The potential habitats of two submerged macrophytes, *Myriophyllum spicatum* and *Hydrilla verticillata* in the river ecosystems, South Korea

Deokjoo Son¹, Kang-Hyun Cho² and Eun Ju Lee¹,*

¹ School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea
² Department of Biological Sciences, Inha University, Incheon 22212, Republic of Korea

Abstract – We examined the environmental factors that characterize the habitats of two submerged macrophytes, *Myriophyllum spicatum* and *Hydrilla verticillata*, in South Korea using generalized additive models (GAMs). We performed a vegetation survey and measured water depth and water velocity from May to September, 2014–2015. Averaged water quality data from the Ministry of Environment’s national water quality measurement network from January 2012 to October 2015 were used for modeling. Potential habitats of *M. spicatum* were linked with chlorophyll *a*, nitrate nitrogen, suspended solids, water temperature, water depth, and water velocity (deviance explained = 28.7%, accuracy = 74%). In the case of *H. verticillata*, electrical conductivity and suspended solids were key habitat factors (deviance explained = 23.4%, accuracy = 75%). Model results were highly consistent with observations and field data. Model performances were evaluated by the accuracy rate, the area under the receiver operating characteristic curve, the kappa value, and field verification, and were in generally good agreement. Few studies have evaluated the developed models using the independent field data. By understanding the environmental factors that characterize the habitats for submerged macrophytes, our results contribute to the development of conservation and management strategies for river ecosystems.

Keywords: : Eurasian watermilfoil / Hydrilla / water condition / habitat suitability / generalized additive model


**Mots-clés** : myriophylle eurasienne / Hydrilla / qualité de l’eau / convenance de l’habitat / modèle additif généralisé

*Corresponding author: ejlee@snu.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution License CC-BY-ND (http://creativecommons.org/licenses/by-nd/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. If you remix, transform, or build upon the material, you may not distribute the modified material.*
1 Introduction

Submerged macrophytes play an important role as a producer in food webs, shelter and forage for other organisms, and as a water quality indicator (Nieder et al., 2004). In addition, submerged macrophytes produce oxygen in stagnant regions and prolong the hydrologic retention time for the removal of particulate nutrients (Nepf et al., 2007). Despite the importance of submerged macrophytes, the formation of dense monotypic stands has adverse effects on the diversity and richness of invertebrates and fish (Buchan and Padilla, 2000). Dense beds of submerged macrophytes can produce organic materials during drought periods (Woo, 2010). In particular, their proliferation can impede water flow, clog inlets of reservoirs, and interfere with recreational activities (Kenneth, 1996).

The spatial distribution of organisms is related to species dispersal and survival at a regional scale. In addition, abiotic conditions (environmental constraints) and biotic interactions (e.g., competition and herbivory) influence species distributions at a local scale (Austin, 2002; Bućas et al., 2013; Chappuis et al., 2014). The occurrence and abundance of submerged macrophytes are influenced by chemical and physical factors, such as water quality, light availability (Dennison et al., 1993), water transparency, water depth (Canfield et al., 1985), channel slope, channel dimensions (O’Hare et al., 2011), and hydrological regime (Franklin et al., 2008). Understanding how diverse environmental factors affect the habitats of submerged macrophytes is important for flow control, sediment transport (Järvell, 2005), and assessments of the ecological condition of rivers (Clayton and Edwards, 2006).

A variety of statistical approaches such as generalized linear models and generalized additive models (GAMs) are important tools for predicting the likely occurrence or distribution of a species (Pearce and Ferrier, 2000; Austin, 2002). In particular, GAMs are used extensively in habitat suitability modeling and identification of the optimal environmental conditions for a given species; data are fitted using a semi-parametric model to predict non-linear responses to the exploratory variables (Elith et al., 2006; Drexler and Ainsworth, 2013; Li and Wang, 2013). GAMs not only have a strong statistical foundation, but can be used to realistically model ecological relationships (Yee and Mitchell, 1991; Sanchez et al., 2008). When the relationship between a species distribution and environmental variables is complex, GAMs are practical and perform as well or better than other types of predictive models (Drexler and Ainsworth, 2013; Li and Wang, 2013). GAMs have been used to examine potential seagrass habitats (Lathrop et al., 2001; Downie et al., 2013), fish production and distributions (Borchers et al., 1997; Buisson and Blanc, 2008; Murase et al., 2009; Solanki et al., 2016), and terrestrial plant distributions (Yee and Mitchell, 1991; Austin and Meyers, 1996; Thuiller et al., 2005), but few studies have used GAMs to examine submerged macrophytes in river ecosystems.

