MODELING FRESHWATER MUSSEL DISTRIBUTION IN RELATION TO
BIOTIC AND ABIOTIC HABITAT VARIABLES IN THE
MIDDLE FORK JOHN DAY RIVER, OREGON

by

Ericka E. Hegeman

A thesis submitted in partial fulfillment
of the requirements for the degree
of
MASTER OF SCIENCE
in
Ecology

Approved:

__________________________  ________________________
Scott W. Miller            Karen E. Mock
Major Professor            Committee Member

__________________________  ________________________
Joseph M. Wheaton          Mark R. McLellan
Committee Member           Vice President for Research and
                           Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012
ABSTRACT

Modeling Freshwater Mussel Distribution in Relation to Biotic and Abiotic Habitat Variables in the Middle Fork John Day River, Oregon

by

Ericka E. Hegeman, Master of Science
Utah State University, 2012

Major Professor: Dr. Scott W. Miller
Department: Watershed Sciences

The habitat requirements of western freshwater mussels, *Anodonta*, *Gonidea*, and *Margaritifera*, remain unclear despite their imperiled status. Freshwater mussels provide a series of ecosystem services including habitat enhancement, substratum stabilization, nutrient cycling, and water clarification, which makes their loss from aquatic ecosystems particularly detrimental. To improve the efficacy of restoration actions targeting these organisms, I used random forest modeling to investigate the biotic and abiotic factors influencing mussel density and distribution throughout a 55-kilometer (km) segment of the Middle Fork John Day River (MFJDR), in northeastern Oregon. Data was collected to characterize the occurrence of mussels with respect to the hierarchical, hydrogeomorphic structure of habitat within reaches of varying valley confinement and channel units nested within these reaches. Data regarding functional habitat features were also included to ensure that models included the wide range of characteristics that mussels need from their environment. By collecting data at both the reach and channel unit scale, I was able to investigate how mussel densities and distributions vary with spatial scale and other biophysical parameters. Throughout the study area, *Margaritifera* density exhibited a unimodal distribution with respect to river km, while *Anodonta* and *Gonidea* density showed a negative
relationship with river km and exhibited higher densities downstream. The large scale, longitudinal trends of *Margaritifera* were related to hydrogeomorphic characteristics at the reach scale, while less than half of the longitudinal variation in *Anodonta* and *Gonidea* were explained by hydrogeomorphic and water quality parameters. At the channel unit scale, all mussel genera responded to the patchy variation in physical habitat characteristics, particularly habitat factors that indicated more stable parts of the channel. Overall, physical habitat characteristics such as woody debris, emergent aquatic vegetation, coarse substratum, and channel morphology were more important than hydraulic, biotic, and chemical variables. These results suggest that at both the reach and channel unit scales, mussel density and distribution are influenced by high flow refugia and the hierarchical structuring of hydrogeomorphic habitat characteristics. These results will assist mussel restoration efforts by providing specific guidance about the types of physical habitat conditions that are suitable for mussels.

(57 pages)
PUBLIC ABSTRACT

Modeling Freshwater Mussel Distribution in Relation to Biotic and Abiotic Habitat Variables in the Middle Fork John Day River, Oregon

by

Ericka E. Hegeman, Master of Science
Utah State University, 2012

Major Professor: Dr. Scott W. Miller
Department: Watershed Sciences

Freshwater mussels are the most threatened taxonomic group in North America with extinction rates that exceed those of many species found in both terrestrial and freshwater ecosystems including fish, birds, and amphibians. Part of the reason that mussels are so threatened is because their larvae are parasitic on fish, making the completion of their life cycle dependent upon healthy fish populations. The imperilment of freshwater mussels is a cause for concern because of the benefits that mussels provide to freshwater ecosystems including habitat enhancement, substratum stabilization, nutrient cycling, and water clarification. Restoration and conservation efforts targeting western freshwater mussels have been constrained by a lack of information about habitat requirements. As a result, I was interested in investigating how mussel density and distribution varied with respect to both biotic and abiotic factors at multiple spatial scales.

I used a modeling approach to determine which habitat parameters were associated with mussel distribution and density throughout a 55-kilometer (km) of the Middle Fork John Day River, Oregon. Parameters included physical stream habitat characteristics, host fish presence, water quality measures, and mussel food quantity and quality. Results of this analysis indicated
that mussels responded to the hierarchical arrangement of physical habitat from the valley segment to the channel unit and that higher densities of mussels were found in parts of the river that were more stable at high flows. I found that the distribution of host fish was not limiting to mussels in this river system and that the overall physical habitat characteristics such as gravel size, silt cover, and woody debris were most important to explaining mussel density and distribution. These results will assist mussel restoration efforts by providing specific guidance about the types of habitat conditions that are suitable for mussels at multiple spatial scales.
I would like to thank the Confederated Tribes of the Umatilla Indian Reservation (CTUIR) for their generous funding of this research. I am indebted to Jayne Brim Box for spending numerous hours snorkeling with me and sharing her vast knowledge of mussels. Christine O’Brien, Scott O’Daniel, Donna Nez, and other CTUIR natural resources staff also helped provide support and assistance to this project.

I am thankful to have had a committee with such a diverse set of backgrounds, each of which has come into play at various points during my thesis work. My major professor, Dr. Scott Miller, has shown incredible patience and perseverance on this project while personally helping me understand everything from the role of spatial scale to quantitative study design and I am very proud to be his first graduate student. Dr. Karen Mock’s infectious optimism and words of wisdom at just the right moment helped me get through some rough periods. Dr. Joe Wheaton was particularly helpful at providing a wide range of ideas regarding how mussels may respond to the geomorphic template of a river at multiple spatial scales.

I would also like to thank field technicians Melissa Prechtl and Dan Weber for helping me snorkel more than 12 kilometers of river and count over 50,000 mussels. My dog, Zoey, has been my constant companion throughout graduate school and has been with me in the field, in the office, and for afternoon trail runs and cross-country skiing. Finally, I never would have been able to make the big step of returning to school without the encouragement and support of my wonderful husband, Matt.

