speas, 2001, "Observations of Size-Related Asymmetries in Diet and Energy Intake of Rainbow Trout in a Regulated River," *Environmental Biology of Fishes* 61:435–444.

Murphy, B.R., D.W. Willis, and T.A. Springer, 1991, "The Relative Weight Index in Fisheries Management: Status and Needs," *Fisheries* 16:30–38.

Muth, R.T., L.W. Crist, K.E. LaGory, J.W. Hayse, K.R. Bestgen, J.K. Lyons, T.P. Ryan, and R.A. Valdez, 2000, *Flow Recommendations for Endangered Fishes in the Green River Downstream of Flaming Gorge Dam*, final report to Upper Colorado River Endangered Fish Recovery Program, U.S. Fish and Wildlife Service, Denver, Colo.

Olden, J.D., and N.L. Poff, 2003, "Redundancy and the Choice of Hydrologic Indices for Characterizing Streamflow Regimes," *River Research and Applications* 19:101–121.

Osmundson, D.T., R.J. Ryel, V.L Lamarra, and J. Pitlick, 2002, "Flow-Sediment-Biota Relations: Implications for River Regulation Effects on Native Fish Abundance," *Ecological Applications* 12:1719–1729.

Quinn, G.P., and M.J. Keough, 2002, *Experimental Design and Data Analysis for Biologists*, Cambridge, UK.

Railsback, S.F., B.C. Harvey, J.W. Hayse, and K.E. LaGory, 2005, "Tests of Theory for Diel Variation in Salmonid Feeding Activity and Habitat Use," *Ecology* 86:947–959.

Railsback, S.F., J.W. Hayse, and K.E LaGory, 2006, *Simulation Analysis of Within-Day Flow Fluctuation Effects on Trout below Flaming Gorge Dam*, Environmental Science Division, Argonne National Laboratory, ANL/EVS/TM-01 EPRI 1012855, final report to Western Area Power Administration.

Reclamation (Bureau of Reclamation), 2005, *Operation of Flaming Gorge Dam Environmental Impact Statement,* U.S. Department of Interior, Bureau of Reclamation, Upper Colorado Region, Salt Lake City, Utah.

Reclamation, 2006, *Record of Decision, Operation of Flaming Gorge Dam Environmental Impact Statement*, U.S. Department of Interior, Bureau of Reclamation, Upper Colorado Region, Salt Lake City, Utah.

SAS Institute Inc., 2002, *SAS Statistical Software, Version 9.1, Users Guide*, Cary, N.C.

USFWS (U.S. Fish and Wildlife Service), 1992, *Final Biological Opinion on the Operation of Flaming Gorge Dam*, Mountain-Prairie Region, Denver, Colo., Nov.

Vinson, M.R., E.C. Dinger, and M. Baker, 2006, *Flaming Gorge Tailwater Aquatic Biota Monitoring Program, 1994–2005*, National Aquatic Monitoring Center, Department of Watershed Sciences, Utah State University, Logan, Utah.

APPENDIX A

SUMMARY STATISTICS OF RELATIVE WEIGHT VALUES CALCULATED FROM INDIVIDUAL TROUT IN THE GREEN RIVER DOWNSTREAM OF FLAMING GORGE DAM, 1990–2006

TABLE A.1 Summary Statistics of Relative Weight Values for Individual Trout for Each Size Class at Tailrace and Little Hole[a]