Recently, river ecosystems in South Korea have experienced channel dredging, channelization, and dam construction for flood control during rainfall periods and to secure water resources during drought periods (Woo, 2010). In particular, the “Four Major Rivers Project” (2009–2012) involved the construction of 16 weirs and three dams in the Han, Geum, Nakdong, and Yeongsan Rivers (Lah et al., 2015). This national project aimed to secure water resources, reduce flooding, improve water quality, and create multipurpose public spaces for local residents (Jun and Kim, 2011). Despite substantial controversy surrounding the effectiveness of this project (Normile, 2010), it drastically changed the natural riverine habitats and enabled the artificial manipulation of the water level and the regulation of water flow (Lee et al., 2014). The modified slow velocity in regulated streams may increase the abundance of macrophytes (Bunn and Arthington, 2002) and devastate habitats for organisms adapted to the natural discharge regime (Dynesius and Nilsson, 1994).

Alterations to hydrological regimes affect the structure and function of aquatic ecosystems, resulting in changes in the spatial distributions of submerged macrophytes (Tian et al., 2015). Submerged macrophytes in the river ecosystems of South Korea were found to change from slow flowing streams to large rivers after weir construction. Myriophyllum spicatum L. (Eurasian watermilfoil) and Hydrilla verticillata (L. f.) Royle (Hydrilla) are native species in South Korea; however, they are fast growing and the most abundant macrophytes in Korean rivers. They are invasive species and strong competitors in Europe, the United States, and South America owing to their rapid and dense growth (Van et al., 1999; Gassmann et al., 2006; Beck et al., 2008). These two species are cosmopolitan angiosperms with extensive worldwide ranges (Zhou et al., 2016) and overrun various habitats, from lentic to lotic systems, and in turn affect flow velocity and nutrient cycling in the water column (Sousa, 2011).

Accordingly, it is necessary to understand the current distributions as well as predict suitable habitats of submerged macrophytes with high invasive potentials for river management and conservation planning. Few studies have predicted the distributions of submerged macrophytes in the rivers using GAMs (Ahmadi-Nedushan et al., 2006; Camporeale and Ridolfi, 2006). Moreover, the developed GAMs have rarely been validated via independent field data (Guisan et al., 2002). The aims of the study were to examine under what environmental conditions M. spicatum and H. verticillata are likely to occur in and to predict and validate their suitable habitats. Our specific objectives were (1) to characterize the relationship between environmental variables and the occurrence of two submerged macrophytes (M. spicatum and H. verticillata) using GAMs, (2) to predict the suitable habitats for these two submerged macrophytes, and (3) to evaluate GAMs by applying the independent field data.

2 Materials and methods

2.1 Study sites and data collection

The study sites were located in the central and southern parts of the Korean Peninsula (33°–39°N, 124°–130°E) which has a temperate climate. The mean annual precipitation is approximately 1300 mm, and the mean annual temperature is 10–14 °C (Korea Meteorological Administration, http://www.kma.go.kr/weather/climate/average_normal.jsp). Two-thirds of the annual precipitation occurs in the summer
The winter season (December–February) tends to be dry and cold, reflecting a typical continental weather pattern (Shin and Chung, 2011).

Surveyed streams and rivers converge into five major rivers: Han River, Geum River, Nakdong River, Yeongsan River, and Seomjin River (Fig. 1). For the vegetation survey, 197 sites in the rivers and tributaries were randomly selected, including 71 sites in the Han River, 43 sites in the Geum River, 46 sites in the Nakdong River, 27 sites in the Yeongsan River, and 10 sites in the Seomjin River. Sites were assigned to water quality monitoring towers operated by the Ministry of Environment. The presence and absence of submerged macrophytes were recorded using a belt transect (50 m × 2 m) parallel to the direction of water flow from May to September in 2014 (June–August); examinations were performed by wading or in a boat with a rake for confirmation (Park et al., 2013).

