Responses of Macroinvertebrate Assemblages to Flow Variations in the Qinghai-Tibet Plateau: Establishment and Application of a Multi-metric Habitat Suitability Model

Zhou Xiongdong¹, Xu Mengzhen², Lei Fakai¹, Wang Zhaoyin², and Luo Yuyan¹

¹State key laboratory of hydroscience and engineering
²Tsinghua University

November 15, 2022

Abstract

Anthropogenic impacts and climate change modify instream flow, altering ecosystem services and impacting on aquatic ecosystems. Alpine rivers and streams on the Qinghai-Tibet Plateau (QTP), are especially vulnerable to disturbance due to a limited taxonomic complexity. The effects of variations in flow have been studied using specific taxa, however, the flow-biota relationships of assemblages are poorly understood. A multi-metric habitat suitability model (MM-HSM) was developed, using biological integrity measures of macroinvertebrate assemblages to substitute for habitat suitability indices (HSI) derived from individual taxa. The MM-HSM was trained using macroinvertebrate data from three representative alpine rivers (the Yarlung Tsangpo, the Nujiang, and the Bai Rivers) on the QTP, and was verified using data from the Lanmucuo River. The model produced reliable predictions using the training dataset (R² = 0.587) and the verification dataset (R² = 0.489), and was robust to inter-basin differences and changes in dataset size. By coupling the MM-HSM with hydrodynamic simulations, the relationship between weighted usable area (WUA) and flow variations (0.11–1.99 m³/s) for macroinvertebrates was established, and a unimodal response pattern (optimal flow Q = 1.21 m³/s) was observed for macroinvertebrate assemblages from the Lanmucuo River. This was in contrast to the skewed unimodal or monotonically increasing relationships observed for individual indicator taxa, supporting our hypothesis that biological integrity varies with changing flow and conforms to the intermediate disturbance hypothesis. The MM-HSM provides a novel framework to quantify species-environment relationships, which may be used for integrated river basin management.

Hosted file


Hosted file

X. D. Zhou†, M. Z. Xu††, F. K. Lei†, Z. Y. Wang†, and Y. Y. Luo†

†State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing, China

Corresponding author: M. Z. Xu (mzxu@mail.tsinghua.edu.cn)

†The first two authors contributed equally to this paper.

Key Points:
- The multi-metric habitat suitability model uses biological integrity measures to substitute for indices derived from individual taxa
- The response pattern for macroinvertebrate assemblages to flow variations is dramatically different from those for individual taxa
- An flow discharge around 1.21 m$^3$·s$^{-1}$ is the most beneficial to macroinvertebrate integrity in the Lanmucuo river in summer

Abstract

Anthropogenic impacts and climate change modify instream flow, altering ecosystem services and impacting on aquatic ecosystems. Alpine rivers and streams on the Qinghai-Tibet Plateau (QTP), are especially vulnerable to disturbance due to a limited taxonomic complexity. The effects of variations in flow have been studied using specific taxa, however, the flow-biota relationships of assemblages are poorly understood. A multi-metric habitat suitability model (MM-HSM) was developed, using biological integrity measures of macroinvertebrate assemblages to substitute for habitat suitability indices (HSI) derived from individual taxa. The MM-HSM was trained using macroinvertebrate data from three representative alpine rivers (the Yarlung Tsangpo, the Nujiang, and the Bai Rivers) on the QTP, and was verified using data from the Lanmucuo River. The model produced reliable predictions using the training dataset ($R^2 = 0.587$) and the verification dataset ($R^2 = 0.489$), and was robust to inter-basin differences and changes in dataset size. By coupling the MM-HSM with hydrodynamic simulations, the relationship between weighted usable area (WUA) and flow variations (0.11–1.99 m$^3$·s$^{-1}$) for macroinvertebrates was established, and a unimodal response pattern (optimal flow $Q = 1.21$ m$^3$·s$^{-1}$) was observed for macroinvertebrate assemblages from the Lanmucuo River. This was in contrast to the skewed unimodal or monotonically increasing relationships observed for individual indicator taxa, supporting our hypothesis that biological integrity varies with changing flow and conforms to the intermediate disturbance hypothesis. The MM-HSM provides a novel framework to quantify species-environment relationships, which may be used for integrated river basin management.

1 Introduction

The Qinghai-Tibet Plateau (QTP), also known as the Third Pole Region, is the largest high-altitude land mass on earth. The region holds the most freshwa-
ter, mainly in the form of glaciers, outside the polar ice caps (Morton, 2011). The QTP is of critical ecological importance, possessing diverse biomes (alpine tundra, montane forest, subtropical dry forest and rainforest; Ni & Herzschuh, 2011), and encompassing several biodiversity hotspots (the Central Asia mountain area, the Himalayas, the Indo-Burma, and the Hengduan mountain area) (Tang et al., 2006). Due to the unique conditions, river and stream biota in the QTP are highly specialized, and are therefore vulnerable to external disturbance (Yao et al., 2012; Favre et al., 2015; Xu et al., 2018). Over the past decades, environmental conditions, especially instream flow, have been intensively modified by anthropogenic activities and climate warming, increasing the risk to ecosystem services in alpine rivers and streams (Poff et al., 1997; Wenger et al., 2011).

As taxa have evolved over long time periods under natural fluctuations in flow, the impacts of flow variations on organisms may be limited (Schneider et al., 2013). However, it is still expected that the morphological, functional and trophic structures of biotic assemblages may be altered if variations in flow exceed a certain threshold and become ‘extreme’ (Gibbins et al., 2007; Bêche et al., 2009). These impacts have been studied in a wide range of biota, from microorganisms to vertebrates, but most studies are on macroinvertebrates (e.g., Lancaster et al., 1990; Rader & Belish, 1999; Suren & Jowett, 2006; Dewson et al., 2007) and fishes (e.g., Bradford & Heinonen, 2008; Murchie et al., 2008; Irvine et al., 2009). Flow variations can indirectly affect biotic assemblages through altering environmental conditions, including hydrodynamics (e.g., stream power, water level, near-bed shear stress, temporal flow fluctuation, and turbulence; Hart & Finelli, 1999; Jowett, 2003; Wedderburn et al., 2012; Blankaert et al., 2013; Zhou et al., 2017), sediment movement processes (e.g., substrate composition, suspended load, channel bed stability; Boubée et al., 1997; Duan et al., 2009; Zhao et al., 2015), physicochemical environmental variables (e.g., water temperature, dissolved oxygen, salinity; Korven & Wilcox, 1964; Wilcock et al., 1998; Lowney, 2000), and the transportation of autochthonous and allochthonous substances (e.g., organic detritus, nutrients, pollutants; Mathooko et al., 2001; Maazouzi et al., 2013; Xu et al., 2014). Notably, responses of biota to variations in flow will vary depending on taxonomic position, and on life stage, which is especially distinct in migratory fishes such as salmonids (Bjornn & Reiser, 1991; Hayes et al., 2019).