Location	Species	Size (mm)	*N*	Min	Max	Means	SD	SE	Q25	Median	Q75
Tailrace and Little Hole	Brown and Rainbow trout	200–299	1194	51.0	196.0	96.1	13.5	0.4	88.0	96.0	103.0
Tailrace	Brown trout	200–299	77	71.0	165.0	93.8	12.5	1.4	87.0	94.0	99.0
Tailrace	Brown trout	300–399	732	63.0	186.6	100.5	11.6	0.4	94.0	100.0	106.0
Tailrace	Brown trout	≥400	1694	58.0	172.0	101.9	12.9	0.3	93.6	102.0	110.6
Tailrace	Rainbow trout	200–299	765	64.0	196.0	97.3	13.6	0.5	90.0	97.0	103.0
Tailrace	Rainbow trout	300–399	4090	53.0	191.0	105.8	12.8	0.2	98.0	106.0	113.4
Tailrace	Rainbow trout	≥400	3564	55.0	217.4	108.3	14.7	0.2	99.0	108.0	117.0
Little Hole	Brown trout	200–299	260	51.0	151.0	93.7	12.2	0.8	86.0	94.0	101.5
Little Hole	Brown trout	300–399	1871	55.0	200.0	105.7	12.2	0.3	98.6	105.0	112.0
Little Hole	Brown trout	≥400	4091	51.0	182.7	105.7	14.5	0.2	97.0	106.0	115.0
Little Hole	Rainbow trout	200–299	92	63.0	153.7	94.6	15.2	1.6	83.8	94.7	103.0
Little Hole	Rainbow trout	300–399	1312	61.0	199.0	109.3	13.6	0.4	101.8	109.2	116.7
Little Hole	Rainbow trout	≥400	1098	54.9	159.0	110.8	15.7	0.5	101.0	111.0	121.0

[a] *N* = Number of trout, Min = Minimum, Max = Maximum, SD = standard deviation, SE = standard error, Q25 = 25th percentile, Q75 = 75th percentile.

APPENDIX B

SPEARMAN RANK CORRELATIONS AMONG SPRING TROUT CONDITION AND OCTOBER THROUGH MARCH FLOW AND MACROINVERTEBRATE VARIABLES IN THE GREEN RIVER DOWNSTREAM OF FLAMING GORGE DAM, 1990–2006

TABLE B.1 Spearman Rank Correlations (r_s) among Trout WR and CPUE and Flow Variables, 1990–2006[a]

	Trout WR Correlated with CPUE and Flow Variables								
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace		Brown Trout at Little Hole		Rainbow Trout at Little Hole	
Variable	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
$CPUE_{Total}$	-0.11	-0.03	-0.65	-0.56	-0.53	-0.58	-0.59	0.07	-0.11
mean_dmean	-0.20	0.05	-0.35	-0.19	-0.10	-0.13	-0.21	0.02	-0.02
med_dmean	-0.08	0.07	-0.41	-0.24	-0.17	-0.20	-0.30	0.03	-0.08
mean_dmin	-0.17	0.12	-0.44	-0.27	-0.23	-0.22	-0.24	0.03	-0.02
mean_dmax	-0.09	0.09	-0.09	0.07	0.16	0.10	-0.02	0.22	0.12
hgt1000	-0.29	-0.05	-0.45	-0.35	-0.24	-0.25	-0.35	-0.14	-0.26
hgt2000	-0.18	0.28	-0.06	0.11	0.16	0.22	0.15	0.25	0.24
hgt3000	-0.04	0.35	-0.03	0.29	0.27	0.28	0.25	0.45	0.57
hgt4000	-0.30	0.08	0.12	0.28	0.28	0.18	0.25	0.50	0.48
mean_hdelta	-0.32	-0.17	0.20	0.16	0.21	0.29	0.06	0.07	-0.22
mean_hdelta%	-0.27	-0.25	0.40	0.30	0.37	0.35	0.21	0.09	-0.13
mean_drange	-0.35	-0.21	0.04	0.07	0.11	0.10	-0.10	-0.01	-0.35
mean_dcv	-0.39	-0.33	0.25	0.17	0.27	0.23	0.10	0.01	-0.18
mean_dskew	0.12	0.07	0.61	0.31	0.32	0.49	0.50	0.39	0.24
mean_ddelta	-0.28	0.05	0.56	0.62	0.58	0.77	0.60	0.31	0.32
mean_ddelta%	-0.16	-0.04	0.62	0.62	0.57	0.72	0.60	0.27	0.27
range_dmean	-0.25	0.18	0.24	0.50	0.43	0.50	0.34	0.43	0.32
cv_dmean	-0.22	-0.02	0.38	0.48	0.40	0.53	0.38	0.28	0.20
skew_dmean	-0.04	0.38	0.52	0.48	0.38	0.57	0.70	0.66	0.75
mean_mdelta	0.09	0.15	0.05	0.24	0.15	0.25	0.02	0.21	0.02
mean_mdelta%	0.15	0.06	0.19	0.28	0.18	0.26	0.04	0.16	-0.09

[a] Statistically significant correlations ($P \leq 0.05$) are indicated in bold text. CPUE was averaged across the two locations in the analyses for combined small trout.