Water quality data were acquired from the Ministry of Environment’s national water quality measurement network (http://water.nier.go.kr) from January 2012 to October 2015. Averaged water quality data calculated from monthly measurements were used for modeling. We chose 14 water environmental variables that had a postulated relationship with the ecological requirements of submerged macrophytes (Gallego et al., 2015; Klippel et al., 2016), namely, water temperature, pH, biochemical oxygen demand (BOD), chemical oxygen demand (COD), suspended solids, total nitrogen, ammonium nitrogen, nitrate nitrogen, total phosphorus, total organic carbon, electrical conductivity, total dissolved nitrogen, total dissolved phosphorus, and chlorophyll a. In each transect, water depth (meter stick) and water flow (Flowwatch; JDC Electronic SA, Yverdon-les-Bains, Switzerland) were measured three times where vegetation was most abundant, and the values were averaged. If submerged macrophytes were lacking, water depth and water flow were measured in triplicate at the midpoint of the transect. We recorded channel structure and confluence points according to in situ survey and stream order data obtained from the Water Resources Management Information System (the Ministry of Land, Infrastructure and Transport, www.wamis.go.kr).

2.2 Model building

All statistical analyses were performed in R (R Development Core Team, 2016). Prior to the statistical analysis, all explanatory variables were log_{10}-transformed to improve normality (Chappuis et al., 2014). Pearson correlation analyses were performed to detect high multicollinearity (Zhao et al., 2014; Wedding and Yoklavich, 2015) and the variables that have correlation coefficient greater than 0.75 with another variable were eliminated (Kuhn and Johnson, 2013) (Tab. S1). Total nitrogen, BOD, COD, ammonium nitrogen, total phosphorus, total organic carbon, and total dissolved nitrogen were excluded because they were highly correlated with suspended solids, nitrate nitrogen, and total dissolved phosphorus. Multivariate statistical methods, such as principal component analysis (PCA), have been recommended to determine the main environmental factors before inclusion in models (Shmueli, 2010; Zhao et al., 2014). Six significant elements of water environmental factors determined via PCA were selected and two elements from in situ measurements (water depth and water velocity) were added to the GAMs (Tab. S2). Finally, eight environmental descriptors were included in GAMs: chlorophyll a, electrical conductivity, nitrate nitrogen, suspended solids, total dissolved phosphorus, water temperature, water depth, and water velocity. All variables describing the physical and chemical properties of water included in the GAMs are described in Table 1.

We used GAMs to approximate the probability of taxon presence with respect to the predictors. A binominal distribution was specified (presence = 1 and absence = 0) with a logit link function relating the dependent variables to the predictors. Such approach allows modeling presence–absence data at survey sites, thereby enabling the probability of species occurrence to be predicted from independent data at unsurveyed sites (Pearce and Ferrier, 2000). The descriptor variables were modeled as cubic splines, with four degrees of freedom for smoothing splines (Lehmann, 1998;
Table 1. Mean, standard error (SE), minimum (Min), and maximum (Max) values for physical and chemical properties of water at 197 sites. Chlorophyll $a$, electrical conductivity, total dissolved phosphorus, nitrate nitrogen, and suspended solids are presented as mean values of monthly estimates from 2012 to 2015, water temperature is presented as the mean from May to October, and water depth and velocity data were obtained at the sampling date.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Mean</th>
<th>Median</th>
<th>SE</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorophyll $a$</td>
<td>mg/m$^3$</td>
<td>14.0</td>
<td>9.6</td>
<td>0.9</td>
<td>0.9</td>
<td>65.1</td>
</tr>
<tr>
<td>Electrical conductivity</td>
<td>µS/cm</td>
<td>328</td>
<td>258</td>
<td>19</td>
<td>77</td>
<td>1515</td>
</tr>
<tr>
<td>Total dissolved phosphorus</td>
<td>mg/L</td>
<td>0.076</td>
<td>0.035</td>
<td>0.009</td>
<td>0.008</td>
<td>1.282</td>
</tr>
<tr>
<td>Nitrate nitrogen</td>
<td>mg/L</td>
<td>2.2</td>
<td>2.1</td>
<td>0.1</td>
<td>0.5</td>
<td>5.9</td>
</tr>
<tr>
<td>Suspended solids</td>
<td>mg/L</td>
<td>9.6</td>
<td>8.3</td>
<td>0.5</td>
<td>1.1</td>
<td>35.9</td>
</tr>
<tr>
<td>Water temperature</td>
<td>°C</td>
<td>23.1</td>
<td>23.5</td>
<td>0.1</td>
<td>17.7</td>
<td>26.0</td>
</tr>
<tr>
<td>Water depth</td>
<td>m</td>
<td>0.61</td>
<td>0.50</td>
<td>0.02</td>
<td>0.10</td>
<td>1.60</td>
</tr>
<tr>
<td>Water velocity</td>
<td>m/s</td>
<td>0.10</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Table 2. Selected environmental variables and deviance explained in GAMs for *Myriophyllum spicatum* and *Hydrilla verticillata*.

