Ecological characterization of habitats colonized by the freshwater gastropod *Viviparus contectus* (MILLET, 1813) (Gastropoda, Prosobranchia) – Theoretical and experimental data

Robert STURM

**Abstract:** Ecological evaluation of the freshwater gastropod *Viviparus contectus* was carried out by application of logistic regression, where the probability of occurrence $p(x)$ of an organism is described as a function of a single environmental variable by the use of presence/absence data. In this contribution, the technique was applied to 2562 malacological data, and regression models were computed for eight environmental variables, respectively. Main parameters calculated with the mathematical concept included the maximum probability of occurrence $p_{\text{max}}$, the optimum range with $p(x)/p_{\text{max}} > 0.75$ as well as the reduction in deviance $R$ (0-100%) indicating the goodness-of-fit of the regression. The correlation between probability of occurrence and environmental variable could be uniformly expressed by symmetric, bell-shaped regression curves. Reduction in deviance commonly varied between 1.43% (BOD5) and 26.05% (content of calcium carbonate). With regard to temperature, pH, nitrate content in the water and BODs *V. contectus* exhibits slight reminiscences of specialization. Concerning the other variables, the snail seems to be marked by a more generalistic behaviour.

**Keywords:** Logistic regression, mathematical modeling, habitat model, *Viviparus contectus*, freshwater gastropod, ecology.

**Introduction**

The significance of freshwater molluscs acting as bioindicators for the evaluation of water quality has been underlined by numerous scientific studies during the past decades (GLÖER & MEIER-BROOK 2003, STURM 2005a, 2007, 2012, 2013, 2016, 2018). Currently, treatment of this essential question represents a fundamental component of aquatic ecology. In general, diversity of a community of aquatic molluscs is determined by a great many physical, chemical, and biological parameters, for what reason it may be regarded as temporally and locally confined specificity of a given ecosystem (e.g., FALKNER 1990, BAADE 1993, STURM 2005b, 2007, 2012). Concerning gastropods and bivalves colonizing freshwater habitats, species-related thresholds of tolerance with respect to abiotic and biotic parameters are commonly characterized by remarkable discrepancies. In the past, this circumstance resulted in a classification of this group of organisms into generalists and specialists (e.g., GLÖER & Meier-BROOK 2003, STURM 2012). Representatives of the first category exhibit enhanced tolerance to natural or artificial fluctuations of diverse ecological conditions, whereas representatives of the second category quickly respond to any physical and chemical changes of their environment, by means of which their bioindicative value
is additionally increased. Detailed knowledge on the habitat requirements of such specialists may be important in several respects: First, it provides useful information for an effective evaluation of water quality (Statzner & Sperling 1993) and, second, it represents a significant foundation for various renaturalization projects (Osborne et al. 1993). Although the distribution of freshwater molluscs colonizing aquatic habitats in Central Europe has been subjected to a comprehensive description in the meantime (Turner et al. 1998, Glöer & Meier-Brook 2003), ecological demands of single species on their habitats as well as specific preferences with regard to physical and chemical parameters have been elucidated very marginally.

Regression models present as indispensable tools supporting the determination of ecological demands raised by single species (Hosmer & Lemeshow 1989). An appropriate mathematical approach to this problem is provided by logistic regression which describes the probability of local occurrence of a given species as function of one or more independent environmental variables (Jongman et al. 1987, Peeters & Gardeniers 1998, Sturm 2004, 2005b). This procedure has already been applied to a high number of ecological problems (e.g., Ter Braak & Van Dam 1989, Odland et al. 1995). Its enlarged use in limnology and freshwater zoology may be registered since about 20 years (e.g., Eyre et al. 1993, Peeters & Gardeniers 1998, Groger 2000, Field et al. 2002, Sturm 2005b, 2007, 2012, 2016).