TABLE B.2 Spearman Rank Correlations (r_s) among Trout RatioWR and CPUE and Flow Variables, 1990–2006[a]

	Trout RatioWR Correlated with CPUE and Flow Variables								
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace		Brown Trout at Little Hole		Rainbow Trout at Little Hole	
Variable	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
$CPUE_{Total}$	0.02	-0.31	0.01	-0.37	-0.29	-0.33	-0.02	-0.05	0.10
mean_dmean	0.01	-0.54	-0.18	-0.45	0.01	-0.45	-0.25	-0.09	0.04
med_dmean	0.02	-0.51	-0.22	-0.47	-0.04	-0.40	-0.26	-0.09	-0.06
mean_dmin	-0.04	-0.40	-0.10	-0.54	-0.12	-0.47	-0.13	-0.19	0.05
mean_dmax	-0.12	-0.70	-0.26	-0.32	-0.07	-0.36	-0.34	-0.07	-0.03
hgt1000	0.18	-0.35	-0.17	-0.42	0.22	-0.28	-0.25	0.03	-0.03
hgt2000	-0.08	-0.43	-0.09	-0.37	-0.05	-0.38	-0.24	-0.20	0.01
hgt3000	-0.38	-0.60	-0.29	-0.53	-0.34	-0.48	-0.22	-0.46	0.00
hgt4000	-0.46	-0.72	-0.06	-0.26	-0.28	-0.64	-0.50	-0.33	-0.08
mean_hdelta	-0.04	-0.36	0.00	0.23	0.36	0.14	-0.14	0.30	0.02
mean_hdelta%	-0.12	-0.30	-0.07	0.26	0.27	0.31	-0.02	0.29	0.06
mean_drange	-0.07	-0.43	-0.08	0.09	0.32	0.07	-0.24	0.25	-0.14
mean_dcv	-0.05	-0.43	-0.05	0.21	0.30	0.21	-0.05	0.36	0.13
mean_dskew	-0.05	0.16	0.24	0.36	-0.16	0.25	0.01	0.04	-0.04
mean_ddelta	-0.40	-0.38	-0.07	0.09	0.13	0.08	0.06	-0.05	0.14
mean_ddelta%	-0.46	-0.25	-0.18	0.06	0.03	0.10	0.06	-0.13	0.02
range_dmean	-0.53	-0.66	-0.20	-0.24	-0.04	-0.33	-0.26	-0.27	-0.13
cv_dmean	-0.43	-0.44	-0.11	0.01	0.09	-0.23	-0.24	-0.14	-0.17
skew_dmean	-0.51	-0.05	0.31	0.05	-0.54	-0.14	0.15	-0.39	0.21
mean_mdelta	-0.30	-0.46	-0.25	-0.20	0.03	-0.30	-0.28	-0.18	-0.29
mean_mdelt%	-0.35	-0.31	-0.24	-0.12	0.04	-0.27	-0.37	-0.21	-0.46

[a] Statistically significant correlations ($P \leq 0.05$) are indicated in bold text. CPUE was averaged across the two locations in the analyses for combined small trout.

TABLE B.3 Spearman Rank Correlations (r_s) between Trout Variables and Macroinvertebrate Variables, 1994–2006[a]

Trout and Macroinvertebrate Variables	Trout Variables Correlated with Macroinvertebrate Variables								
	Brown and Rainbow Trout at Tailrace and Little Hole	Brown Trout at Tailrace		Rainbow Trout at Tailrace		Brown Trout at Little Hole		Rainbow Trout at Little Hole	
	200–299 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm	300–399 mm	≥400 mm
WR									
October macroinvertebrates	0.10	0.32	0.05	0.40	-0.10	0.32	-0.03	0.20	0.05
January macroinvertebrates	-0.37	-0.43	-0.45	-0.75	-0.55	-0.32	-0.52	-0.26	-0.44
April macroinvertebrates	0.02	-0.36	0.13	-0.19	0.16	-0.45	-0.26	-0.41	-0.41
Total macroinvertebrates	0.02	-0.43	-0.10	-0.47	-0.27	0.10	-0.15	0.02	-0.20
RatioWR									
October macroinvertebrates	-0.40	0.03	0.13	0.15	0.17	0.22	0.50	0.17	0.05
January macroinvertebrates	0.71	0.35	0.44	0.24	0.22	0.73	0.43	0.89	0.27
April macroinvertebrates	0.41	-0.19	-0.08	0.09	-0.09	0.22	0.13	-0.09	-0.26
Total macroinvertebrates	0.38	0.47	-0.08	0.30	0.22	0.42	0.38	0.30	-0.18