<table>
<thead>
<tr>
<th>Species</th>
<th>Environmental variable</th>
<th>Significance (p-value)</th>
<th>Deviance explained</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Myriophyllum spicatum</em></td>
<td>Chlorophyll $a$</td>
<td>&lt;0.001</td>
<td>28.7%</td>
</tr>
<tr>
<td></td>
<td>Nitrate nitrogen</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Suspended solids</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water temperature</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water depth</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water velocity</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td><em>Hydrilla verticillata</em></td>
<td>Electrical conductivity</td>
<td>&lt;0.001</td>
<td>23.4%</td>
</tr>
<tr>
<td></td>
<td>Suspended solids</td>
<td>0.010</td>
<td></td>
</tr>
</tbody>
</table>

Wood, 2000), using the “mgcv” library for GAMs (Wood, 2006). The most parsimonious model for each species with the fewest variables was chosen using a stepwise selection procedure (Buisson et al., 2008). To determine the best fit model, Akaike’s information criterion (AIC) was used as a goodness-of-fit statistic (Zuur and Pierce, 2004; Buisson et al., 2008; Sanchez et al., 2008). Models with a smaller AIC were able to explain the residual deviance better than those with a larger AIC. The best model accounted for the most variation in the data using the fewest predictors (Burnham and Anderson, 1998).

2.3 Validation of the predictive performance of models

To generate a confusion matrix, the Youden index was chosen as the threshold probability, which was the maximum difference between sensitivity (the probability of correct classification as positive) and specificity (the probability of correct classification as negative) (Jiménez-Valverde and Lobo, 2007; Freeman and Moisen, 2008), using the “SDMTools” library for model accuracy (VanDerWal et al., 2014). Coordinates of observations and predictions based on the confusion matrix were projected to World Geodetic System 84 (WGS84) using QGIS (QGIS Development Team, 2016).

Model accuracy was assessed by two measures: Cohen’s kappa and the area under the receiver operating characteristic (ROC) curve (AUC) (Zimmermann et al., 2007). Cohen’s kappa is the most common method for determining the accuracy of presence–absence predictions based on a selected threshold probability, regardless of variation in prevalence (Segurado and Araujo, 2004; Allouche et al., 2006). Landis and Koch (1977) suggested the following interpretation of kappa values: excellent agreement, >0.75; good agreement, 0.40–0.75; and poor agreement, <0.40. Another method for assessing the accuracy of models uses the AUC value as a threshold-independent criterion (Fielding and Bell, 1997). To construct ROC curves, all possible thresholds were used to classify the scores into confusion matrices, and the sensitivity and specificity were estimated for each matrix (Allouche et al., 2006). According to Swets (1988), AUC values were interpreted as follows: excellent, >0.90; good, 0.80–0.90; fair, 0.70–0.80; poor, 0.60–0.70; fail, 0.50–0.60. Moreover, in
situ observations of the 41 sites were randomly performed to evaluate the accuracy of the model for field verification.