Ericka Hegeman
CONTENTS

Page

ABSTRACT ......................................................................................................................... iii
PUBLIC ABSTRACT ........................................................................................................... v
ACKNOWLEDGMENTS ...................................................................................................... vii
LIST OF TABLES .............................................................................................................. ix
LIST OF FIGURES ........................................................................................................... x
INTRODUCTION ................................................................................................................ 1
METHODS ......................................................................................................................... 5
    Study area .................................................................................................................... 5
    Study design ............................................................................................................... 7
    Mussel and habitat surveys ......................................................................................... 8
    Statistical analysis ..................................................................................................... 12
RESULTS .......................................................................................................................... 16
    Mussel density and distribution ................................................................................. 16
    Spatial patterns in biotic and abiotic habitat variables ................................................ 20
    Mussel habitat associations ....................................................................................... 24
DISCUSSION ..................................................................................................................... 30
    Mussel habitat relationships ...................................................................................... 30
    Unexpected results ..................................................................................................... 34
CONCLUSIONS ................................................................................................................ 37
REFERENCES ..................................................................................................................... 39
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Channel unit and reach scale variables included in development of random forest models of mussel density organized by variable type. The range, mean, and CV are reported for the smallest spatial scale at which data was collected; the channel unit or reach. Asterisk (*) symbol indicates variable measured or computed only at the reach scale.</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Variables included in the best channel unit (top) and reach (bottom) scale models for all three genera. The percent variance explained with the addition of river km at the channel unit scale to account for spatial autocorrelation is shown. Variable relationships are summarized as positive (+), negative (-), and flat (f). Variables for each model are listed in their order of importance, as indicated by the percent increase in mean squared error (MSE) ± the standard deviation (SD). Variable name abbreviations are explained in Table 1.</td>
<td>25</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Map of the upper MFJDR illustrating the longitudinal distribution of wide and narrow valley segments along the 55-km study area. Inset highlights the hierarchical study design where two high and two low gradient reaches were located within a segment and sampled at the channel unit scale.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Longitudinal patterns of reach scale mean (± 1 within reach standard deviation) mussel density for (a) <em>Margaritifera</em>, (b) <em>Anodonta</em>, and (c) <em>Gonidea</em> throughout the 55-km study area. Black circles denote reaches located in wide valley segments and white circles denote reaches in narrow segments. Note the different y-axis scales among genera.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Box plots of (a) <em>Margaritifera</em>, (b) <em>Anodonta</em>, and (c) <em>Gonidea</em> density within narrow (white) and wide (dark grey) valley confinement. Y-axis is the log of mussel density plus the smallest non-zero value and circles indicate outliers beyond the first and third quartiles. Note the different y-axis scales among genera.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Box plots of (a) <em>Margaritifera</em>, (b) <em>Anodonta</em>, and (c) <em>Gonidea</em> density within glides, pools, riffles, and runs. Y-axis is the log of mussel density plus the smallest non-zero value and circles indicate outliers beyond the first and third quartiles. Note the different y-axis scales among genera.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Longitudinal variation of physical habitat variables (reach average) (a) bankfull width to depth ratio, (b) sinuosity, (c) LWD density, (d) percent silt, (e) percent emergent vegetation, (f) D&lt;sub&gt;16&lt;/sub&gt; particle size, (g) percent slope, and (h) Pfankuch bed stability throughout the study area. Black circles denote wide valley segments and white circles denote narrow valley segments.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Longitudinal variation in the spatial extent percent linear extent of (a) pool, (b) glide, (c) riffle, and (d) run channel units within individual reaches. Black circles denote wide valley segments and white circles denote narrow valley segments.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Longitudinal variation in water quality variables of (a) specific conductance and (c) water temperature and biotic variables of (b) relative coldwater fish abundance, (d) relative coolwater fish abundance, and (e) autotrophic index (reach average) throughout the study area. Black circles denote wide valley segments and white circles denote narrow valley segments.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Partial dependence plots of the variables in the final channel unit scale random forest models for each of the three genera. Plots indicate mussel density based on each predictor variable in the best models after averaging out the effects of all other predictor variables in the model. Rug plots indicate deciles of data for each predictor variable. Variable importance for river km not shown and relative importance of variables for each model decreases from left to right as reported in</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. .............................................................................................................................................. 26

9 Partial dependence plots of the variables in the final reach scale random forest models for each of the three genera. Plots indicate mussel density based on each predictor variable in the best models after averaging out the effects of all other predictor variables in the model. Rug plots indicate deciles of data for each predictor variable. Relative importance of variables for each model decreases from left to right. .............................................................................................................................................. 27

10 Variable importance plots from the entire set of variables used to construct best models at the (a) channel unit and (b) reach scale. ........................................................................................................... 29
INTRODUCTION

Freshwater mussels are among the most threatened faunal groups worldwide (Haag and Rypel, 2011; Régnier, Fontaine & Bouchet, 2009) and in North America, where the greatest mussel species diversity occurs (Williams, Warren & Cummings, 1993), recent extinction rates exceed the rates of all other aquatic and terrestrial fauna (Ricciardi & Rasmussen, 1999). Freshwater mussels are rapidly declining because of habitat destruction (Fuller, 1974; Williams et al., 1993; Wilcove et al., 1998), impoundments and dam construction (Vaughn & Taylor, 1999; Watters, 2000), pollution (Aldridge, Payne & Miller, 1987; Naimo, 1995; Wilcove et al., 1998), and overharvest (Claassen, 1994; Strayer et al., 2004). In addition, due to their dependence on host fishes to complete their life cycle (Fuller, 1974; Watters, 1992), mussels are threatened both directly by human impacts and indirectly through the decline of their host fish populations (Bogan, 1993). The loss of mussels from freshwater ecosystems is particularly detrimental because of the ecosystem services they provide, including nutrient cycling (Vaughn & Hakenkamp, 2001; Vaughn, Gido & Spooner, 2004), habitat creation for other benthic organisms (Spooner & Vaughn, 2006; Limm & Power, 2011) and water filtration (Kryger & Riisgård, 1988; Strayer et al., 1994; Howard & Cuffey, 2006).

In the western United States, only a small number of quantitative, peer-reviewed freshwater mussel studies have been conducted, but several of these studies have described distinctive distributional trends and responses to hydrogeomorphic (e.g. shear stress, slope, substratum composition, channel morphology) habitat characteristics. For example, surveys of mussel distribution in the South Fork Eel River, California (Howard & Cuffey, 2003), and the Middle Fork John Day River (MFJDR), Oregon (Howard, 2005; Brim Box et al., 2006), have shown genus-specific distributional trends at scales ranging from the watershed to the channel unit scale. In particular, *Margaritifera* are more widely distributed than *Anodonta* and *Gonidea,*
which tend to be restricted to lower gradient, higher order river segments (Howard & Cuffey, 2003; Howard, 2005; Brim Box et al., 2006). It has been suggested that the composition and distribution of host fishes, spatial patterns in food resource availability, and disparate metabolic rates among genera may be the cause for these large scale, longitudinal gradients (Bauer, Hochwald & Silkenat, 1991; Howard, 2005; Brim Box et al., 2006), but little empirical data exist to support these hypotheses. At smaller spatial scales, Howard & Cuffey (2003) and Howard (2005) found that Anodonta and Margaritifera were positively associated with both pools and runs. In contrast, Vannote & Minshall (1982) found Margaritifera and Gonidea almost exclusively in runs in the Salmon River, Idaho, highlighting the uncertainty about channel unit scale habitat requirements for western freshwater mussels.

Stream habitats are organized in a nested hierarchical manner such that larger scale habitat features influence the presence of smaller scale habitat types (Frissell et al., 1986; Montgomery & Buffington, 1998), which results in predictable distributional patterns of biota across the landscape (Poff, 1997). Investigating species-habitat relationships at multiple spatial scales can help clarify the role of different habitat characteristics at individual spatial scales. Such multi-scale habitat investigations are common with both fishes (Torgersen et al., 1999; Baxter & Hauer, 2000; Torgersen & Close, 2004) and invertebrates (Parsons, Thoms & Norris, 2003; Hutchens et al., 2009), but few studies have been conducted with freshwater mussels and most investigations tend to focus on a single spatial scale (but see Howard & Cuffey, 2003; McRae, Allan & Burch, 2004; Hopkins, 2009). Based on the multi-scale species habitat relationships seen with fish and invertebrates, I expected that mussels would respond to the hierarchical structuring of stream habitat in similar ways.

Studies of mussels conducted at single spatial scales also provide evidence that mussels respond to different habitat characteristics at scales ranging from the watershed to the sub-meter. For example, at the watershed scale, topographic relief and soil erosion potential have been
shown to influence mussel distribution (Arbuckle & Downing, 2002) and these large scale factors modify reach scale shear stress and subsequent substratum stability (Howard & Cuffey, 2003; Gangloff & Feminella, 2007), resulting in mussels being found in areas providing high energy flow refugia. At the sub-meter scale, these larger scale habitat characteristics influence substratum size (Vannote & Minshall, 1982; Vaughn, 1997; McRae et al., 2004) and riparian vegetation such as the presence of sedge root mats (Howard & Cuffey, 2003) indicating that mussels have differential habitat use with spatial scale. In addition to the hierarchical arrangement of physical habitat in streams, habitat models of freshwater mussels could benefit from the inclusion of functional habitat characteristics such as host fish presence, water quality parameters, and food availability (Newton, Woolnough & Strayer, 2008; Strayer, 2008). Since mussels respond to different habitat variables at different spatial scales, a habitat model that incorporates both biotic and abiotic habitat factors at multiple spatial scales may clarify the factors causing the unique, multi-scale distributional trends of western freshwater mussels.