Ecologists have sought to quantify biotic responses to flow variations by combining physical habitat mapping and habitat suitability analysis. Instream Flow Incremental Methodology (IFIM) is one of the most prevalent methods for establishing instream flow criteria, and this method has been successfully applied to evaluate fish habitat and for environmental flow management (e.g., Brooks, 1997; Vadas & Orth, 2001). The IFIM established quantitative relationships between habitat suitability index (HSI) and environmental variables, using a range of analytical techniques including empirical suitability curves (Raleigh et al., 1986), the fuzzy logic model (Van Broekhoven et al., 2006), statistical regressions (e.g. generalized additive models; Costa et al., 2012), machine learning
(e.g., artificial neural networks; Park et al., 2003), and the maximum entropy model (Ashford et al., 2014). Despite the methodological refinement of HSI quantification, IFIM is mainly focused on a specific species or a narrowly constrained taxonomic group (usually fishes) for specific economic or conservation targets (e.g., Lamouroux et al., 1998; Monton et al., 2007; Yao & Rutschmann, 2015). Further research is required to establish an integrated HSI for broader taxonomic groups such as macroinvertebrate assemblages. Most species within an assemblage are equally important to performing ecological services (Dee et al., 2019), and therefore the ‘best way’ to assess an ecosystem is to evaluate the degree to which it supports and maintains balanced, integrated and adaptive biotic assemblages (Karr, 1981). The use of a biological integrity index, instead of a single species, has good potential as an ecological indicator, since it describes ecosystem health from a multi-dimensional perspective taking into account biodiversity, and the morphological and functional structure of assemblages (Karr & Dudley, 1981).

Quantitative flow-biota relationships are generally focused on a single species rather than a broader biotic assemblage. Further research is needed to understand how river or stream assemblages respond to flow variations. We have developed a multi-metric habitat suitability model (MM-HSM) using measures of biological integrity including biodiversity, morphological structure, functional structure, and tolerance to pollution, as well as trophic characteristics. The model was trained and verified using data describing macroinvertebrates collected from four representative rivers in the QTP. Further, we coupled this MM-HSM with a 2D hydrodynamic simulation to investigate the response pattern of biotic assemblages to instream flow variations, and to test our hypothesis that an intermediate flow discharge would allow the optimal biological integrity of a flowing-water ecosystem (Connell, 1978). The MM-HSM will help to unravel a predictable pattern of biotic response to flow variations, and determining optimal flow for ‘biological integrity’ will benefit the management of river basins.

2 Materials and Methods

2.1 Field investigation and macroinvertebrate sampling

The ecology and geomorphology of four QTP rivers (the Yarlung Tsangpo River in the south, the Nujiang River in the southeast, and the Bai and the Lanmucuo Rivers in the northeast) were examined during 2014–2018 (Figure 1a). These rivers represent typical alpine river types: fast-flowing, deeply incised, cobble- and boulder-bed rivers (the Yarlung Tsangpo and the Nujiang Rivers; Zhou et al., 2017; Zhou et al., 2019a), and slow-flowing, silt- and gravel-bed rivers (the Bai River; Zhou et al., 2019b). Macroinvertebrates were sampled from 68 sampling sites distributed across the Yarlung Tsangpo, the Nujiang and the Bai Rivers for MM-HSM training, and from 25 sampling sites on the middle reach of the Lanmucuo River for MM-HSM verification. All sampling sites were in rural areas and free from anthropogenic disturbance. Bathymetry and hydrodynamics were measured in detail on the Lanmucuo River, for further hydrodynamic simulation (Figures 1b–1d).
Figure 1  (a) Distribution of the sampling sites on the Yarlung Tsangpo, the Nujiang, the Bai and the Lanmucuo Rivers on the QTP. (b) The Lanmucuo River Basin. (c) Distribution of the nine sections used to measure hydrodynamics, along which the 25 sampling sites were located. (d) The Lanmucuo River is a meandering river with mild flow, gravel- and cobble-substrate, and riffle-pool bed structures (Fig. 1b). The photo was taken on 19 July 2014, showing a typical cross section (MB09) of the Lanmucuo River.

Three replicate macroinvertebrate samples were collected from each sampling site (1 m$^2$) using a kick-net (mesh size: 420 μm). Specimens were sorted and preserved in 75% ethanol in the field and identified to genus level using an optical microscope in the laboratory. Forty-one macroinvertebrate indices were calculated (Table S1, columns 1–2) to measure biological integrity including indices describing: biodiversity (eight indices), morphological structure (16 indices), functional structure (nine indices), tolerance to pollution (five indices) and trophic structure (three indices). Indices describing trophic structure of macroinvertebrate assemblages are not typically used in traditional Biological Integrity Indexing Systems (e.g., Hilsenhoff, 1988; Kerans & Karr, 1994). However, the capacity of food-web indices to reflect the trophic characteristics of ecosystems have been proven (e.g., O’Gorman et al., 2019), and these indices were applied in this study. The response of each index to increasing disturbance was compared with previous studies and follows a monotonic pattern (Table S1, columns 3–4). Before further analysis in the MM-HSM, all indices were normalized, and their response patterns were mathematically reoriented to show a consistent decrease as disturbance levels increase (Table S1, column 5).