3 Results

3.1 GAM response curves

We defined the ranges of environmental factors with respect to species occurrences. The models explained 28.7% and 23.4% of the observed variation in selected variables for *M. spicatum* and *H. verticillata*, respectively (Tab. 2). For *M. spicatum*, we found that chlorophyll *a*, nitrate nitrogen, suspended solids, water temperature, water depth, and water velocity were significant variables in the GAM. Based on the response curves, the probability of *M. spicatum* presence increased as nitrate nitrogen increased, and decreased as water temperature and suspended solids increased (Fig. 2A). The response curves for chlorophyll *a* indicated a sharp increase in the predicted presence of *M. spicatum* from 0 to ∼20 mg/m³ and a decline at higher concentrations. The presence probabilities of *M. spicatum* for water velocity decreased for values of up to ∼0.4 m/s and then increased in relatively rapid flows. *Myriophyllum spicatum* was distributed in a wide range of water depths up to 1 m and decreased in deeper water. Electrical conductivity and suspended solids were important parameters determining the *H. verticillata* distribution (Tab. 2). *Hydrilla verticillata* presence was negatively related to electrical conductivity and suspended solid concentrations (Fig. 2B). We also observed differences in the predicted potential habitats between the two species (Fig. 3). *Myriophyllum spicatum* was widely distributed in the Han River and Nakdong River, and its predicted distribution was similar to its current distribution. *Hydrilla verticillata* was abundant everywhere, whereas its distribution was predicted to be decreased in Nakdong River and increased in Yeongsan River.

3.2 Model validation and field verification

To assess model performances for each species, we compared predicted potential habitats with observed habitats

![A) Myriophyllum spicatum](image1.png)

![B) Hydrilla verticillata](image2.png)

Fig. 2. Response curves of (A) *Myriophyllum spicatum* and (B) *Hydrilla verticillata* for environmental gradients in GAMs. The vertical axes represent the probabilities of occurrence, and shaded bands show the 95% confidence interval. Rug plots on the x-axis show data points. Chla, chlorophyll *a*; EC, electrical conductivity; NO₃N, nitrate nitrogen; SS, suspended solids; WT, water temperature; WD, water depth; WV, water velocity.
using the confusion matrix (Tab. 3). We observed accuracy rates of 0.74 for *M. spicatum* and 0.75 for *H. verticillata*. For *M. spicatum*, we observed AUC = 0.84 and kappa = 0.46 when we applied a threshold of 0.536. For *H. verticillata*, we observed AUC = 0.79 and kappa = 0.39 when the threshold was 0.492. Based on the field verification, we observed accuracy rates of 76% for the two species (Tab. 4). In addition, we established AUC values of 0.75 for *M. spicatum* and 0.82 for *H. verticillata*, and kappa values of 0.44 for *M. spicatum* and 0.25 for *H. verticillata*.

4 Discussion

In this study, we identified factors that characterize the suitable habitats of *M. spicatum* and *H. verticillata* using GAMs, namely, chlorophyll *a*, electrical conductivity, nitrate nitrogen, suspended solids, water temperature, water depth, and water velocity. GAMs are very useful for describing the complex relationships between response variables and environmental factors (Cheng and Gallinat, 2004); however, our models explained low proportions of deviance (Tab. 2). This can be explained by the patchy distribution of submerged macrophytes, which cannot be fully explained by the selected variables (Lehmann, 1998). Nevertheless, based on the model evaluation procedures, the response curves for each variable enable us to infer general trends and we expect to adapt models to other locations around the world.

4.1 GAM results and environmental factors

The abundance and distribution of submerged macrophytes in river ecosystems are related to water quality conditions (Nieder et al., 2004), water depth, and water velocity (Sousa, 2011). We found that variables associated with water environmental factors were important determinants of the distributions of *M. spicatum* and *H. verticillata*, especially chlorophyll *a*, electrical conductivity, nitrate nitrogen, suspended solids, water temperature, water depth, and water velocity. Gradients of these variables also determine suitable habitats for submerged macrophytes in previous studies (Dodkins et al., 2005; Lacoul and Freedman, 2006; Franklin et al., 2008).