Using quantitative hierarchical habitat data to guide mussel restoration efforts can provide a systematic description of locations to target at multiple spatial scales. To date, management and conservation efforts have been constrained by a lack of quantitative information regarding species-habitat relationships, particularly west of the continental divide where Unionoidea diversity consists of three genera, Anodonta, Gonidea, and Margaritifera (Brim Box et al., 2006). To effectively protect and restore western freshwater mussel populations we need to further develop an explicit understanding of the habitat parameters that sustain the growth, reproduction, and, survival of mussel populations across multiple life stages.

The goal of this study was to provide information on habitat use by western freshwater mussels using both the functional habitat needs and the hierarchical structuring of hydrogeomorphic habitat throughout a 55-kilometer (km) section of the upper MFJDR. First, I quantified the distribution and density of freshwater mussels throughout the study area. Next, I
assessed the relative importance of both biotic and abiotic habitat predictors to explain patterns in mussel density and distribution. Finally, I investigated how mussel density and distribution changed with respect to the hydrogeomorphic template of the MFJDR at multiple spatial scales: the sub-watershed, the reach, and the channel unit. Ultimately, the findings from this study will help target locations in other watersheds where similar habitat characteristics may be found and, thus, where suitable habitat for mussel restoration efforts may occur.
METHODS

Study area

The study was conducted in a 55-km section of the upper MFJDR located in northeastern Oregon (Fig. 1). The MFJDR flows for 117 km and drains a watershed of approximately 2,050 km² before joining the North Fork John Day River. Elevations in the watershed range from 2,480 m in the headwaters of the Blue Mountains to 670 m at the confluence. Runoff in this snowmelt-dominated system typically occurs from March to May, with a mean daily peak discharge of 60 m³/s. From August to November, base flows (<1.5 m³/s) dominate, although low flows can last through winter. The MFJDR is also susceptible to late winter and early spring rain-on-snow events that produce short duration high flow events comparable to peaks occurring during spring runoff. Data collected for this study occurred after the second highest flood event in the 80-year mean daily flow gage record (129 m³/s) with a recurrence interval of 49 years (USGS gage no. 14044000).

The upper 55-km section of the MJFDR was selected for study because it contains all three genera of western freshwater mussels (Brim Box et al., 2006) and has been the subject of previous research on geomorphologic habitat controls (McDowell, 2001) and salmonid and lamprey habitat use (Torgersen et al., 1999; Torgersen & Close, 2004). In addition to providing data for this study, these investigations found biotic responses to small scale geomorphic patterns that are likely influenced by segment level variation in valley confinement. Specifically, the upper portion of the MFJDR flows through fourteen alternating valley segments of wide, narrow, and intermediate confinement, which were delineated based upon valley width, tributary junctions, and cultural features such as bridges (McDowell, 2001).
The landscape of this semi-arid watershed is dominated by mixed ponderosa conifer forest with lesser amounts of perennial grasslands and shrub plant communities (Kauffman, Thorpe & Brookshire, 2004). Riparian vegetation consists of woody species such as willow, hawthorn, alder, and wild rose, as well as various sedges and grasses (Torgersen et al., 1999; Beschta & Ripple, 2005). Historical modifications to the watershed include dredge mining, channel straightening, and road construction (McDowell, 2001; Torgersen & Close, 2004) and current land use modification consists primarily of cattle grazing and logging. Compared to
narrow valley segments, wide valley segments have experienced greater human modification, which has resulted in channelization, decreases in sinuosity, and loss of large woody debris (McDowell, 2000).

**Study design**

I used a stratified random sampling design to evaluate mussel density and distribution with respect to a variety of habitat parameters at multiple spatial scales from the sub-watershed to the channel unit. Specifically, I stratified by valley confinement (wide or narrow) at the coarsest spatial scale and slope (low or high) at smaller spatial scales to randomly select reaches within individual valley segments. Overall, I selected 46 reaches, which comprised 18 reaches within the five wide segments and 28 reaches within the eight narrow segments.

McDowell (2001) noted that narrow valley segments have a valley width of 10 bankfull channel widths or less and wide segments have a valley width of 10 to 20 times bankfull width. Therefore, I used the ratio of valley width (GIS derived) to bankfull width (field measured) to empirically verify McDowell’s classifications. Based upon these delineations, I found that the lone intermediate segment had an average valley width of 13 times bankfull width and subsequently reclassified this segment to wide.

Slope was determined by partitioning narrow and wide segments into 200 m reaches for which slope was calculated in ArcGIS 10 using a one-meter digital elevation model (DEM) (Watershed Sciences, 2006) and then categorized as low (<0.5%) or high (>0.5%). I randomly selected two high and two low gradient reaches from all possible reaches within each segment, except for eight reaches where property access was not granted (Fig. 1). Reach length was determined as a function of bankfull width (20 times bankfull) with a minimum reach length of 200 meters and a maximum of 300 meters (Heitke et al., 2011). The minimum reach length was
established to ensure that sufficient geomorphic variability was sampled within each reach, while the infeasibility of surveying reaches longer than 300 m necessitated a maximum reach length.

Mussel and habitat surveys

To determine species-habitat relationships at multiple spatial scales, I collected a series of habitat data in conjunction with mussel surveys. Specifically, four main categories of habitat data were collected: hydrogeomorphic, biotic, water quality, and human impacts (Table 1); all of which have been identified as important by previous research (Bauer et al., 1991; Watters, 1992; Hastie, Boon & Young, 2000; Vaughn & Taylor, 2000; Howard & Cuffey, 2003; McRae et al., 2004; Howard, 2005; Galbraith & Vaughn, 2009; Galbraith, Blakeslee & Lellis, 2012). The hydrogeomorphic parameters included a broad array of physical habitat characteristics such as substratum size, emergent vegetation, channel morphology, slope, and hydraulic forces. Biotic parameters of host fish presence and mussel food quality and quantity were collected to assess potential differences in metabolic rate among mussel genera. Water quality parameters and a measure of human impacts were also included to account for potential limitations on mussel distribution.

Visual snorkel surveys were used to assess the composition of individual mussel genera for each channel unit (pool, riffle, run, or glide) during the summer of 2011. Snorkel surveys began at the downstream end of each reach with a pair of snorkelers moving upstream approximately two meters apart and searching all possible habitats until the entire channel unit was systemically searched (Howard & Cuffey, 2003). In wide reaches, up to five passes were used to search a channel unit, while in narrow reaches, only one pass was necessary. Snorkeling was used except in very shallow areas (<10 cm) where mussels could be detected by wading or using a clear-bottom bucket in more turbulent areas. Mussels were identified to genus based on shell morphology and mantle margins. While Margaritifera falcata and Gonidea angulata can be
identified to species, this level of taxonomic identification was not attempted with *Anodonta* due to recent genetic analysis suggesting a potential taxonomic restructuring of this genus (Chong *et al*., 2008; Mock *et al*., 2010); all analyses were subsequently conducted at the genus level.

Linear mussel density (number of mussels per meter length of channel) was used as the model response variable for the two spatial scales modeled: the channel unit and the reach. Both linear and areal mussel densities were computed and were strongly correlated ($r > 0.9$ for all three genera, $P < 0.0001$). However, linear mussel density was selected because the majority of mussels were found along the banks and including the area of the river bottom dampened the high density trends in wider parts of the river. The precision of visual mussel surveys was quantified by resurveying 14 randomly selected channel units of varying mussel density and channel unit type and calculating the coefficient of variation (CV) to compare error among genera. The average CV among all genera was moderately low (0.32) with *Gonidea* having the lowest CV (0.14), followed by *Margaritifera* (0.28), and *Anodonta* (0.56).