[a] Statistically significant correlations ($P \leq 0.05$) are indicated in bold text. WR = relative weight, ratioWR = change in WR from fall to spring. Macroinvertebrate variables were averaged across the two locations (Tailrace and Little Hole) in the analyses for combined small trout in the first column.

APPENDIX C

TECHNICAL DESCRIPTION OF A RANDOM FOREST MODEL AND AN EXAMPLE RUN USING R

C.1 TECHNICAL DESCRIPTION OF THE RANDOM FOREST MODEL

Variables

A random forest model describes a response variable (dependent variable) using a set of descriptor variables (independent variables). For the trout data, the random forest model was used to describe the relationships among a set of independent variables (CPUE, flows, and macroinvertebrate abundance) and a dependent variable WR/ratioWR.

Regression Tree

A random forest model is an extension of a regression trees. A regression tree is a descriptive and predictive model between a dependent variable and several independent variables. It is an upside down tree with the "root" at the top. The root contains all observations (about two-thirds of the observations in the random forest). These observations are then divided into two groups or branches using the independent variable that yields the two smallest within-group (residual) sums of squares within the dependent variable. Each new group is divided further until a predefined minimum node size is reached.

Random Forest Regression Model

A random forest is a collection of regression trees in which each tree is built on a bootstrapped subsample of the data (about two-thirds of the data). The remaining samples are called out-of-bag (OOB) data and are used to calculate the predictive power of each variable in the model and also to obtain a measure of the total variance explained. The observations are split into two groups using the best of a randomly chosen subset of descriptor variables (about one-third of the total number of predictors). Like the regression tree, each new group is further split into two new subgroups until the observations cannot be divided further.

A Tree in the Forest—an Example

A random forest is an average of many trees (in our case, 2,000 trees), and the forest cannot be illustrated as a summary tree. In Figure C.1 we show one of the 2,000 trees that together build the forest. In this particular tree, 16 observations of WR were divided into two groups based on the value of CPUE in the same year. The new left group consists of six observations of WR for which the CPUE is lower than 617.8, whereas the right group consists of ten observations of WR for which the CPUE is higher than 617.8. The group on the right hand side with ten observations was further divided into two groups based on values of skew_dmean in the same years. The ten observations were divided into one group of four WR values from years when skew_dmean was lower than -0.46 and one group of six values of WR