Main objective of the present contribution consists in an extension of our knowledge with regard to freshwater mollusc ecology by modeling the habitat requirements of the snail *Viviparus contectus* (Millet, 1813) which represents one of the largest gastropods colonizing home waters. For this purpose, the logistic regression technique was applied to published malacological datasets that included information on both the geographic distribution of the aquatic gastropod and site-related physical and chemical parameters. Validation of the generated regression models was conducted by comparing theoretical predictions with results of malacological field investigations that had been set about between 1998 and 2006.

**Materials and Methods**

**Short characterization of the genus *Viviparus* and available malacological data**

According to the scientific reviews of Glöer & Meier-Brook (2003) and Turner et al. (1998) gastropods belonging to the family Viviparidae are marked by both separate sexuality and viviparous way of life. Besides the pond snail *Lymnaea stagnalis* (Linnaeus, 1758) they can be assigned to the largest indigenous freshwater snails. In the open literature the four species *Viviparus contectus* (Millet, 1813; Fig. 1), *V. viviparus* (Linnaeus, 1758), *V. ater* (Christofori & Jan, 1832) and *V. acerosus* (Bourguignat, 1862) are described which exhibit different distributions in Central Europe (Glöer & Meier-Brook 2003). From an ecological point of view, they typically live in waters with moderate turbulence and neutral to slightly acidic pH value. Concerning *V. contectus* the spectrum of colonized habitats commonly ranges from ponds covered with dense vegetation to swamps, moor pools and side branches of rivers. Thereby, size and colour of the shells varies according to the supply with CaCO3 and nutritive substances.

In Austria *V. contectus* could be detected at 154 of 2562 sample points published in the
scientific literature thus far. For the generation of appropriate regression models all malacological data points listed in the literature were used. For most sample points collection of chemical and physical water parameters has to be regarded as incomplete, so that only eight environmental variables could be implemented into model development.

**Fig. 1**: Stereoscopic photographs showing the front and back of the shell of *V. contectus* with its typical shape and mouth geometry. The height of the shells measures about 4.5 cm.

**Basic concepts of logistic regression**

In principle, the mathematical concept of logistic regression allows the precise analysis of the relationship between a dependent variable of binary character (e.g., 0 or 1) and one or more independent parameters (*Ter Braak & Looman* 1986, *Hosmer & Lemeshow*...
A special case of this method is represented by the so-called "presence/absence curve", where the probability of local occurrence of a given species, \( p(x) \), is simply expressed as function of a certain environmental variable \( x \). In general, \( p(x) \) can be written as

\[
p(x) = \frac{e^{(\beta_0 + \beta_1 x + \beta_2 x^2)}}{1 + e^{(\beta_0 + \beta_1 x + \beta_2 x^2)}},
\]

where \( \beta_0, \beta_1 \) and \( \beta_2 \), respectively, denote the regression coefficients with \( \beta_0 \) acting as constant term. For \( \beta_0 \neq 0 \) the function noted above describes a symmetric, bell-shaped curve (Gaussian logit-curve), whereas for \( \beta_0 = 0 \) \( p(x) \) produces a graph with sigmoid increase or decrease. The optimal range of a certain environmental variable with high probability of occurrence of a species can be determined by the parameters \( u \) and \( t \) (JONGMAN et al. 1987) in the following way:

\[
u = \frac{\beta_1}{\sqrt{-2\beta_2}}
\]

and

\[
t = \pm \frac{1}{\sqrt{2\beta_2}}.
\]

The parameter \( u \) defines the position of the mode of the probability function with \( p'(x) = 0 \), whereas the parameter \( t \) represents the range of the independent variable between the two turning-points \( (p''(x) = 0) \) of the function. Within the interval \( u \pm t \) the quotient of local and maximal probability of occurrence \( (p(x)/p_{\text{max}}) \) adopts values between 0.75 and 1.00 (PEETERS & GARDENIERS 1998).