Our results demonstrated that the occurrence of *M. spicatum* increased for low concentrations, and decreased for high concentrations of chlorophyll *a*. In addition, the probability of *M. spicatum* presence decreased as water temperature increased. A higher temperature within optimal ranges usually promotes a higher chlorophyll *a* concentration and productivity as well as a greater abundance of submerged macrophytes (Barko et al., 1986). However, competition for light between aquatic plants and phytoplankton may limit plant growth and even result in the disappearance of taxa (Rybicki and Landwehr, 2007; Borne et al. and Puijalon, 2011). Jones et al. (1983) also found that phytoplankton-created low water clarity inhibits *M. spicatum* growth. Furthermore, the adaptability of submerged macrophytes to low temperatures may play a role in interspecific competition because the optimal water temperature for submerged macrophytes is 28–32 °C (Barko et al., 1986).

The occurrence of *H. verticillata* was high when electrical conductivity was low in oligotrophic water (Fig. 2B). This agrees with the findings of Sousa et al. (2009) and Cook and Lüönd (1982) who found *H. verticillata* develops well in oligotrophic waters. Plant growth usually increases as the
concentration of nutrients in water and sediment increases (Van et al., 1999; Yu et al., 2010; Sousa, 2011). Moreover, Kennedy et al. (2009) found that H. verticillata can thrive not only in eutrophic waters, but also in oligotrophic waters. According to Sousa (2011), however, eutrophic conditions may have negative effects on H. verticillata growth via the proliferation of plankton, which compete with submerged macrophytes for light and nutrients. Electrical conductivity, as a measure of the chemicals summary variable (Heegaard et al., 2001), may affect macrophyte composition and be unfavorable for submerged macrophytes that are sensitive to eutrophication (Thomaz et al., 2003; Lauridsen et al., 2015). The occurrence of M. spicatum was high when nitrate nitrogen was high in water (Fig. 2A). Nitrogen is a key element for aquatic plants, which use nitrate as a nitrogen source (Bornette and Puijalon, 2011) and M. spicatum was related to with higher nitrate availability (Ali and Soltan, 2006).

We detected abundant M. spicatum in water at depths of 0.7–1.0 m, and a decreased abundance in deeper water (Fig. 2A). This result concurs with a previous study; Nichols and Shaw (1986) reported that M. spicatum is typically found in water ranging from 1 to 4 m deep. Angradi et al. (2013) observed that the optimal depth for submerged macrophytes is 1.2 m and <1 m in turbid conditions. However, an increase in water depth causes a light deficiency for submerged macrophytes (Bornette and Puijalon, 2011), thereby decreasing the rate of photosynthesis. Shallow water bodies allow more light penetration, provided that the water is not turbid (Narumalani et al., 1997). Lower water clarity owing to sediments, turbidity, and nutrients can reduce the water depth and spatial distribution for growth and survival of submerged macrophytes (Dar et al., 2014; Patrick et al., 2014). Consequently, as depth increases or water clarity decreases, light availability for photosynthesis may diminish (Lacoul and Freedman, 2006).

Based on the velocity response curve, we observed a decrease in M. spicatum abundance for values of up to ~0.4 m/s and an increase for higher values (Fig. 2A). On the other hand, M. spicatum has been found in the stream at water velocities of 2 m/s (Nichols and Shaw, 1986). A number of water velocity readings in this study were at 0 m/s, with an average of 0.1 m/s and a median of 0 m/s. The regulated rivers and streams in study sites were close to lentic conditions; accordingly, we could not precisely evaluate the velocity response of M. spicatum. However, in general, the biomass and richness of submerged macrophytes are higher at 0.3–0.4 m/s, and lower at higher velocities because of failure to establish and colonize (Lacoul and Freedman, 2006). Moderate flow can encourage submerged macrophyte growth by continuously washing photosynthetic tissues covered with epiphytic algae (Strand and Weisner, 1996; Lehmann, 1998).