A physical habitat survey measuring hydrodynamic and substrate conditions
(i.e., the environmental variables: flow velocity \( v \), water depth \( h \), and median particle diameter \( D_{50} \)) was conducted at each sampling site. Flow velocity, water depth, and cross-sectional discharge were measured using a portable acoustic Doppler velocimeter (FlowTracker 2, SonTek, USA). For sites with a coarse substrate, we took photographs of boulders and large cobbles and analyzed grain size distribution (GSD) graphically using BASEGRAIN (Detert & Weitbrecht, 2013). At sites with a fine sediment substrate, a grab was used to collect a sample of the substrate, and this was then sieved in the laboratory to determine GSD and median particle diameter (\( D_{50} \)). Flow variations in the Lanmucuo River were estimated using data from two hydrometric stations (the Jungong Station which is downstream, and the Maqu Station which is upstream; Figure S1a). Summer flow in the study reach varied from 0.11–1.99 m\(^3\) · s\(^{-1}\) based on average daily flow for a 14-year period (Figure S1b).

2.2 MM-HSM training, verification and further evaluation

The model framework was constructed based on the MM-HSM (Figure 2a) and hydrodynamic simulations (Figure 2b). The MM-HSM produced MM-HSI values (0–1, reflecting biological integrity) as a function of velocity \( v \), water depth \( h \) and median particle diameter \( D_{50} \), and the hydrodynamic simulation yielded predicted values of \( v \), \( h \), and \( D_{50} \) for a given flow discharge (\( Q \)). MM-HSI values were derived by coupling the results of the MM-HSM and the hydrodynamic simulations. Predicted weighted usable area (WUA) values generated in response to different flow conditions were obtained by implementing the framework under different flow discharge (\( Q \)) scenarios.

![Figure 2 Model framework for this study. (a) The multi-metric habitat suitability model (MM-HSM), (b) the hydrodynamic simulation model, and (c) coupling of the two models to generate weighted usable area (WUA) as a function of flow discharge (\( Q \)).](image-url)
The MM-HSM was trained using data from 68 sampling sites in the Yarlung Tsangpo, the Nujiang, and the Bai Rivers, and was verified using data from 25 sampling sites in the Lanmucuo River. Forty-one biological integrity indices were calculated (after normalization and reorientation) using the macroinvertebrate data. Pairwise Spearman correlation analyses were performed to investigate the consistency of responses to increasing disturbance. Indices showing negative relationships ($R < -0.3$) were omitted from further procedures. The MM-HSI for macroinvertebrates was then synthesized as the weighted average of the filtered biological integrity indices. As the indices were all normalized, the MM-HSI values ranged from 0 to 1. Weight for each index was determined using goodness of fit ($R^2$) to the physical environmental variables including $v$, $h$ and $D_{50}$, using generalized additive models (GAMs; R package “mgcv”, function “gam”; Wood, 2011). Model performances on the training/verification data were evaluated by comparing MM-HSI values calculated from biotic data (observed) with those estimated from corresponding environmental variables (predicted).

To further test the model, we simultaneously altered the proportion of data from the Bai River in the training dataset, to test for inter-basin differences, and altered the total size of the training dataset. The performance of the MM-HSM in response to the combined effects of both inter-basin differences and the size of the training dataset were analyzed using 1000 Monte-Carlo replications for each case.

Response of the macroinvertebrate assemblages to flow variations were quantified using the WUA calculated as:

$$WUA = \sum_{i=1}^{n} MM - HSI_i \times A_i \quad (1)$$

where $WUA$ is the weighted usable area for the study reach under a specific flow discharge, $n$ is the total number of sites in the hydrodynamic simulations on the study reach, $MM-HSI_i$ is the MM-HSI value for the $i$-th site, and $A_i$ is the area of the $i$-th site. WUA quantifies the spatial proportions in the study reach that can be preferentially utilized by macroinvertebrate assemblages, and a larger WUA represents a higher biological integrity of the ecosystem (Kelly et al., 2015).

We also developed three other habitat suitability models (HSMs) for specific indicator taxa. Indicator value analysis was performed using the Lanmucuo River macroinvertebrate data (R package “labdsv”, function “IndVal”; Roberts, 2019). Three indicator taxa were identified, including *Polypedilum* (Diptera: Chironomidae; indicator value = 0.79, $p < 0.001$) in shallow riffles, *Brachycentrus* (Trichoptera: Brachycentridae; indicator value = 0.58, $p < 0.001$) and *Limnophila* (Diptera: Limoniidae; indicator value = 0.54, $p < 0.001$) in deep pools. Normalized abundance data were used as HSI surrogates (labeled as $HSI_P$, $HSI_B$ and $HSI_L$) for these three taxa, and were linked with $v$, $h$ and $D_{50}$ using GAMs. Finally, we compared changes in WUAs for given variations in flow for macroinvertebrate assemblages (labeled as $WUA_{MM}$) with those for the indicator taxa (labeled as $WUA_P$, $WUA_B$, and $WUA_L$), to see whether the response of macroin-
vertebrate assemblages differed from those of specific taxa.

2.3 Hydrodynamic simulations

The hydrodynamic simulation was performed using a two-dimensional unstructured grid model, named T elemac2d. The simulation is based on the solution of the de Saint-Venant equations, including the continuity equation (Equation 2) and momentum equations (Equations 3a and 3b):

\[
\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + v \frac{\partial h}{\partial y} = 0 \tag{2}
\]

\[
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -g \frac{\partial \eta}{\partial x} + \frac{1}{h} \left( \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \right) - \frac{\tau_{bx}}{h} + f_{\text{cor}} \tag{3a}
\]

\[
\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -g \frac{\partial \eta}{\partial y} + \frac{1}{h} \left( \frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} \right) - \frac{\tau_{by}}{h} - f_{\text{cor}} u \tag{3b}
\]

where \(u\) and \(v\) are the velocity (m·s\(^{-1}\)) at the \(x\) and \(y\) direction, respectively; \(t\) is time (s); \(g\) is gravitational acceleration (m·s\(^{-2}\)); \(\rho\) is the density of water (kg·m\(^{-3}\)); \(h\) is water depth (m); \(\eta\) is the water surface elevation (m); \(f_{\text{cor}}\) is the Coriolis parameter (0 was chosen); \(\tau_{xx}, \tau_{xy}, \tau_{yx}, \tau_{yy}\) are the depth-integrated Reynolds stresses, which were calculated from the \(k-\varepsilon\) turbulence model; and \(\tau_{bx}\) and \(\tau_{by}\) are the shear stresses on the bed and flow interface, which were calculated using the Striker bottom friction law.