Statistical significance of the regression coefficients \( \beta_1 \) and \( \beta_2 \) was evaluated by application of the so-called likelihood-ratio test (HOSMER & LEMESHOW 1989), where the predictive power of the normal regression model (Eq. (1)) and the null model \( (\beta_1 = \beta_2 = 0) \) is subjected to a comparison. If the difference between both approaches is negligible, the respective coefficients are eliminated from the model. Goodness-of-fit of the regression model was determined with the help of the parameter \( R \) which is defined by the equation

\[
R = \left( 1 - \frac{D_1}{D_0} \right) \cdot 100\%.
\]

In Eq. (4) the parameter \( D_1 \) denotes the deviation of the adapted model (incl. \( \beta_1 \) and \( \beta_2 \)) from binary data, whereas \( D_0 \) represents the respective deviation of the null model (PEETERS & GARDENIERS 1998). High values of \( R \), indicating good adaptation of \( p(x) \) to the binary data, results from low values of \( D_1 \) and vice versa.

For computation of the regression models the presence \( (= 1) \) and absence \( (= 0) \) of \( V.\ contectus \) was tested for all malacological data and plotted against a certain environmental variable. Respective calculations according to the formulae noted above were conducted with the help of specific work-sheets programmed in MS-Excel®. The regression coefficients \( \beta_0, \beta_1 \) and \( \beta_2 \) were estimated by application of a self-programmed iteration procedure.
Results

Results of descriptive statistics are summarized in Fig. 2. Statistical parameters such as median, minimum, maximum, 25%-quartile and 75%-quartile were determined for both the entirety of malacological data and the sample points with occurrence of *V. contectus*. Comparison between respective box-plots being based on all data and those being based on *Viviparus*-related data clearly indicate that especially for water temperature, pH value, electric conductivity, CaCO$_3$ content and total hardness significant differences of the ranges are noticeable. Whilst the medians of the pH values are very similar, the medians of water temperature, nitrate content and BOD$_5$ partly exhibit remarkable discrepancies. With regard to the remaining environmental variables, only smaller differences between the ranges of measured values and medians are recognizable.

**Fig. 2 a-h:** Box-plots for the statistical evaluation of single environmental variables under incorporation of all malacological data available in the scientific literature and those data with occurrence of *V. contectus*, respectively. The black boxes range from the first to the third quartile, whereas the ends of the lines mark the minimum and the maximum of the data. The white line indicates the position of the median.
For all environmental variables described above a successful adaptation of the logistic regression models to the presence/absence data could be obtained (Tab. 1, Fig. 3). With the help of likelihood-ratio tests significance of the regression coefficients $\beta_1$ and $\beta_2$ ($p < 0.01$) was established in all cases, wherefore the produced functions exhibit uniform symmetry and bell shape (Gaussian logit-curves). Goodness-of-fit, however, is subject to

![Logistic regression models of the single environmental variables for the presentation of eventual habitat preferences of *V. contectus*. For a validation of the models experimental data from diverse field studies were used (e.g., PATZNER & ISARCH 1999, STURM 2000a).](https://www.zobodat.at)
remarkable variations which is above all reflected by respective oscillations of the parameter \( R \) adopting values between 1.43% in the case of BOD\(_5\) and 26.05% in the case of CaCO\(_3\) content in the water (Tab. 1). The regression curves determined for the single environmental variables exhibit discrepancies with regard to their maxima (\( p_{\text{max}} \)) which range from 0.30 to 0.55. In addition, they are characterized by significant variations concerning their ratio of half-width and total data range. Whilst the logit-curves for water temperature, pH value and O\(_2\) content are distinguished by a comparatively narrow peak, for the nitrate content and BOD\(_5\) the model produced regression functions with large half-widths. In Tab. 1 the optimal living conditions of \( V. \text{contectus} \), which are defined by the interval \( u \pm t \), are listed for all environmental variables. Within this range the quotient \( p(x)/p_{\text{max}} \) commonly adopts values > 0.75. In other words, the local probability of occurrence of the gastropod at least amounts to 75% of the maximal probability of occurrence.