### 4.2 Model validation and field verification

We did not develop GIS-based predictive maps to identify potential distribution for submerged macrophytes because it was difficult to construct a bathymetric map of all rivers and streams in South Korea. We present maps showing areas of agreement between observations and predictions. Although we did not generate spatial maps interpolated with predicted probabilities, we were able to detect areas with abundant submerged macrophytes in the five rivers and characterize the suitable habitats of them.

Correct and incorrect predictions in a confusion matrix indicate the strength of predictions (Peters et al., 2007). The misclassified sites in the predicted distribution according to GAMs were related to channel structure, rather than water quality. Most of these sites were confluence points, i.e., sites at which two channels met, each carrying independent influxes and sediment discharge (Benda et al., 2004). To examine false positive errors (commission error; observation = 0 and prediction = 1), we considered the characteristics of survey sites at confluence points that connected relatively larger

---

**Table 3.** Comparison of predicted and observed distributions of Myriophyllum spicatum and Hydrilla verticillata during model building. Predicted values were obtained from the fitted probability of presence using the Youden index to apply a threshold probability: 0.536 for M. spicatum and 0.492 for H. verticillata.

<table>
<thead>
<tr>
<th>Species</th>
<th>Contents</th>
<th>Predicted absence</th>
<th>Predicted presence</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. spicatum</em></td>
<td>Observed absence</td>
<td>91</td>
<td>23</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Observed presence</td>
<td>29</td>
<td>54</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>120</td>
<td>77</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>Correct prediction</td>
<td>(91 + 54)/197 = 0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of commission</td>
<td>23/114 = 0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of omission</td>
<td>29/83 = 0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kappa</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>H. verticillata</em></td>
<td>Observed absence</td>
<td>115</td>
<td>13</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>Observed presence</td>
<td>37</td>
<td>32</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>152</td>
<td>45</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>Correct prediction</td>
<td>(115 + 32)/197 = 0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of commission</td>
<td>13/128 = 0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of omission</td>
<td>37/69 = 0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kappa</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

tributaries to main water bodies. Confluences have been described as biodiversity hot spots with physical heterogeneity and habitat complexity owing to diverse physical, chemical, and biological attributes resulting from tributary streams (Kiffney et al., 2006; Rice et al., 2006). At these sites, we expected to observe submerged macrophytes; however, their distributions were discontinuous and bed sediment size and flow properties were unstable (Rice et al., 2006). Slow velocity conditions and the maintenance of stable substrates are necessary for macrophytes colonization (Riis and Biggs, 2003). False positive results imply that submerged macrophytes did not have an opportunity to establish to a suitable habitat (Buchan and Padilla, 2000). We observed false negative errors (omission error; observation = 1 and prediction = 0) at study sites that were typically downstream of confluence points connecting relatively small tributaries and irrigation ditches with the potential for submerged macrophytes dispersal. Myriophyllum spicatum and H. verticillata are found in rivers, lakes, irrigation ditches, and other waterways (Netherland, 1997; Eiswerth et al., 2000). Their primary dispersal strategy is vegetative reproduction by fragmentation, moving through small waterways and then establishing at channel junctions. Vegetative reproduction by stem fragmentation is an efficient mechanism for dispersal, colonization, and overwintering (Sousa, 2011; Xie et al., 2013). The conditions downstream of small tributaries are sufficiently stable for submerged macrophytes inhabitation. Submerged macrophytes habitats are regulated by a variety of factors, varying within not only whole streams, but also smaller stream reaches (Riis et al., 2001).

We observed high accuracy rates for each species model (0.74 for M. spicatum and 0.75 for H. verticillata), but low proportions of variation explained by the models. We observed kappa values (which were dependent on a threshold) of 0.46 for M. spicatum and 0.39 for H. verticillata, indicating a fair model fit. In addition, we observed AUC values (independent of threshold values) of 0.84 for M. spicatum and 0.79 for H. verticillata, indicating satisfactory predictive ability. According to the AUC and kappa values, model performance was good for both species. Field verification to validate favorable potential habitats for M. spicatum and H. verticillata confirmed model performance, supporting their good prediction abilities, based on accuracy rates, AUC, and kappa values, except for the kappa value (0.25) of H. verticillata.