Riverbed deformation was calculated using the overall mass balance equation (Equation 4; Parker et al., 2000):

\[
(1 - p^{'}) \frac{\partial Z_b}{\partial t} + \frac{\partial Q_{bs}}{\partial x} + \frac{\partial Q_{bn}}{\partial y} = 0 \tag{4}
\]

where \(p^{'}\) is a parameter that depends on the porosity of the bed material (\(p^{'} = 0.05\) in this study); and \(Q_{bs}\) and \(Q_{bn}\) are the bed-load flux, which are calculated by the bed-load equation (Equation 5; Van Rijn, 1984a, 1984b, 1993):

\[
Q_b = 0.005 U h \left( \frac{U - U_{cr}}{|(S - 1)g D_{50}^{1.2}} \right)^{2.4} \left( \frac{D_{50}}{D} \right)^{1.2} \tag{5}
\]

where \(D_{50}\) is the median particle size (mm), and \(U_{cr}\) is the threshold current velocity, which is calculated by Equation 6:

\[
U_{cr} = 8.5 (D_{50})^{0.6} \log_{10} \left( \frac{4h}{D_{50}} \right) \tag{6}
\]

A numerical model of the river channel was meshed using a triangular grid with a size around 0.5 m. The total number of nodes and elements were 29050 and 55342, respectively. The elevation of each node was interpolated using the Blue Kenne (CHC-NRC, 2010) and measured topography data from 2018 (Figure S3). Topography information was generated using real-time kinematic positioning (RTK GPS) (Huaxing A16, HI Target, China) and an unmanned aerial vehicle (Hero 4, DJI, China). The sediments in the model were divided into nine representative groups with median particle diameters of 0.1, 0.5, 1, 2, 4, 8, 16, 32 and 64 mm (Figure S4a). The implicit finite volume method (FVM) was used to discretize the governing equations for flow and sediment transport.
with a curvilinear non-orthogonal grid. The turbulence model, bed deformation equation, and the sediment transport models were internally coupled with hydrodynamics. The convergence is guaranteed when the maximum residual error is less than $10^{-9}$. To accommodate to the size of the grids, the time step of this simulation was set at 0.25 s.

The initial and boundary conditions were set in both the inlet and outlet sections. The inlet was set by the flow rate versus time. At the outlet, the stage-discharge curve (Figure S4b) was set, and zero gradient outflow boundaries were adopted for the variables of velocity and turbulent kinetics. A solid wall boundary condition was applied to the side boundary. Seventeen flow discharge scenarios were implemented, including 0.11, 0.21, 0.30, 0.40, 0.52, 0.64, 0.73, 0.83, 1.04, 1.15, 1.21, 1.33, 1.45, 1.52, 1.65, 1.85 and 1.99 m$^3$·s$^{-1}$.

3 Results and Discussion

3.1 Macroinvertebrate characteristics of the study rivers

A total of 65820 macroinvertebrate specimens were collected and identified, belonging to 84 families and 235 genera. The result of a canonical correlation analysis (CCA) for 93 sampling sites constrained to $v$, $h$ and $D_{50}$ is shown in Figure 3. The first two axes explained a total of 80.43% of the variance, marking the critical contributions that hydrodynamic conditions and substrate type made to macroinvertebrate-environmental relationships on the QTP. Samples from different rivers tended to assemble in the CCA depending on river type and environmental conditions. Samples from the Yarlung Tsangpo and the Nujiang Rivers were relatively clustered. Both rivers are mountainous, deeply incised, and have large-sized substrate (0.4073 ± 0.4055 m, hereinafter mean ± sd) and fast flow (0.60 ± 0.46 m·s$^{-1}$). Samples from the meandering Bai and Lanmucuo Rivers were more dispersed, despite of a large overlap in their 95% envelopes. In addition, the 95% envelope extension around the Lanmucuo samples were surrounded by the three others, indicating that macroinvertebrate genera from the Lanmucuo were also found in the other three rivers.
Figure 3 Canonical correlation analysis (CCA) on macroinvertebrate samples in this study. Axes one and two explain 41.11% and 39.32% of the variance in macroinvertebrates, respectively. Dashed ellipses represent 95% envelope extensions of the datasets. Environmental variables: flow velocity ($v$), water depth ($h$) and median particle diameter of the substrate ($D_{50}$).

After initial analysis there were 26 macroinvertebrate indices used in the MM-HSI synthesis. This included five indices for biodiversity, eight for morphological structure, seven for functional structure, four for tolerance to pollution, and two for trophic structure (see Figure S2 for Spearman pre-filtering). Inter-basin value ranges, GAM goodness of fit and suggested weights are listed in Table 1. Indices reflecting different aspects of biological integrity differed markedly in weighting, ranging from 0.1% (relative abundance of the most dominant taxa) to 8.5% (relative abundance of Plecoptera).