For a validation of the regression models experimental data originating from diverse field studies (e.g., PATZNER & ISARCH 1999, STURM 2000) were additionally plotted into the diagrams of Fig. 3. Local probabilities of occurrence were determined by dividing the number of sample points with presence of \( V. \text{contectus} \) by the total number of sample points given for a studied lake or pond. In the case of water temperature, CaCO\(_3\) content and O\(_2\) content good adaptation of the data points by the regression curves can be observed. For the remaining variables partly larger differences between model and experimental data are noticeable.

**Tab. 1: Results of logistic regression modeling.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \beta_0 )</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( R(%) )</th>
<th>( u )</th>
<th>( t )</th>
<th>opt. Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>temperature (°C)</td>
<td>-22.03</td>
<td>2.37</td>
<td>-0.04019</td>
<td>17.58</td>
<td>20.09</td>
<td>2.98</td>
<td>17.11-23.7</td>
</tr>
<tr>
<td>pH value</td>
<td>-115.03</td>
<td>31.30</td>
<td>-336.788</td>
<td>25.76</td>
<td>6.91</td>
<td>0.33</td>
<td>6.58-7.24</td>
</tr>
<tr>
<td>elect. conduct. (µS/cm)</td>
<td>-4.05</td>
<td>0.02</td>
<td>-0.00004</td>
<td>24.87</td>
<td>262.55</td>
<td>52.46</td>
<td>210.09-315.01</td>
</tr>
<tr>
<td>CaCO(_3) content (mg/L)</td>
<td>-3.22</td>
<td>0.04</td>
<td>-0.00019</td>
<td>32.82</td>
<td>119.71</td>
<td>28.48</td>
<td>91.23-148.19</td>
</tr>
<tr>
<td>total hardness (mg/L)</td>
<td>-2.56</td>
<td>0.04</td>
<td>-0.00014</td>
<td>29.17</td>
<td>141.99</td>
<td>34.77</td>
<td>107.22-176.76</td>
</tr>
<tr>
<td>O(_2) content (mg/L)</td>
<td>-24.53</td>
<td>5.52</td>
<td>-0.31491</td>
<td>2.99</td>
<td>8.76</td>
<td>1.26</td>
<td>7.5-10.02</td>
</tr>
<tr>
<td>nitrate content (mg/L)</td>
<td>-1.64</td>
<td>0.80</td>
<td>-0.29511</td>
<td>3.11</td>
<td>1.35</td>
<td>1.30</td>
<td>0.05-2.65</td>
</tr>
<tr>
<td>BOD(_5) (mg/L)</td>
<td>-1.44</td>
<td>0.92</td>
<td>-0.24251</td>
<td>1.43</td>
<td>1.96</td>
<td>1.44</td>
<td>0.52-3.4</td>
</tr>
</tbody>
</table>

**Discussion**

Based on the mathematical concept of logistic regression it could be demonstrated that distribution of the Central-European freshwater gastropod *Viviparus contectus* is partly controlled by the environmental conditions being prevalent in a given ecosystem. The computed regression models unanimously revealed a relationship between the occurrence of probability of \( V. \text{contectus} \) and the considered chemical or physical variable, whereby symmetric and bell-shaped functions were produced (Fig. 3). Best correspondence between regression curves on the one hand and malacological data published in the literature on the other could be found in descending order for the CaCO\(_3\) content in water, total hardness, electric conductivity, water temperature and the pH value with respective values for \( R \) ranging from 15 to 26% (Tab. 1). For the remaining environmental variables computed values for \( R \) varied between 1.4 and 3% which corresponded to low goodness-of-fit.
of the related regression models. With regard to water temperature, pH value, nitrate content and BOD$_5$ \textit{V. contectus} shows rather narrow tolerance intervals compared to the total range of the global data and thus tends to develop a more specialistic behaviour.