Prior to mussel sampling, a series of habitat variables were measured at each channel unit within a reach. Channel unit types of pool, riffle, and run were identified using channel morphology and surface turbulence according to Heitke *et al*. (2011), while glides were identified as deep, uniform, laminar flow areas as described by Torgersen (2007). The length of each channel unit was measured and the maximum and minimum water depth along the thalweg was determined. In addition, wetted width was measured at three evenly spaced locations within each channel unit and bankfull width and maximum bankfull depth were measured at a single transect. Pebble counts were conducted in a zigzag manner with a minimum of 20 pebbles counted per channel unit (Wolman, 1954). The minimum pebble count for each channel unit was increased when less than five channel units were present in a reach to ensure a combined minimum of 100 pebbles per reach. Cumulative substratum size class categories (Harrelson, Rawlins & Potyondy,
1994) were calculated ($D_{50}$ only for the channel unit scale) and Shannon’s diversity was computed from the pebble counts at both spatial scales.

In addition to quantitative habitat measurements, several qualitative measures of physical habitat were obtained. Bed stability was assessed for each channel unit using a modified Pfankuch channel stability evaluation (Pfankuch, 1975) in which lower scores indicate greater bed stability. We did not include the upper and lower bank components of Pfankuch channel stability evaluation and instead used only the streambed component since it may be more applicable to benthic communities (Death & Winterbourn, 1994). Percent silt cover was visually scored from one (<25% of the bottom covered with silt) to five (100% of the bottom covered with silt and substratum interstitial spaces completely filled). Similarly, percent emergent aquatic vegetation was also scored one (<25% of both banks covered by emergent aquatic vegetation) to five (100% of both banks covered); with the left and right bank each allocated 50% of the total. Finally, the linear density of large woody debris (LWD) was computed for each channel unit and reach. Our definition of LWD included LWD >1 m in length and >10 cm in diameter (Heitke et al., 2011), as well as any large woody plants and root wads in contact with the channel at high flows. LWD density was calculated based upon the number of LWD pieces per length of channel unit or reach. Median values of bankfull width to depth, emergent aquatic vegetation, maximum depth, minimum depth, Pfankuch bed stability, and silt were calculated for use in the reach scale models. Percent channel unit type was calculated based upon the total length of each channel unit type within a reach.

Several habitat variables were measured only at the reach scale and, thus, were only included as predictors at this larger spatial scale. I used a rapid habitat bioassessment to assess human impacts with evaluation criteria selected from both the US Environmental Protection Agency Rapid Habitat Bioassessment Protocol (Barbour et al., 1999) and the US Department of Agriculture Stream Visual Assessment Protocol (USDA Natural Resources Conservation Service,
1998); higher values in these assessments indicated less human influence. A series of water quality and mussel food parameters were collected during a four-day period in late September. Specific conductance, pH, and water temperature were measured once per reach using a YSI Model 85 Handheld meter. As a measure of potential food quantity and quality, seston and chlorophyll \( a \) were sampled by filtering measured quantities of water at each reach through glass fiber filters using a GeoPump\textsuperscript{TM} peristaltic pump. Seston ash free dry mass (AFDM) was determined by ashing filters in a muffle furnace using the methods of Hauer & Lamberti (2006). Chlorophyll \( a \) concentration was determined by hot ethanol extraction and spectrophotometric analysis (Hauer & Lamberti, 2006). The autotrophic index was computed by dividing AFDM (mg/L) by chlorophyll \( a \) (mg/L), where lower values indicated higher quality seston food resource availability.

Additional reach scale habitat variables were obtained by compiling spatially explicit datasets obtained from published sources or derived from a geographic information system (GIS). Specifically, fish abundance data (Torgersen \textit{et al.}, 2006) was spatially aligned to sampled reaches in ArcGIS version 10 and relative abundance was quantified for coldwater (salmonids) and coolwater (castomids and cyprinids) assemblages by dividing reach level abundance values for each assemblage by the maximum possible value per reach. Sinuosity was calculated in ArcGIS by dividing reach length by valley length using a one-meter DEM (Watershed Sciences, 2006). The median August water temperature at each reach was obtained from forward looking infrared (FLIR) aerial surveys provided by the Oregon Department of Environmental Quality from 2003 (Watershed Sciences, 2004). August water temperatures were considered potentially limiting to mussels since flow is low and temperatures are high, conditions that have been shown to limit mussel distribution (Golladay \textit{et al.}, 2004). Median temperature values were assumed appropriate since within reach temperature variation was low (CV ≤ 0.01). Hydraulic characteristics of shear stress and relative bed stability (RBS) were calculated using slope,
bankfull depth, and substratum size at the reach scale. Since bankfull width was >10 times bankfull depth for the majority of the study reaches, I assumed a rectangular channel existed where the hydraulic radius was equivalent to bankfull depth and used this variable when calculating shear stress (Anderson & Anderson, 2010). RBS was calculated using guidance from Kaufmann et al. (1999) by dividing the substratum $D_{50}$ by the average critical diameter at bankfull.

Statistical analysis

Random forest modeling (Breiman, 2001; Liaw & Wiener, 2002) was used to quantify the relationship between linear mussel density and habitat variables at the channel unit and reach scales. Random forest is a tree-based tool that uses bootstrap sampling and fits many classification or regression trees to a data set (Breiman, 2001). At every split of the data, or node on a tree, a random set of variables is used from a bootstrap sample of the data. The results from all trees are averaged resulting in a single prediction from the many binary trees. The algorithm is robust to outliers, prevents over fitting, and can handle a large number of categorical and continuous variables (Breiman, 2001; Cutler et al., 2007). I ran random forest in regression mode and increased the default number of trees to 5000 per model run to increase stability. Model performance was assessed with percent variance explained, which is an internal cross-validated metric defined as 1- (mean squared error)/(variance (response)), and can be thought of as a pseudo r-squared (Pang et al., 2006). Variable importance was assessed using the percent increase in mean square error (MSE), with higher MSE values indicating greater variable importance (Goodwin et al., 2008).

Using all possible variables at either the channel and reach scale, models were developed for each of the three genera by successively removing variables until the percent variance explained was maximized. In addition to these “best” models, comparative variable importance
plots were prepared using all variables at each spatial scale to compare the relative importance of individual predictor variables across all three genera. Variables tested in the channel unit scale models included only hydrogeomorphic habitat variables such as substratum size, channel unit type, large woody debris, and emergent vegetation that were explicitly measured for each channel unit. At the reach scale, I used variables from all four habitat categories (Table 1) such as relative host fish abundance, water quality, hydraulic forces, and human impacts. Since correlations existed between some variables, I investigated Pearson correlation coefficients between all variables. For pairs of highly correlated variables \( r > 0.7 \), I removed the variable with the lowest MSE through the variable selection process. Models were then optimized to minimize model error by modifying the default number of variables tested at each node. I ran each model ten times to obtain mean values for percent variance explained, percent increase in MSE, and standard deviation (SD) (Table 2). Partial dependence plots were examined to investigate the relationship of individual predictors with the response variable by holding the effects of all other predictors in the model constant (Cutler et al., 2007).