Table 1 Average values (mean ± sd) of macroinvertebrate biological integrity indices for each of four different river basins on the QTP. Weighting of values used in the MM-HSI synthesis were determined using goodness of fit ($R^2$) to the physical environmental variables ($v$, $h$, and $D_{50}$) using GAMs. Abbreviations: YR for the Yarlung Tsangpo River, NR for the Nujiang River, BR for the Bai River, and LR for the Lanmucuo River. Bold superscript represents significant ($p < 0.05$) ANOVA and post hoc results, and italic superscript represents significant Kruskal-Wallis and post hoc test results.
<table>
<thead>
<tr>
<th>Macroinvertebrates Index</th>
<th>Yarlung Tsango</th>
<th>Nujiang</th>
<th>Bai</th>
<th>Lammucuo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plecoptera richness</td>
<td>1.4±2.1</td>
<td>1.2±2.0</td>
<td>0.4±0.5</td>
<td>0.5±0.8</td>
</tr>
<tr>
<td>Trichoptera richness</td>
<td>1.3±1.8</td>
<td>2.8±3.4</td>
<td>0.6±0.7</td>
<td>1.6±1.1</td>
</tr>
<tr>
<td>EPT(^1) richness</td>
<td>4.5±4.6</td>
<td>6.5±7.6</td>
<td>2.3±2.6</td>
<td>4.1±2.2</td>
</tr>
<tr>
<td>RA(^2) of 1st dominant taxa</td>
<td>0.48±0.18</td>
<td>0.52±0.22</td>
<td>0.55±0.2</td>
<td>0.40±0.13</td>
</tr>
<tr>
<td>RA of EPT</td>
<td>0.37±0.30(^{NR})</td>
<td>0.18±0.21(^{YR})</td>
<td>0.26±0.33</td>
<td>0.16±0.14</td>
</tr>
<tr>
<td>RA of Ephemeroptera</td>
<td>0.24±0.21</td>
<td>0.13±0.16</td>
<td>0.15±0.24</td>
<td>0.11±0.12</td>
</tr>
<tr>
<td>RA of Plecoptera</td>
<td>0.09±0.17</td>
<td>0.01±0.03</td>
<td>0.09±0.20</td>
<td>0.01±0.03</td>
</tr>
<tr>
<td>RA of Trichoptera</td>
<td>0.05±0.09</td>
<td>0.04±0.05</td>
<td>0.02±0.04</td>
<td>0.04±0.07</td>
</tr>
<tr>
<td>RA of Diptera</td>
<td>0.47±0.30(^{NR, LR})</td>
<td>0.78±0.22(^{YR, BR, LR})</td>
<td>0.36±0.29(^{NR})</td>
<td>0.22±0.20</td>
</tr>
<tr>
<td>RA of Chironomidae</td>
<td>0.44±0.36(^{NR, LR})</td>
<td>0.77±0.23(^{YR, BR, LR})</td>
<td>0.32±0.28(^{NR})</td>
<td>0.18±0.20</td>
</tr>
<tr>
<td>EPT/(Chironomidae+Oligochaeta)</td>
<td>1.99±3.03(^{NR})</td>
<td>0.45±0.94(^{YR, BR})</td>
<td>3.86±10.42</td>
<td>1.55±2.64</td>
</tr>
<tr>
<td>FC(^3) richness</td>
<td>1.1±1.0</td>
<td>1.8±2.2</td>
<td>0.9±0.9</td>
<td>1.3±1.0</td>
</tr>
<tr>
<td>SC richness</td>
<td>1.8±2.0(^{NR, LR})</td>
<td>3.7±3.0(^{YR, BR})</td>
<td>0.8±0.9(^{NR, LR})</td>
<td>3.5±1.9(^{YR})</td>
</tr>
<tr>
<td>PR richness</td>
<td>2.7±1.9</td>
<td>3.3±2.7</td>
<td>1.8±1.9</td>
<td>3.1±1.6</td>
</tr>
<tr>
<td>RA of SC</td>
<td>0.08±0.13(^{LR})</td>
<td>0.1±0.12(^{LR})</td>
<td>0.12±0.2</td>
<td>0.18±0.13</td>
</tr>
<tr>
<td>GC+FC richness</td>
<td>6.7±2.8</td>
<td>9.5±5.7</td>
<td>6.0±2.2</td>
<td>7.9±2.3</td>
</tr>
<tr>
<td>cn(^4) Richness</td>
<td>4.7±4.3</td>
<td>6.6±7.5</td>
<td>2.3±2.5(^{LR})</td>
<td>5.3±2.5(^{BR})</td>
</tr>
<tr>
<td>RA of cn</td>
<td>0.22±0.21(^{NR})</td>
<td>0.08±0.1(^{YR, LR})</td>
<td>0.21±0.25(^{LR})</td>
<td>0.34±0.20</td>
</tr>
<tr>
<td>Biological Index</td>
<td>5.00±1.32</td>
<td>5.63±0.68(^{LR})</td>
<td>5.83±2.41(^{LR})</td>
<td>4.57±0.64</td>
</tr>
<tr>
<td>RA of sensitive taxa</td>
<td>0.17±0.22</td>
<td>0.04±0.07(^{LR})</td>
<td>0.27±0.32</td>
<td>0.1±0.1(^{NR})</td>
</tr>
<tr>
<td>Sensitive taxa richness</td>
<td>3.1±3.5</td>
<td>3.4±4.8</td>
<td>2.2±2.2</td>
<td>3.1±1.6</td>
</tr>
<tr>
<td>Tolerant taxa richness</td>
<td>0.8±1.0</td>
<td>0.5±0.8(^{BR})</td>
<td>1.5±0.8(^{NR, LR})</td>
<td>0.7±0.7(^{BR})</td>
</tr>
<tr>
<td>Numbers of food chains</td>
<td>18.5±13.6</td>
<td>21±13.5(^{BR})</td>
<td>10.0±4.9(^{NR, LR})</td>
<td>20.5±8.5(^{YR})</td>
</tr>
<tr>
<td>Average food chain length</td>
<td>3.2±0.3(^{BR, LR})</td>
<td>3.3±0.4(^{BR, LR})</td>
<td>2.8±0.3(^{YR, NR})</td>
<td>2.9±0.3(^{YR})</td>
</tr>
</tbody>
</table>

Notes:


Indices reflecting EPT (Ephemeroptera, Plecoptera, and Trichoptera) composition and diversity (e.g., relative abundance of Plecoptera, relative abundance of EPT, EPT richness, Ephemeroptera richness), and indices related to pollution tolerance (e.g., relative abundance of sensitive taxa, biological index, tolerant taxa richness) showed predictable relationships between the biotic data and environmental variables (hydrodynamic and substrate conditions). These indices were weighted more heavily, accounting for 40.5% of the total weighting. This coincided with the understanding that EPT composition and richness,
and the tolerance of macroinvertebrates to pollution, are key indicators of river and stream ecosystem health (Barbour et al., 1999). EPT-related indices performed well in distinguishing between pristine and disturbed rivers, since the larvae/nymphs of these insects are commonly found in rivers and streams with few anthropogenic disturbances (Kerans et al., 1992). Indices relating to the pollution tolerance of macroinvertebrates also performed well. Hilsenhoff (1988) generated the biotic index, by assigning a value to each taxon, ranging from 0 (sensitive) to 10 (tolerant). This method has proved to be effective in quantifying pollution in aquatic ecosystems (e.g., Kerans & Karr, 1994; Reynolds et al., 1995; Masese et al., 2009; Li et al., 2010). Many ecologists suggest that the tolerance value assigned to a specific taxon may vary with environmental conditions. Therefore, in this study, we determined the tolerance value for each taxon using previously assigned values from several studies from different parts of China (Wang et al., 2003; Wang et al., 2004; Duan et al., 2010).