According to GLÖER & MEIER-BROOK (2003) species belonging to the Central-European genus \textit{Viviparus} are commonly characterized by low tolerance with respect to the pH value, thereby preferring slightly acidic conditions. \textit{V. contectus} preferably occurs in moor pools and isolated side branches of rivers, where pH values < 7.0 are measured very frequently (PATZNER & ISARCH 1999, STURM 2000). Concerning the BOD$_5$, for which the preference interval ranges from 0.52 to 3.4 mg/L (Tab. 1), the significance of the snail as indicator for clean waters (quality class I-II, BAUR 1998) is additionally underlined. The circumstance is also confirmed by the enhanced sensitiveness of the gastropod with respect to the nitrate concentration in the water. Preference of \textit{V. contectus} for higher water temperatures (optimal range: 17-23 °C) is largely consistent with field data, although this ecological specificity has not been outlined in literature hitherto.

The unpretentiousness with respect to the amount of ions dissolved in the colonized water is largely supported for \textit{V. contectus} and \textit{V. viviparus} (GLÖER & MEIER-BROOK 2003), because both species tolerate a maximal salt concentration in the water of 0.3 to 0.4%. Any decrease of the CaCO$_3$ content does not inevitably result in a diminution of the population density, but frequently leads to a considerable corrosion of the shell with increased mortality (Ribi et al. 1986). Regarding the remaining results confirming or disproving literature data are not available so far, because main interest in association with the gastropod was among other focused on reproduction (e.g., GEBHARDT & RIBI 1987) and insertions of heavy metals in various tissues (e.g., BARSYTE-LOVEJOY 1999).

Specialists commonly exhibit low tolerances with respect to any modifications of their environment and therefore tend to colonize ecological niches, where prevailing environmental conditions may be regarded as constant over longer periods of time. Adaptation of organisms to highly specific environmental conditions was described by numerous theoretical models in the past and thus represents a central topic in modern ecology (e.g., REUTTER et al. 2003, STURM 1998, 2000b, 2003, 2004, 2005a, 2012, 2016, 2018). According the results of the present contribution \textit{V. contectus} only seems to develop a high grade of specialization with respect to water temperature, the content of acids and nitrates in the water and the biological oxygen demand. With regard to the other environmental variables the species may be classified as generalist. However, the highly limited number of chemical and physical variables provides a rather restricted ecological picture of the gastropod and thus should be enlarged in near future.

\textbf{Zusammenfassung}

Ökologische Charakterisierung von Habitaten der Süßwasserschnecke \textit{Viviparus contectus} (MILLET, 1813) (Gastropoda: Prosobranchia) – Theoretische und experimentelle Daten. Die ökologische Bewertung der Süßwasserschnecke \textit{Viviparus contectus} wurde unter Anwendung der logistischen Regression durchgeführt, bei welcher die Auftrittswahrscheinlichkeit eines Organismus als Funktion einer einzelnen Umweltvariablen auf Basis von Präsenz-/Absenzdaten beschrieben wird. Im vorliegenden Beitrag wurde das mathematische Verfahren unter Einbindung von insgesamt 2562 malakologischen Datensätzen ausgeführt, wobei entsprechenende Regressionsmodelle für acht Umweltvariablen erstellt wurden. Die aus der logistischen Regression gewonnenen Parameter umfassten unter anderem die maximale Auftrittswahrscheinlichkeit $p_{\text{max}}$, die optimale Spannweite einer Variablen mit
$p(x)/p_{\text{max}} > 0,75$ und die Anpassungsgüte $R$ (0-100 %). Die Korrelation zwischen Auftrittswahrscheinlichkeit und Umweltvariable konnte auf einheitliche Weise durch symmetrische, glockenförmige Regressionskurven zur Darstellung gebracht werden. Die Anpassungsgüte varierte zwischen 1,43 % (BSB5) und 26,05 % (Carbonatgehalt im Wasser). In Bezug auf die Wassertemperatur, den pH-Wert, den Nitratgehalt im Wasser und den BSB5 zeigt $V. \text{contectus}$ leichte Anklänge einer Spezialisierung. Hinsichtlich der übrigen Umweltvariablen scheint die Schnecke eher durch generalistisches Verhalten gekennzeichnet zu sein.

Literatur