Due to the longitudinal distribution of mussels throughout the study area, I suspected that spatial autocorrelation might exist in the data, which would violate the statistical assumption of independence among observations. In particular, locations with high mussel density tended to be clustered together, which could result in overinflated model performance if not addressed. To quantify the extent of spatial autocorrelation, I used Moran’s I to test the residuals from the best models against river km. I assessed spatial autocorrelation by using a threshold Moran’s I z-value of >1.96 at the 0.05 alpha level. This test indicated significant autocorrelation for all three genera among channel units, but not among reaches. Therefore, I added river km as a predictor variable to the best channel unit scale models and then re-ran the Moran’s I test using the residuals from these modified models. After adding river km, the effect of spatial autocorrelation was greatly reduced for Anodonta \( z = 2.4, P < 0.01 \) and Margaritifera \( z = 2.5, P < 0.01 \) to nearly the level.
Table 1. Channel unit and reach scale variables included in development of random forest models of mussel density organized by variable type. The range, mean, and CV are reported for the smallest spatial scale at which data was collected; the channel unit or reach. Asterisk (*) symbol indicates variable measured or computed only at the reach scale.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Name</th>
<th>Range</th>
<th>Mean</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hydrogeomorphic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BkfWtoD</td>
<td>Ratio of bankfull width to depth</td>
<td>3.2 – 63.8</td>
<td>17.6</td>
<td>0.6</td>
</tr>
<tr>
<td>ChanUnit</td>
<td>Channel unit type</td>
<td>Pool, riffle, run, glide</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>ChanUnitH*</td>
<td>Channel unit diversity</td>
<td>0 – 1.4</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Confine*</td>
<td>Valley confinement(^1)</td>
<td>Wide or Narrow</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>D(_{16})</td>
<td>D(_{16}) particle size (mm)</td>
<td>2 – 54</td>
<td>22</td>
<td>0.6</td>
</tr>
<tr>
<td>D(_{50})</td>
<td>D(_{50}) particle size (mm)</td>
<td>2 – 512</td>
<td>78</td>
<td>0.7</td>
</tr>
<tr>
<td>D(_{84})*</td>
<td>D(_{84}) particle size (mm)</td>
<td>37 – 512</td>
<td>176</td>
<td>0.5</td>
</tr>
<tr>
<td>EmergVeg</td>
<td>% emergent aquatic vegetation</td>
<td>1 – 5</td>
<td>3</td>
<td>0.4</td>
</tr>
<tr>
<td>LWD</td>
<td>Large woody debris (no./m)</td>
<td>0 – 0.4</td>
<td>0.3</td>
<td>2.0</td>
</tr>
<tr>
<td>MaxDepth</td>
<td>Maximum water depth</td>
<td>20 – 150</td>
<td>62</td>
<td>0.4</td>
</tr>
<tr>
<td>MinDepth</td>
<td>Minimum water depth (cm)</td>
<td>10 – 65</td>
<td>29</td>
<td>0.3</td>
</tr>
<tr>
<td>%Glide*</td>
<td>Percent glide</td>
<td>0 – 68</td>
<td>11</td>
<td>1.8</td>
</tr>
<tr>
<td>%Pool *</td>
<td>Percent pool</td>
<td>0 – 100</td>
<td>26</td>
<td>1.1</td>
</tr>
<tr>
<td>%Riffle*</td>
<td>Percent riffle</td>
<td>0 – 73</td>
<td>36</td>
<td>0.6</td>
</tr>
<tr>
<td>%Run*</td>
<td>Percent run</td>
<td>0 – 63</td>
<td>24</td>
<td>0.8</td>
</tr>
<tr>
<td>%Side</td>
<td>Percent side channel</td>
<td>0 – 27</td>
<td>4</td>
<td>2.1</td>
</tr>
<tr>
<td>PebbleH*</td>
<td>Substratum size diversity</td>
<td>1.1 – 2.5</td>
<td>2.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Pfankuch</td>
<td>Bed stability index</td>
<td>15 – 60</td>
<td>38</td>
<td>0.3</td>
</tr>
<tr>
<td>RBS*</td>
<td>Relative bed stability</td>
<td>0.3 – 9.2</td>
<td>1.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Silt</td>
<td>Silt cover (%)</td>
<td>1 – 5</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>Sinuosity*</td>
<td>Sinuosity</td>
<td>0.96 – 2.31</td>
<td>1.18</td>
<td>0.2</td>
</tr>
<tr>
<td>Shear*</td>
<td>Shear stress (N/m(^2))</td>
<td>9.31 – 89.51</td>
<td>36.74</td>
<td>0.5</td>
</tr>
<tr>
<td>Slope*</td>
<td>Slope (%)</td>
<td>0.16 – 1.38</td>
<td>0.55</td>
<td>0.5</td>
</tr>
<tr>
<td>ValtoBkfl</td>
<td>Valley to bankfull width ratio</td>
<td>3.3 – 84.7</td>
<td>17.5</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Biotic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChlA*</td>
<td>Chlorophyll a (mg/L)</td>
<td>0.65 – 36.85</td>
<td>3.66</td>
<td>1.4</td>
</tr>
<tr>
<td>AFDM*</td>
<td>Seston ash free dry mass (mg/L)</td>
<td>0.0003 – 0.0186</td>
<td>0.0023</td>
<td>1.2</td>
</tr>
<tr>
<td>Autotroph*</td>
<td>Autotrophic index</td>
<td>0.15 – 2.46</td>
<td>0.88</td>
<td>0.7</td>
</tr>
<tr>
<td>ColdWF*</td>
<td>Coldwater fish abundance(^2)</td>
<td>0 – 1</td>
<td>0.45</td>
<td>0.6</td>
</tr>
<tr>
<td>WarmWF*</td>
<td>Coolwater fish abundance(^2)</td>
<td>0 – 1</td>
<td>0.75</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Water quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct*</td>
<td>Specific conductance (mS/cm)</td>
<td>0.097 – 0.166</td>
<td>0.131</td>
<td>0.2</td>
</tr>
<tr>
<td>Temp03*</td>
<td>FLIR water temperature (°C)(^3)</td>
<td>19.1 – 23.6</td>
<td>21.6</td>
<td>0.04</td>
</tr>
<tr>
<td>pH*</td>
<td>pH</td>
<td>6.4 – 8.9</td>
<td>8.1</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Human impact</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPAScore*</td>
<td>Rapid bioassessment score</td>
<td>20 – 50</td>
<td>40</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\(^1\)McDowell, 2001
\(^2\)Torgersen et al. 1999; Torgersen et al. 2006
\(^3\)Watershed Sciences, 2004
of significance and for *Gonidea* (*z* = 0.5, *P* = 0.64) it was reduced below the level of significance. As a result, I concluded that spatial autocorrelation was sufficiently addressed by including river km at the channel unit scale.

To investigate whether mussel density varied among channel unit types within wide or narrow valley segments, I conducted non-parametric Wilcoxon rank sum tests (stats package, R Development Core Team, 2011) for each genus. While attempts were made to utilize a two-way ANOVA with the confinement by channel unit type interaction, such tests were not compatible with the highly right skewed, negative binomial distribution of mussel density. Wilcoxon rank sum tests were used to test the hypothesis (*α* = 0.10) that the central tendency in mussel abundance was the same between both wide and narrow valley segments and then post-hoc tests were conducted between channel unit types within each segment.
RESULTS

*Mussel density and distribution*

I observed all three genera of western freshwater mussels in the MFJDR: *Anodonta*, *Gonidea*, and *Margaritifera*. At least one individual mussel was found in each surveyed reach for a total of 53,823 individuals. Among the 46 surveyed reaches, *Margaritifera* was the most abundant with 46,248 individuals distributed across 96% of sampled reaches and 89% of channel units. *Anodonta* was found in nearly the same number of reaches (89%); however, they occurred in a smaller percentage of channel units (59%) and far fewer individuals were encountered (7,103). In contrast, both the overall abundance (472 individuals) and distribution of *Gonidea* among reaches (30%) and channel units (14%) was much more limited than the other two genera. Among all reaches sampled, the greatest average density was exhibited by *Margaritifera* (3.6 mussels/m, SD = 3.9), which exceeded both *Anodonta* (0.52 mussels/m, SD = 1.1) and *Gonidea* (0.03 mussels/m, SD = 0.08).