In contrast, the five indices with the lowest weightings, including taxa richness, Trichoptera richness, predators richness, filtering collectors richness, and relative abundance of the most abundant taxa, demonstrated weak relationships with the physical environmental variables, and accounted for only 5.7% of the total weighting. Taxa richness was calculated to contribute the fifth least to the MM-HSI, which contradicts an intuitive understanding that high diversity is a key indicator of ecosystem health (Karr, 1999; Vörösmarty et al., 2010). We attributed this result to the meso-spatial scale used in this study to quantify biodiversity responses to ecosystem health. Biodiversity on a large-scale, does not necessarily equate to biodiversity on a small-scale. A typical example of 'low diversity but high diversity' can be found in macroinvertebrate assemblages in debris-deposit systems in the Nujiang River (Zhou et al., 2019a). As diversity represents the whole ecosystem, it will increase as ecosystem health improves, since a healthy ecosystem is habitat-heterogenous (Karr, 1999; Wang et al., 2015). An increase in habitat heterogeneity will result in an increase in diversity by supporting organisms in different niches (Tews et al., 2004). At a smaller scale, diversity may depend on variation in the environmental variables selected to describe the status of ecosystem. In this study, we selected two hydrodynamic variables (v and h), and a substrate variable (D_{50}) for multi-metric habitat suitability modeling, as these were the most important three physical variables for stream and river macroinvertebrates in previous studies (Beisel et al., 1998; Duan et al., 2009). Even though the inclusion of extra environmental variables may improve model precision, there is a risk of model overfitting (Hawkins, 2004).

In this study, we included biological integrity indices representing trophic characteristics of macroinvertebrates by using the number of food chains and average food chain length in the MM-HSI. This trial was conducted in view of the key roles that trophic levels play in stabilizing and maintaining the ecosystem (Rooney et al., 2006). These indices were a fairly good fit with the environmental variables (R^2 = 0.502 for numbers of food chains, and 0.484 for average food chain length), and contributed moderately to the MM-HSI (weights = 0.045 and
Comparing with EPT- or pollution-related indices, trophic indices are less important to the model, which may be attributed to the following: (1) The accurate assignment of trophic units to taxa in this study was difficult, despite combining multiple methods including, gut content analysis, $^{13}$C and $^{15}$N stable isotope analysis, allometric diet breadth analysis, and a literature review, since all these methods have shortcomings in determining predator-prey relationships (Petchey et al., 2008); (2) the influence of the hydrodynamic and substrate variables on trophic characteristics were indirect, such as changing inputs, retention of nutrients, and dynamics of primary productivity (Jones & Smock, 1991; Olsen & Townsend, 2005); (3) the two trophic indices selected, reflect the general complexity of food webs, and have controversial relationships with a stable and healthy ecosystem. It is generally understood that complex ecosystems are more stable (McCann, 2000), however, this idea was challenged by Gardner and Ashby (1970) and May (1972) using stochastic mathematical approaches, and suggesting that high complexity may lead to a sudden drop of stability at a threshold value. Loreau and Mazancourt (2013) stressed that intra-species dynamics controlled inter-species complexity on system stability, while Rooney et al. (2006) pointed out that system asymmetry and species asynchrony may play critical roles in maintaining a stable ecosystem. As a result, quantifying trophic values in a complex food web system is difficult, especially when using a simple index system.

3.2 MM-HSM training and verification

The MM-HSM successfully predicted 58.7% of the variation from the observed data from the training dataset (Figure 4a), and 48.9% of the variation from the verification dataset (Figure 4b). The MM-HSI observed values for the dataset used to train the model ranged from 0.071 to 0.772, and the predicted values ranged from 0.147 to 0.591. Sites from both the Yarlung Tsangpo and the Nujiang Rivers varied from low biological integrity values (MM-HSI < 0.2) to high values (MM-HSI > 0.6) (MM-HSIs for the Yarlung Tsangpo River varied from 0.118 to 0.772, and MM-HSIs for the Nujiang River varied from 0.093 to 0.634). The MM-HSIs for sites from of the Bai River, were dispersed at either the low (0.157 ± 0.055) or high (0.580 ± 0.072) ends of the MM-HSI distribution, with no values in the 0.276 to 0.497 range. When verifying the model, MM-HSIs from Bai River sites ranged from 0.298 to 0.642, with most values in an intermediate range (MM-HSI: 0.2–0.6).
Figure 4 Multi-metric habitat suitability index (MM-HSI) comparisons between observed and predicted values using the multi-metric habitat suitability model (MM-HSM) for (a) training and (b) verification performance. Dashed ellipses represent the 95% envelope extensions of the datasets. Areas in grey represent 95% confidence intervals.