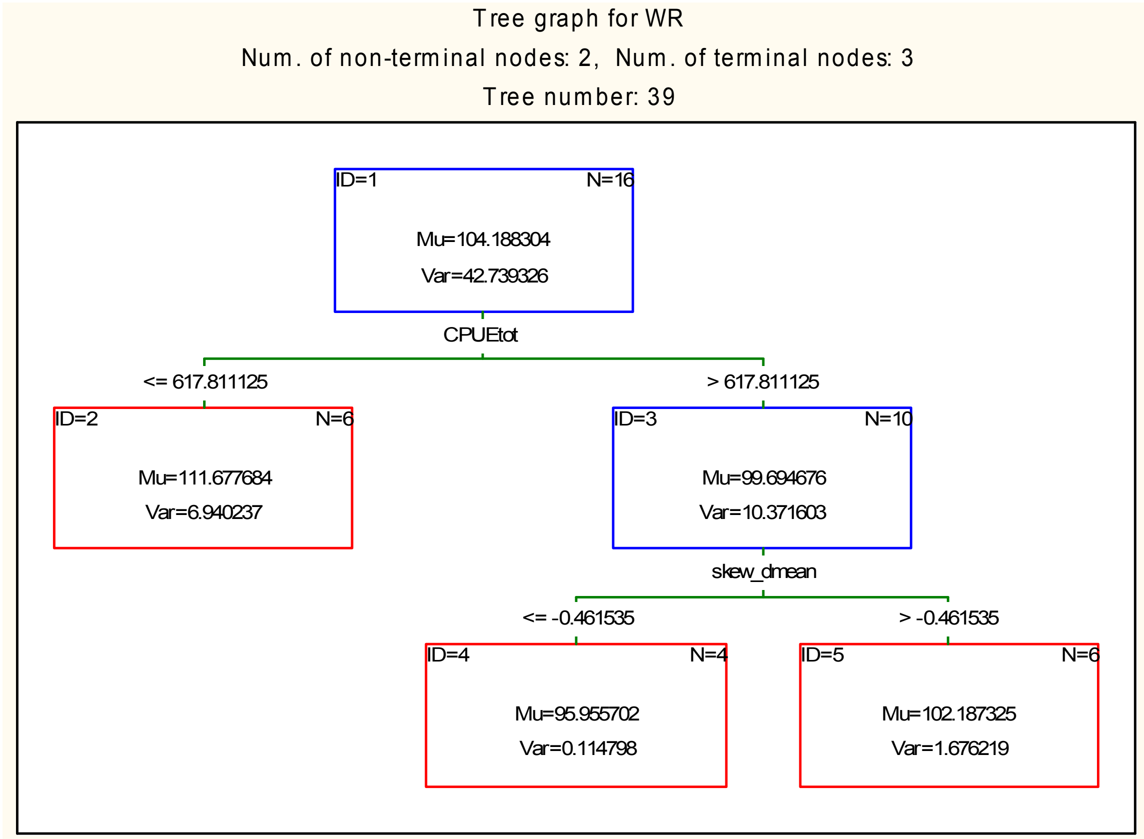


FIGURE C.1 One of 2,000 Trees That Build the Random Forest Model of WR for Large Brown Trout at Tailrace (ID = group number; *N* = the number of observations within a group; Mu = variable importance measure [not used in this report]; Var = the variance in the dependent variable within a group. The independent variables used for the two divisions in this particular tree are CPUE and skew_dmean; the value used to divide the observations into two groups is shown above each box.)

from years when skew_dmean was greater than -0.46. These three resulting groups could not be divided any further.

Random Forest Parameters That May Influence the Outcome of the Runs

- The size of forest (*ntree*)—set to 2,000 trees (default is 500).

- The number of independent variables randomly sampled as candidates at each split (*mtry*)—set to 7 (default is total number of independent variables divided by three = 7 for our Model 1).

- The minimum number of observations in each terminal node (*nodesize*)—set to 5 observations (default is 5).

C.2 OUTPUT FROM A RANDOM FOREST RUN

Output from a random forest run consists of four separate components, as described below.

1. Mean Square Error

Mean square error (MSE) is defined as the sum of squared residuals divided by the sample size.

2. Pseudo-R^2

The pseudo-R^2 is the percent variance explained (% Var Explained). It is defined as 1 – (mean square error/variance [response]). The pseudo-R^2 is an indication of how well the independent variables in the model are able to explain the variation in the dependent variables *relative to other models on the same data set*. Thus, it should not be interpreted independently or compared across datasets, but should be instead to compare different sets of independent variables within each group. Pseudo-R^2 can also be viewed as an adjusted R^2 (using mean squares instead of sum of squares) that uses the OOB estimates of MSE. The value of the pseudo-R^2 is very small, zero, or even negative when there is very little or no explanatory power in the independent variables and when the model performs no better than at random. A model with a pseudo-R^2 of 70% or greater is considered a rather good model (Liaw 2009).

3. Variable Importance

The random forest generates two different measures of variable importance: (1) Permutation Importance (%IncMSE; prediction accuracy on the OOB portion of the data) and (2) IncNodePurity (residual sum of squares, or the fit of each variable to the modeled tree). IncNodePurity is not reported in this report because %IncMSE is easier to interpret.