Although mussels were found at every reach sampled, each genus exhibited unique trends with respect to longitudinal distribution and channel unit preference. *Margaritifera* exhibited a unimodal distribution with the highest density at river km 35 and secondary peaks at river km 18 and 49 (Fig. 2). *Margaritifera* density also peaked in narrow valley segments (P = 0.06) (Fig. 3), with elevated densities found in riffles and runs compared to pools and glides (P < 0.009) (Fig. 4). Both *Anodonta* and *Gonidea* density peaked in the downstream portion of the study area; however, *Anodonta* were ubiquitously distributed, while *Gonidea* density was an order of magnitude lower than the other two genera and its distribution was constrained to below river km 21. *Anodonta* showed no statistically significant trends with valley confinement (P = 0.3) and insufficient data was available to test whether *Gonidea* density differed with regards to valley
Figure 2. Longitudinal patterns of reach scale mean (± 1 within reach standard deviation) mussel density for (a) *Margaritifera*, (b) *Anodonta*, and (c) *Gonidea* throughout the 55-km study area. Black circles denote reaches located in wide valley segments and white circles denote reaches in narrow segments. Note the different y-axis scales among genera.
Figure 3. Box plots of (a) *Margaritifera*, (b) *Anodonta*, and (c) *Gonidea* density within narrow (white) and wide (dark grey) valley confinement. Y-axis is the log of mussel density plus the smallest non-zero value and circles indicate outliers beyond the first and third quartiles. Note the different y-axis scales among genera.
Figure 4. Box plots of (a) *Margaritifera*, (b) *Anodonta*, and (c) *Gonidea* density within glides, pools, riffles, and runs. Y-axis is the log of mussel density plus the smallest non-zero value and circles indicate outliers beyond the first and third quartiles. Note the different y-axis scales among genera.
confinement. *Gonidea* exhibited a preference for glides (*P* < 0.03), while *Anodonta* showed a preference for both glides and runs (*P* < 0.05).

*Spatial patterns in biotic and abiotic habitat variables*

Several habitat variables exhibited large scale, longitudinal trends indicating a consistent shift in habitat from the headwaters to the downstream portion of the study area. Above river km 45, bankfull width to depth ratios were among the lowest observed (Fig. 5), pools comprised the majority of reach lengths (Fig. 6), and sinuosity, LWD, and percent silt cover exhibited their highest levels. In contrast, as bankfull width to depth ratios increased downstream, a more diverse array of channel unit types were observed, substratum coarsened, the prevalence of emergent vegetation increased, and both LWD and sinuosity decreased.

Several biotic and water quality habitat predictors exhibited similar large scale, longitudinal trends. Specific conductance exhibited minor variability over a low range of values and systematically increased downstream of river km 45 (Fig. 7). The relative abundance of coldwater fishes peaked at a trough in water temperature near river km 35 and water temperature generally increased in a downstream direction. In contrast, coolwater fishes exhibited a nearly opposite pattern with the lowest densities near river km 35 and higher densities both upstream and downstream of this location. Several variables (e.g. water depth, EPA score, substratum diversity, and pH) exhibited low variability and showed no clear longitudinal trends (data not shown).

In addition to these longitudinal trends, several habitat variables covaried with the large scale, geomorphic structure of the MFJDR and followed the alternating pattern of valley confinement between wide and narrow segments. In reaches located in wide valley segments, channel gradient was generally low, pools were the dominant channel unit type comprising 43% of average reach length, and LWD was rare. In contrast, reaches in narrow valley segments were
Figure 5. Longitudinal variation of physical habitat variables (reach average) (a) bankfull width to depth ratio, (b) sinuosity, (c) LWD density, (d) percent silt, (e) percent emergent vegetation, (f) $D_{50}$ particle size, (g) percent slope, and (h) Pfankuch bed stability throughout the study area. Black circles denote wide valley segments and white circles denote narrow valley segments.
Figure 6. Longitudinal variation in the spatial extent percent linear extent of (a) pool, (b) glide, (c) riffle, and (d) run channel units within individual reaches. Black circles denote wide valley segments and white circles denote narrow valley segments.
characterized by steeper gradients, riffles as the dominant channel unit type, and higher densities of both LWD and percent emergent aquatic vegetation, especially above river km 30. Despite higher slope and associated shear in narrow valley segments, Pfankuch bed stability was higher (low values) in reaches located in narrow valley segments. Seston quantity and quality, as measured by the autotrophic index, chlorophyll $a$, and AFDM, was highly variable throughout the study area, showed no strong longitudinal trend, and was consistently a poor predictor for all genera.
Mussel habitat associations

At the channel unit scale, *Anodonta* and *Margaritifera* models performed similarly with 48% and 43% variance explained, respectively, while the *Gonidea* model had the poorest performance with 25% variance explained. River km was consistently among the most important variables at the channel unit scale (Table 2) along with water depth, bankfull width to depth ratio, substratum size, bed stability, and emergent vegetation. Densities of all genera increased among channel units containing a greater percentage of emergent vegetation and decreased as percent silt increased or channel stability decreased (*Margaritifera* and *Gonidea* only) (Fig. 8). While *Anodonta* and *Gonidea* were both found in higher densities within channel units greater than one meter deep, *Margaritifera* exhibited higher densities in wide and shallow channel units, as measured by the ratio of bankfull width to depth.

The model with the greatest percent variance explained occurred with the reach scale *Margaritifera* model (80% variance explained), while the *Anodonta* (32%) and *Gonidea* (44%) models exhibited moderate performance. While the reach scale models included all four categories of predictor variables, the same types of hydrogeomorphic habitat characteristics were generally important to explaining mussel density at both spatial scales. Substratum size remained important at the reach scale and, for all three genera, mussel density increased with greater substratum size, as indicated in the partial dependence plots (Fig. 9). However, the relationship of substratum size with *Anodonta* and *Gonidea* was weak. The *Margaritifera* model included only two variables, LWD and $D_{16}$, both of which were positively related to mussel density. Similar to the channel unit scale model, *Anodonta* density was positively related to deep reaches with greater emergent vegetation, although maximum depth was replaced by percent pool as a top predictor at this scale. Both the *Anodonta* and *Gonidea* models included a positive relationship with specific conductance, and the *Gonidea* model was almost entirely driven by this single variable.
Table 2. Variables included in the best channel unit (top) and reach (bottom) scale models for all three genera. The percent variance explained with the addition of river km at the channel unit scale to account for spatial autocorrelation is shown. Variable relationships are summarized as positive (+), negative (-), and flat (f). Variables for each model are listed in their order of importance, as indicated by the percent increase in mean squared error (MSE) ± the standard deviation (SD). Variable name abbreviations are explained in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Percent variance explained (%)</th>
<th>Percent variance with river km (%)</th>
<th>Variable</th>
<th>Relationship</th>
<th>Percent increase in MSE (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel Unit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Margaritifera</em></td>
<td>27</td>
<td>43</td>
<td>(River km)</td>
<td>+</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Silt</td>
<td>-</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pfankuch</td>
<td>-</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BkflWtoD</td>
<td>+</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EmergVeg</td>
<td>+</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td><em>Anodonta</em></td>
<td>35</td>
<td>48</td>
<td>MaxDepth</td>
<td>+</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(River km)</td>
<td>-</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EmergVeg</td>
<td>+</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td><em>Gonidea</em></td>
<td>7</td>
<td>25</td>
<td>(River km)</td>
<td>-</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ds0</td>
<td>-</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Silt</td>
<td>-</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MaxDepth</td>
<td>+</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EmergVeg</td>
<td>+</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Reach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Margaritifera</em></td>
<td>80</td>
<td>n/a</td>
<td>LWD</td>
<td>+</td>
<td>61</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D16</td>
<td>+</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td><em>Anodonta</em></td>
<td>32</td>
<td>n/a</td>
<td>%Pool</td>
<td>+</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EmergVeg</td>
<td>+</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ds0</td>
<td>+/f</td>
<td>8</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Silt</td>
<td>-</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conduct</td>
<td>+</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><em>Gonidea</em></td>
<td>44</td>
<td>n/a</td>
<td>Conduct</td>
<td>+</td>
<td>77</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D16</td>
<td>+/f</td>
<td>30</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 8. Partial dependence plots of the variables in the final channel unit scale random forest models for each of the three genera. Plots indicate mussel density based on each predictor variable in the best models after averaging out the effects of all other predictor variables in the model. Rug plots indicate deciles of data for each predictor variable. Variable importance for river km not shown and relative importance of variables for each model decreases from left to right as reported in Table 2.
Figure 9. Partial dependence plots of the variables in the final reach scale random forest models for each of the three genera. Plots indicate mussel density based on each predictor variable in the best models after averaging out the effects of all other predictor variables in the model. Rug plots indicate deciles of data for each predictor variable. Relative importance of variables for each model decreases from left to right.
Comparative variable importance plots of all predictor variables at the reach scale (Fig. 10) indicated additional patterns that were not present in the best models. In particular, while channel unit type and confinement were rarely present in the best models, these habitat characteristics ranked high in the comparative models of all three genera. For example, the ratio of valley width to bankfull width (i.e. confinement) was the last variable removed during variable selection for the reach scale *Margaritifera* model. For channel unit type, percent riffle was the highest ranked for *Margaritifera*, percent pool and percent riffle \((r = 0.64, P < 0.0001)\) were the highest for *Anodonta*, and percent glide was the highest for *Gonidea*. 
Figure 10. Variable importance plots from the entire set of variables used to construct best models at the (a) channel unit and (b) reach scale.
The hierarchical arrangement of habitat patches within lotic systems drives the distributional patterns of biota across a riverscape (Fausch et al., 2002) and can result in scale-dependent habitat relationships (Poff, 1997). Previous multi-scale habitat research with stream fishes has indicated that the large scale, geomorphic structure of valley confinement influences smaller scale habitat features such as pools and spawning gravels (Benda et al., 1992). For freshwater mussels, researchers have successfully used a multi-scale approach to assessing habitat needs, which has clarified the variability of habitat controls from the reach to the sub-meter scale (Howard & Cuffey, 2003). I used the hierarchical structuring of habitat in conjunction with functional habitat features to investigate the variability in western freshwater mussel density to a myriad of habitat predictors including hydrogeomorphology, host fishes, food quality and quantity, and water quality parameters at multiple spatial scales. Through adopting this approach, I was able to identify scale specific habitat relationships within and among genera that have direct implications for the management and restoration of western mussel populations.