Verification of the performance of the MM-HSM as a function of both inter-basin differences (quantified by the proportion of Bai River data in the training dataset) and the size of the training dataset (10–40 sites) is shown in Figure 5. As the Bai River dataset is limited to 12 sites, verification of the model using a proportion of data from Bai River was also limited. However, the influence of inter-basin differences and size of the dataset on the performance of the MM-HSM were still adequately tested as: (1) when the sample size of the training dataset was ten (n = 10), the goodness of fit (R^2) fluctuated around 0.22 regardless of the proportion of data from Bai River; (2) as the size of the training dataset increased from 10 to 40, goodness of fit of the model increased at a slower rate; (3) as the size of the training dataset increased from 10 to 40, the influence of the proportion of Bai River data on the goodness of fit was ‘amplified’; (4) goodness of fit for the model tended to be unchanged when the proportion of Bai River data exceeded 30%. The results showed that both inter-basin differences and the size of the training dataset influenced the performance of the MM-HSM, and that these two factors were intercorrelated. Inter-basin differences, were not apparent when the size of the training dataset was small (e.g., n = 10), however the influence on the model increased as the size of the dataset increased. The performance of the MM-HSM was enhanced as the size of the training dataset increased, however it reached to an asymptotic level when the dataset size exceeded 40.
Differences between river basins (inter-basin) is an important factor that should be considered when untangling the species-environment relationships in especially a large spatial scale since biogeography is an integrative discipline that addresses questions, from the perspective of historical evolution, about species distribution and global patterns of diversity (Wiens & Donoghue, 2004). Previously, ecologists may (as reviewed by Poikane et al., 2016) or may not have (e.g., Dolédec et al., 1998; Statzner et al., 2001) considered biogeography when establishing regional evaluation systems. In this study, the influence of inter-basin differences on the performance of the MM-HSM was limited. This was attributed to: (1) Biodiversity indices are generally considered to be sensitive to biogeography (Heino, 2002), however biodiversity indices made only a limited contribution to the MM-HSI. In this study five biodiversity indices including: taxa richness, EPT richness, Ephemeroptera richness, Plecoptera richness, and Trichoptera richness, made a 21% contribution to the MM-HSI. In particular, taxa richness only contributed 0.8% to the MM-HSM. (2) Morphological structure indices (31%), pollution tolerance indices (24%), and functional structure indices (22%) all contributed strongly to the MM-HSI, and these indices were only weakly influenced by inter-basin differences after nondimensionalization (ratio transformation). (3) Macroinvertebrates are ubiquitous and their movement range is limited to within a mesohabitat scale (around one to several m²; Beisel et al., 1998). Therefore, local environmental conditions (e.g., hydrodynamics and substrate type) play an important role in influencing the diversity and composition of assemblages (Mykrä et al., 2007; Múrria et al., 2017). This was supported by the results of a CCA (Figure 3). Although the Lanmucuo and the Bai Rivers are located in the same region, macroinvertebrate assemblages
were strongly influenced by substrate type. Where sampling sites from the two rivers had a cobble substrate assemblages were similar, however where the sediment was composed of silt, assemblages differed between the two rivers (Zhou et al., 2019b).

3.3 Hydrodynamic simulation and the WUA response of macroinvertebrates to variations in flow

Results of simulations on HSIs for macroinvertebrate assemblages, Polypedilum, Brachycentrus, and Limnophila under low ($Q = 0.52 \text{ m}^3\text{ \cdot s}^{-1}$), medium ($Q = 1.21 \text{ m}^3\text{ \cdot s}^{-1}$), and high ($Q = 1.85 \text{ m}^3\text{ \cdot s}^{-1}$) discharge are presented in Figures 6a–6l. The corresponding $v$, $h$, and $D_{50}$ distributions are provided as supporting material Figures S5a–S5i. The overall suitability index (OSI, calculated as an area-weighted HSI) for macroinvertebrate assemblages ranged from 0.27 (when $Q = 0.11 \text{ m}^3\text{ \cdot s}^{-1}$) to 0.32 (when $Q = 1.21 \text{ m}^3\text{ \cdot s}^{-1}$). These values were similar to the OSI range for Limnophila (0.20–0.34), but lower than those for Polypedilum (0.33–0.45), and higher than those for Brachycentrus (0.01–0.24). The MM-HSI could be treated as an integration of the HSIs for the assembled taxa, therefore, variations of the site-specific MM-HSI in each discharge scenario were fairly compromised comparing with the corresponding HSIs for the specific indicator taxa. For example, when $Q = 1.21 \text{ m}^3\text{ \cdot s}^{-1}$, the site-specific MM-HSI values ranged from 0.01–0.51 (0.32 ± 0.10), while those for Polypedilum, Brachycentrus, and Limnophila ranged from 0–0.88 (0.38 ± 0.23), 0–1 (0.19 ± 0.21), and 0.01–0.91 (0.34 ± 0.23), respectively.

Changes in weighted usable area (WUA) values in response to flow for macroinvertebrate assemblages are shown in Figure 7a. Macroinvertebrate assemblages had a unimodal response in WUA to changes in flow discharge, with an optimal flow ($Q$) of 1.21 m$^3$·s$^{-1}$ and a WUA maximum value of 9100 m$^2$. This supported our hypothesis that for macroinvertebrate assemblages in alpine rivers during summer, an intermediate flow discharge resulted in the highest biological integrity, based on indices of biodiversity, morphological structure, functional structure, pollution tolerance, and trophic structure. This response pattern fits within the framework of the intermediate disturbance hypothesis (Connell, 1978), although our results are based on biological integrity, rather than biodiversity. Generally, the disturbance caused by a high flow discharge inhibited macroinvertebrate assemblages, while at a low flow assemblages tended to be dominated by a few taxa (Roxburgh et al., 2004). Both extremes (high and low flow) may result in a decline in the biological integrity of macroinvertebrates. In contrast, an intermediate flow discharge provides suitable niches for most macroinvertebrates. In this case, the optimal flow discharge, $Q = 1.21 \text{ m}^3\text{ \cdot s}^{-1}$ for the highest biological integrity, was
Figure 6 Distributions of habitat suitability indices (HSIs) for (a–c) macroinvertebrate assemblages, (d–f) *Polypedilum*, (g–i) *Brachycentrus*, and (j–l) *Limnophila* in the study reach of the Lanmucuo River under different flow conditions: 0.52 m$^3\cdot$s$^{-1}$ (column 1), 1.21 m$^3\cdot$s$^{-1}$ (column 2), and 1.85 m$^3\cdot$s$^{-1}$ (column 3).

also the median hydrological flow condition in the Lanmucuo River in summer according to the gauged data (60% recurrence rate; Figure S1b). This discharge fosters the formation of gravel- and cobble- beds in the Lanmucuo River (Carling, 1988). A mixture of gravel and cobbles is the ideal substrate for macroinvertebrates, providing numerous living spaces (Duan et al., 2009), and a relatively
stable bed that can persist even during periods of relatively high flow (Zhao et al., 2015). In addition, an intermediate discharge allows for nutrient transportation and retention in the river, and supports the trophic needs of well-developed macroinvertebrate assemblages (Ensign & Doyle, 2005).