%IncMSE (permutation importance) measures the difference in prediction accuracy on the OOB data before and after permutation of each predictor variable, averaged over all trees. If an independent variable is important for the dependent variable, the prediction accuracy of the model decreases substantially after permutation of this independent variable (seen as a high %IncMSE value). If, on the other hand, the independent variable is not associated with the dependent variable, the independent variable is either not included in the tree (importance is zero), or it is included in the tree by chance (results only in a small random change in prediction accuracy). By chance, %IncMSE can have a negative value when the permuted variable is better suited than the original variable for splitting the data, and thereby increases the prediction accuracy of the model. All variables whose importance is a negative, zero, or small positive value (within the range of the negative values) are not informative in the model (Stroble et al. 2009).

4. Partial Dependence Plots

Partial dependence plots are used to describe the relationships between a descriptor variable and the response variable after accounting for the (average) joint effect of the other descriptor variables.

C.3 EXAMPLE OF HOW TO PREPARE DATA AND RUN A RANDOM FOREST MODEL IN R

Prepare data file (Troutdata.txt) with the following columns: WR; RatioWR; Trout Size; Trout Species; Location; Year; CPUE Variables; Flow Variables; and Macroinvertebrate Variables.

1. Read data:
> getwd()
> setwd("…")
> Trout <- read.delim("Troutdata.txt")

2. Summarize to examine data:
Summary statistics:
> summary(Trout)

3. Prepare data for the random forest:
Evoke the random forest module
>require(randomForest)

Set random seed to obtain the same results in every run on the same data
> set.seed(112233)

Impute missing data (in this case CPUE in 1996 at LH; empty cells are replaced with estimated values calculated from a proximity matrix; since this post occurs six times in the sheet, we obtain equally many values for CPUE; using Excel®, we can replace these imputed values with an average value of the imputed CPUE values)
> Data.imp <- rfImpute(WR ~ ., data=WR)

Remove rows with missing data in dependent variable and select variables:
> WR.imp = subset(Data.imp, complete.cases(WR), select = c(1,3:45))
> ratioWR.imp = subset(Data.imp, complete.cases(ratioWR), select = c(2:45))

4. Run the random forest (specify variables and cases):
Model 1: CPUE and Flows (.rf1)
> set.seed(112233)
> WRBTR23.rf1 <- randomForest(WR ~., data = WR.imp [WR.imp$Location=="TR" & WR.imp$Species=="B" & WR.imp$Size=="23", c(1,8,9:28)], ntree = 2000, importance = TRUE, proximity = TRUE)

Model 2: CPUE, Flows and Macroinvertebrates (.rf2)
> set.seed(112233)
> WRBTR23.rf2 <- randomForest(WR ~., data = WR.imp [WR.imp$Year>1994 & WR.imp$Location==“TR” & WR.imp$Species==“B” & WR.imp$Size==23, c(1,8,9:28,29,34,39,44)], ntree = 2000, importance = TRUE, proximity = TRUE)

5. Request Output:
Obtain measure of how well the independent variables can explain the variation in the dependent variable (% Var Explained, also called pseudo-R^2):
> WRBTR23.rf1

Plot error rate MSE of the random forest:
> plot(WRBTR23.rf1, lty = 1)

Plot the importance of independent variables (%incMSE—external, based on OOB, % contribution to forecasting; and IncNodePurity (or Gini)—internal, contributions to fit the modelled tree.
> varImpPlot(WRBTR23.rf1)

List %incMSE and IncNodePurity:
> round(importance(WRBTR23.rf1), 2)

Plot partial dependence plots to illustrate marginal effects for each independent variable (the relationship between the dependent variable and a given independent variable is averaged within the joint values of the other independent variables as they are represented in a tree structure; that is, the other independent variables are being “held constant”).
> partialPlot(WRBTR23.rf1, WR.imp, mean_dmean)

C.4 LITERATURE CITED

Liaw, A., 2009, R-help mailing list communication, July 21, 2009.

Stroble, C., J. Malley, and G. Tutz, 2009, *An Introduction to Recursive Partitioning*, Technical Report Number 55, Department of Statistics, University of Munich.

Environmental Science Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 240
Argonne, IL 60439-4847

www.anl.gov

U.S. DEPARTMENT OF ENERGY

Argonne National Laboratory is a U.S. Department of Energy
laboratory managed by UChicago Argonne, LLC