**Mussel habitat relationships**

Spatial patterns in mussel density were associated with habitat heterogeneity both within and among reaches. At the sub-watershed scale, mussel distributional patterns were consistent with results from other western studies (Howard & Cuffey, 2003; Howard, 2005; Brim Box et al., 2006) in that *Margaritifera* exhibited higher densities near the headwaters, while *Anodonta* and *Gonidea* densities peaked in downstream reaches. *Margaritifera* appeared particularly responsive to large scale geomorphic gradients with significantly greater densities in narrow valley segments. This large scale relationship with confinement is likely a function of the finer scale habitat conditions created by narrow valley segments, which tend to be higher in slope, dominated by faster channel units of riffles and runs, and contain higher LWD density. The
observed relationship to confinement is similar to the findings of Baxter & Hauer (2000), where higher bull trout redd density was found in low gradient sections of confined river segments where upwelling was common, but locally where downwelling occurred. Although valley confinement was not present in any of the final models, the ratio of valley width to bankfull width was the last variable to be removed from the reach scale *Margaritifera* model. This result highlights the importance of the large scale geomorphic structure of the MFJDR in shaping finer scale habitat structure, which *Margaritifera* appear to respond.

At finer spatial scales, *Margaritifera* densities were elevated in reaches containing higher densities of LWD and coarser substratum. LWD has been shown to positively influence mussel density by providing flow refuge to parts of the channel that might otherwise experience scour during high flow events (Palmer *et al.*, 1996; Jones & Byrne, 2010). LWD can also promote habitat heterogeneity (Frissell *et al.*, 1986; Hilderbrand *et al.*, 1997) at the reach scale by increasing channel unit diversity. At smaller scales, LWD creates a more heterogeneous distribution of substratum size classes, which is connected with increases in biodiversity (Allan, 2004). Such reach scale changes to the habitat template have been associated with higher densities of juvenile salmonids (Roni & Quinn, 2001), which are the presumed host fish of western *Margaritifera*. Consequently, the mechanism by which LWD promotes local mussel recruitment is hard to disentangle and likely results from the interaction of substratum stability, habitat heterogeneity, and proximity to host fishes.

Conversely, *Anodonta* and *Gonidea* did not exhibit differences in density between wide and narrow valley segments, although smaller scale preferences for channel unit types were found. Rather, both *Anodonta* and *Gonidea* density was positively related to specific conductance values above 0.14 mS/cm, indicating that these two genera are responding negatively to the low ionic concentrations observed higher in the watershed. Johnson & Brown (2000) found similar patterns where conductivity values below 0.025 mS/cm limited mussel distribution due to
insufficient amounts of calcium for shell formation. In contrast, very high values (0.8 mS/cm) of conductivity have been shown to limit mussel distribution due to pollution (McRae et al., 2004), although none of the values reported in this study were high enough to be considered detrimental to mussels. Conductivity exhibited a strong longitudinal pattern of increasing value downstream that mirrors the increase in *Anodonta* and *Gonidea* density in the downstream portion of the study area. This pattern may be partially explained by the more dilute urine produced by *Margaritifera* compared to *Anodonta*, indicating superior ion regulation, as seen in Europe mussel species (Dietz et al., 1996; Evans, 2009). If this relationship persists with western genera, the low densities of *Gonidea* and *Anodonta* near the headwaters may be due to their weak osmoregulatory abilities, which, in contrast, allow *Margaritifera* to thrive in these low ion waters.

Despite large scale distributional patterns, mussel density was also highly variable within individual reaches, indicating that mussels were responding to habitat heterogeneity at both large and small spatial scales (Palmer et al., 2000). Similar to the reach scale, mussels were most sensitive to hydrogeomorphic variables at the channel unit scale. However, this variation did not appear related to the specific channel unit type in which mussels were found, but rather substratum characteristics (e.g., fine sediment levels, substrate stability, and emergent vegetation) and, to a lesser extent, channel dimensions, which are habitat features that are known to be highly variable at small spatial scales (Salmon & Green, 1983; Layzer & Madison, 1995; Vaughn & Taylor, 1999; Howard & Cuffey, 2003).

Although mussels appeared to be responding to the hierarchical arrangement of habitat patches within the MFJD, the nature of mussel habitat associations did not appear scale dependent in all instances. Positive relationships with substratum size were present in four of the six models, including all models at the reach scale. All genera in this study were positively associated with larger substratum sizes, particularly coarse gravel and small cobbles, and locations with less silt. This result is consistent with other studies indicating that mussels appear
to have a preference for locations where substratum size is large enough to remain stable at high flows, but velocity is fast enough to prevent excessive siltation (Vannote & Minshall, 1982; Salmon & Green, 1983; Vaughn, 1997; Vaughn & Taylor, 1999; Howard & Cuffey, 2003; McRae et al., 2004). In addition, locations dominated by larger-sized particles provide greater interstitial spaces that can benefit juvenile feeding activity (Yeager, Cherry & Neves, 1994). While previous investigations have found that the role of substratum size has weak or no relationships with mussel occurrence (Strayer & Ralley, 1993; Layzer & Madison, 1995; Strayer, 1999), viewing substratum size at large spatial scales may explain the better performance of this variable within my models.