![Figure 7](image)

Figure 7 Changes in weighted usable area (WUA) in response to variations in flow for (a) macroinvertebrate assemblages \(\text{WUA}_{\text{MM}}\), (b) \textit{Polypedilum} \(\text{WUA}_{\text{P}}\), (c) \textit{Brachycentrus} \(\text{WUA}_{\text{B}}\), and (d) \textit{Limnophila} \(\text{WUA}_{\text{L}}\).

Changes in weighted usable area (WUA) in response to variations in flow for each of three indicator taxa (\textit{Polypedilum}, \textit{Brachycentrus}, and \textit{Limnophila}) are shown in Figures 7b–7d. Each taxon had a unique response in WUA to changes in flow discharge. \textit{Polypedilum} had a left-skewed, unimodal response, indicating a preference for low flow in shallow riffles. \textit{Brachycentrus} and \textit{Limnophila} had monotonically increasing responses, indicating preferences for high flow in pools. We attributed the monotonic patterns of \textit{Brachycentrus} and \textit{Limnophila} to that the ecological amplitudes of these two taxa were relatively large, due to their pool preferences, against the limited, real environmental gradient in
the Lanmuquo River, leading to insufficient exhibitions of the whole unimodal patterns (Lepš & Šmilauer, 2003). The response pattern of macroinvertebrate assemblages is essentially a weighted combination of each taxon (Wright, 1995), which is regulated by inter-taxon interactions within assemblages (Feng et al., 2017). Our results showed the importance of assessing the whole macroinvertebrate assemblage, and the utility of biological integrity as a measure of river health. However, if only a few taxa were monitored, selecting highly sensitive taxa with a predictable response to the environmental gradient (e.g., EPT taxa; Park et al., 2003; Li et al., 2009), or selecting taxa with high trophic values would give the best results (e.g., predatory fishes; Rooney et al., 2006; Daugherty et al., 2009).

With the effects of anthropogenic activities and climate warming, flow regimes in rivers and streams on the QTP are currently undergoing intensive modification. On one hand, inter-basin water transfer projects have reduced flow discharge by withdrawing water from original channels (e.g., 8 billion m$^3$ per year during the first stage of the western route of the South-to-North Water Diversion Project; Li et al., 2017). On the other hand, glacier degradation triggered by warmer temperatures resulted in an increase in flow discharge (estimated as 1.0–7.2% at 1.5–4.0 °C; Wang et al., 2021). Both factors threaten the vulnerable river and stream ecosystems of the QTP. By using structural and functional characteristics of aquatic biotic assemblages, the MM-HSM provides a holistic and quantitative overview of the impacts of different flow regimes. Estimating optimal flow discharge is an ecologically meaningful alternative to traditional hydrology-based models, hydraulic rating models, or single taxon-based habitat simulations (Tharme, 2003). Optimal discharge values based on macroinvertebrate assemblages may differ between alpine and lowland rivers. It is possible that optimal discharge for an alpine river may be lower than that for a similar lowland river, as the poor assemblage complexity of an alpine aquatic ecosystem may not allow for intensive variations in flow (Hamerlik & Jacobsen, 2012). Optimal discharge may vary with season, and periodic fluctuations in flow discharge may be better for the development of macroinvertebrate assemblages compared with a permanent steady flow. In accordance with the flood pulse concept (Junk et al., 1989), a predictable, regular high flow is the primary driving force for the existence, productivity, and interactions among macroinvertebrate assemblages in river-floodplain systems. This flood pulse improves basin-scale habitat heterogeneity by creating riparian ecotones along perennial water bodies, and by stimulating biota to develop adaptations for efficient utilization of habitats and resources, rather than depend on permanent water (Sparks et al., 1990). Periodic fluctuations in flow discharge also significantly improved trophic stability (indicated by the higher persistence of top predators to perturbation) of river ecosystems, as shown by Power et al. (1995) using food web simulations based on Lotka-Volterra dynamics. Therefore, in future studies, we plan to combine within-year flow fluctuations and seasonal optimal flow into the MM-HSM framework, which can be applied to the integrated management and restoration of river basins.
4 Conclusions

A multi-metric habitat suitability model (MM-HSM) was developed using biological integrity indices for macroinvertebrate assemblages from the QTP. The model was robust to differences between river basins and changes in dataset size. For a typical meandering river on the QTP, the relationship between WUA and flow discharge (0.11–1.99 m$^3$·s$^{-1}$) was unimodal for macroinvertebrate assemblages, and the optimal flow discharge was 1.21 m$^3$·s$^{-1}$. This unimodal relationship between WUA and flow discharge for macroinvertebrate assemblages, contrasted with the skewed unimodal or monotonically increasing relationships observed for individual indicator taxa, supporting the hypothesis that flow-biota relationships conform to the intermediate disturbance hypothesis on the QTP. The establishment of the MM-HSM provided a quantitative and holistic way to couple biological and environmental factors from river and stream ecosystems. The optimal flow discharge derived from this model offers an ecologically meaningful value which can be applied to the management of rivers and streams.

Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (NSFC) (No. 91547204, No. 51779120), and the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (No. 2019QZKK0903). The macroinvertebrate indices, and hydrodynamic and substrate records for the 93 sampling sites used in this study are available from datasets ‘Macroinvertebrate Indices, Hydrodynamics, and Substrate in QTP’ (https://doi.org/10.7910/DVN/KNK92K). The digitalized geomorphological model for the study reach of the Lanmucuo River are available from datasets ‘Lanmucuo Geo. file’ (https://doi.org/10.7910/DVN/G09ZOU). The hydrodynamic simulations and corresponding MMI/HSIs evaluations for the study reach of the Lanmucuo River under 18 flow discharge scenarios are available from datasets ‘Hydrodynamic simulations and HSIs’ (https://doi.org/10.7910/DVN/1QDJMY). We thank Jiukai Tang, Luo Sun, Wei Liu, Xiangchao Wang, Xiaoli Fan, Xing Liu, Yufeng Yang, and Yutong Zheng (alphabetically) for the field assistances in macroinvertebrate sampling and/or environment survey. We also thank Karen Tricklebank, PhD, from Liwen Bianji (Edanz) (www.liwenbianji.cn/ac) for editing the English draft of this manuscript.

References


Evolution, 34(8), 746–758. https://doi.org/10.1016/j.tree.2019.03.010


28