Although GAMs do not provide superior predictive performance compared with other models such as multivariate adaptive regression splines and boosted regression trees (Austin, 2007), they are flexible enough to model relationships between occurrences of submerged macrophytes and environmental factors (Murase et al., 2009). Habitats for submerged macrophytes are characterized by a complex set of physical, chemical, and biological parameters. In this study, the modeling of potential habitats for submerged macrophytes was in good agreement, despite only considering water chemicals, water depth, and water velocity. However, a reasonable possibility of prediction errors is the reason why the realized niches of submerged macrophytes were not completely explained by the variables that we selected and added to GAMs. We measured water depth and water velocity once in normal conditions, and we used water chemical data that were averaged over four years. Even though a lack of long-term monitoring data for water depth and velocity at the study sites is a source of uncertainty, we were able to overcome uncertainties by surveying diverse environmental conditions, varying from tributaries to rivers.

Buchan and Padilla (2000) found that the most important factors affecting the presence–absence of submerged macrophytes were associated with water quality rather than with dispersal potential. Nonetheless, to improve the predictive accuracy, it is also necessary to consider physical factors, such

<table>
<thead>
<tr>
<th>Species</th>
<th>Contents</th>
<th>Predicted absence</th>
<th>Predicted presence</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. spicatum</em></td>
<td>Observed absence</td>
<td>23</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Observed presence</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>27</td>
<td>14</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Correct prediction</td>
<td>(23 + 8)/41 = 0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of commission</td>
<td>6/29 = 0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of omission</td>
<td>4/12 = 0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kappa</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>H. verticillata</em></td>
<td>Observed absence</td>
<td>28</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Observed presence</td>
<td>8</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>36</td>
<td>5</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Correct prediction</td>
<td>(28 + 3)/41 = 0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of commission</td>
<td>2/10 = 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error of omission</td>
<td>8/11 = 0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUC</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kappa</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Comparison of predicted and observed distributions of *Myriophyllum spicatum* and *Hydrilla verticillata* at the model confirmation stage. Predicted values were obtained from the fitted probability of presence using the Youden index to apply a threshold probability: 0.536 for *M. spicatum* and 0.492 for *H. verticillata*. 

as flow regime, channel connectivity, channel slope, channel bed, shoreline conditions, and land cover type in the basin (Buchan and Padilla, 2000; Patrick et al., 2014). Biological factors, including competition, herbivory, and disease, are also important habitat determinants (Lacoul and Freedman, 2006). No habitat suitability model is a complete representation of reality and these models should be validated for applications using real-world data by predictive performance evaluations focusing on the reduction of omission errors (Liu et al., 2009; Gastón and García-Viñas, 2013). The distribution of M. spicatum and H. verticillata with high invasiveness could not be compared before and after water regulation in South Korea; however, we are able to predict the distribution using water quality data and adapt the models throughout the world.

4.3 Perspectives

We inferred the habitat characteristics of M. spicatum and H. verticillata using GAMs based on field survey data at the catchment scale. We observed that water chemicals, e.g., chlorophyll a, suspended solids, nitrate nitrogen, water temperature, and electrical conductivity are important factors determining the occurrences of submerged macrophytes. Moreover, we validated the developed models by applying independent field data. Our research has practical implications for the prevention or delay of the aggressive spread of M. spicatum and H. verticillata by providing a basis for river management strategies, such as information about water chemicals to improve water quality in priority areas (Barko et al., 1986). These results are also helpful to sustain aquatic ecosystem functions and biodiversity in regulated hydrological conditions by identifying priority areas for monitoring and management.

Supplementary Material

Supplementary Tables.
The Supplementary Material is available at https://www.kmae journal.org/10.1051/kmae2017044/olm.

References


O’Hare JM, O’Hare MT, Gurnell AM, Dunbar MJ, Scarlett PM, Laizé C. 2011. Physical constraints on the distribution of macrophytes


Solanki HU, Bhatpuria D, Chauhan P. 2016. Applications of generalized additive model (GAM) to satellite-derived variables and fishery data for prediction of fishery resources distributions in the Arabian Sea. Geocarto Int 32: 30–43.