The relationship between mussel density and bed stability was further indicated by the positive relationship with locations having less silt and more emergent vegetation. The high mobility of silt has been found to create unsuitable habitat for mussels (Layzer & Madison, 1995; Morales et al., 2006). At the channel unit scale, Margaritifera and Gonidea were associated with channel units having less silt while Anodonta showed the same relationship at the reach scale. Silt can inhibit the growth of mussels by reducing oxygen absorption through clogging gills and blocking photosynthesis, which can reduce food availability (Brim Box & Mossa, 1999; Poole & Downing, 2004). Similarly, the positive relationship with percent emergent vegetation further confirms the finding that mussels prefer more stable locations, since emergent vegetation has been shown to stabilize banks and nearby substratum (Levine, 2000; Howard & Cuffey, 2003). While many surrogates for stability were included in the models, my calculations of shear stress and RBS were not included in any of the best models. Stability, as measured by hydraulic variables, has been shown to be important in structuring mussel habitat (Howard & Cuffey, 2003; Gangloff & Feminella, 2007; Allen & Vaughn, 2010), such that mussels tend to be found in locations with lower hydraulic forces at high flows. However, it is possible that the reach scale
calculations were too coarse of a measurement, and mussels in this system are influenced by shear stress and similar hydraulic forces at finer spatial scales.

*Unexpected results*

A number of habitat variables that were predicted to be important were absent from the best models. Although percent channel unit type ranked relatively high on the comparative variable importance plots of each genus, only percent pool was retained in the best model of *Anodonta* at the reach scale. Contrary to previous studies in the MFJDR, which occurred in the middle section of the watershed, where the greatest density of mussels were found in pools (Howard & Cuffey, 2003; Howard, 2005), I found channel unit preferences specific to each genus. For example, *Margaritifera* were found more often in higher velocity, shallower channel units of riffles and runs and exhibited a similar preference for channel units with a greater bankfull width to depth ratio. This is consistent with the findings in other systems where *Margaritifera* have been found to be poor vertical migrators and, as a result, may be avoiding the high depositional rates in slower channel units (Vannote & Minshall, 1982; Johnson & Brown, 2000). Channel morphology trends were also present with *Anodonta* and *Gonidea*, where they both exhibited a preference for deeper channel units and reaches, indicating that these two genera may be exploiting the smaller scale, high flow refugia within pools as described by Howard & Cuffey (2003). Overall, I concluded that the conditions creating bed stability within channel units of all types was more important than the actual channel unit classifications.

Most conspicuously lacking from the reach scale models were host fish presence and food quality and quantity, which were the hypothesized causes for the longitudinal spatial structuring of mussels. Host fish requirements for western mussels are still being determined, but *Anodonta* are considered host fish generalists that utilize a wide range of coolwater fishes such as cyprinids (Mock *et al.*, 2004; Brim Box *et al.*, 2006; O'Brien, 2012). In contrast, *Margaritifera*
are considered host fish specialists that require coldwater salmonids to complete their life cycle (Murphy, 1942; Karna & Millemann, 1978). Interestingly, the peak in coldwater fish abundance at river km 35 corresponds with the peak in *Margaritifera* density, and while this variable was not retained in the best model, coldwater fish abundance ranked relatively high in the comparative variable importance plot for *Margaritifera*. The lack of host fish abundance in the models may be explained by the fact that both coldwater and coolwater fish assemblages were ubiquitously distributed throughout the study area and these broadly dispersed fish populations, including highly mobile juvenile salmonid populations (Hartman & Brown, 1987; Kahler, Roni & Quinn, 2001) appear adequate to maintain existing mussel populations. By comparison, the loss of host fishes in the nearby Umatilla River watershed have been implicated in the local extirpation of *Margaritifera* (Brim Box *et al.*, 2006). The host fish abundance that was included in my models was a single temporal measurement summarized to the reach scale, which may have been insufficient at capturing the complex interaction between mussel reproduction timing, host fish movement patterns, and habitat use. Future work should include this type of detailed data to clarify this complex relationship.

As suggested by previous research (Brim Box *et al.*, 2006), I hypothesized that food resource availability might explain differences in the spatial distribution of *Anodonta*, *Gonidea*, and *Margaritifera* because of differential metabolic rates among genera (Bauer *et al.*, 1991). However, no measure of seston quality and quantity was present in any of the best models and these variables generally ranked low in the comparative variable importance plots. The single measurement of seston in autumn may have played a role in the poor performance of this variable since mussel filtration rates can vary with temperature (Aldridge, Payne & Miller, 1995) and seasonality (Howard & Cuffey, 2006). As a result of these factors, mussel growth fluctuates within a given year, and climatic conditions can influence growth and subsequent food demands between years (Schöne *et al.*, 2007), so a more complete picture of temporal variation in food
availability could help clarify the role of this variable. In addition to filtering water for food particles, mussels may feed on organic particles present within the sediment (Yeager et al., 1994; Nichols et al., 2005) and selectively ingest algae and bacteria (Nichols & Garling, 2000), so our seston measurement may have not captured the full range of potential food resources.
CONCLUSIONS

This study provides one of the few quantitative analyses of freshwater mussel habitat requirements in the western United States. The explanatory models provide important information regarding genus-specific habitat associations across multiple spatial scales, which can be used to identify locations of suitable habitat for restoration and conservation in nearby watersheds by using a hierarchical approach. Specifically, this information can be used to systematically guide the selection of river segments, reaches, and subsequent channel units where mussels will be translocated as part of restoration efforts. While the results from this study may be used to broadly guide restoration and management of western mussels, validation with an external data set should be conducted before using model results in a predictive manner. The data that was collected as part of this research also provides quantitative information regarding mussel distribution and density and associated biotic and abiotic habitat data that can be used as baseline information to guide monitoring and management of western freshwater mussels in the MFJDR. While this study investigated mussel density patterns at scales ranging from the watershed to the channel unit, mussels are known to also respond to habitat at the sub-meter scale (Layzer & Madison, 1995; Hastie et al., 2000; Howard & Cuffey, 2003). Consequently, the multi-scale models used in this study may have benefited from inclusion of smaller, sub-meter scale habitat data to investigate the full range of habitat scales capable of influencing western freshwater mussels. For example, the hydraulic predictors of bankfull shear stress and relative bed stability were consistently poor predictors, despite their high predictive capability in other studies when quantified at the channel unit and sub-channel unit scales (Howard & Cuffey, 2003; Gangloff & Feminella, 2007). My use of reach scale average hydraulic forces likely resulted in the poor predictive capability of these variables because of the coarse scale of these measurements relative to the fine scale variability commonly exhibited by river hydraulics. Ideally, sub-channel unit or
sub-meter scale habitat and mussel data would be included in future studies by nesting data collected at this spatial scale within individual channel units, reaches, and sub-watersheds to understand the hierarchical effects of habitat at a wider range of spatial scales.

Model performance for all three genera may have been limited by several additional factors other than the scale at which data was collected. For example, the study area was located at the upper edge of *Gonidea*’s distributional range, which may have restricted my ability to model this genus. In addition, sampling error associated with visual snorkel surveys differed by genus and could have influenced model results. Lastly, there are numerous legacy effects in the MFJDR from historical and current human activities including dredge mining, logging, grazing, and road construction (McDowell, 2000) that may have had a strong influence on mussel populations, and these parameters may not have been captured in the rapid habitat bioassessment.

Successful conservation and restoration will depend upon further study to clarify the causes of the large scale longitudinal trends of western freshwater mussels. Principal among these are the importance of conductivity and differential ion regulation among genera, as well as metabolic differences and food requirements of western freshwater mussels. Understanding how these variables might influence the density and distribution of western freshwater mussels would be best accomplished through coupling observational field studies with laboratory experiments.

In conclusion, I developed multi-scale habitat models that incorporated a wide range of chemical, physical, and biological predictors to inform the conservation and restoration of freshwater mussels. The main goal of this work was to identify relationships between the variation in mussel density and the factors that might be associated with these patterns. This approach was successfully used to describe genus-specific habitat requirements from the sub-watershed, the reach, and down to the channel unit scale. By continuing to build upon the results of this research, we will improve our ability to manage and restore western freshwater mussels.


