Development of empirical standard weight equation for Pursak chub *Squalius pursakensis*, an endemic cyprinid species of Northwestern Anatolia

Şule Sülün¹, Sercan Başkurt¹, Özgür Emiroğlu¹, Daniela Giannetto²*, Ali Serhan Tarkan², Sevan Ağdamar², Özcan Gaygusuz³, Zeynep Dorak³, Hamdi Aydın⁴, Arzu Çiçek⁵

¹Eskişehir Osmangazi Üniversitesi, Fen Fakültesi, Biyoloji Bölümü, Eskişehir, Turkey.
²Muğla Sıtkı Koçman Üniversitesi Su Ürünleri Fakültesi 48000 Kotekli/Muğla, Turkey.
³İstanbul Üniversitesi, Su Ürünleri Fakültesi, 34470 Laleli, Fatih, Istanbul, Turkey.
⁴Kocaeli Üniversitesi, Gazanfer Bilge Meslek Yüksekokulu, 41500, Karamursel, Kocaeli, Turkey.
⁵Çeve Sorunları Uygulama ve Araştırma Merkezi, Anadolu Üniversitesi, Eskişehir, Turkey.

*Corresponding author: email: danielagiannetto@gmail.com

Number of pages: 25
Number of Tables: 1
Number of Figures: 4
Abstract

Indices of condition enable the evaluation of the well-being of fish, with the assumption that heavier fish of a given length are in better condition. Relative weight ($W_r$) is one of these indices and it is calculated by comparing the actual weight of a specimen with the ideal weight of a specimen in the same length of the same species in good physiological condition, i.e. standard weight ($W_s$). In this research, over the distribution range length and weight data on Pursak chub *Squalius pursakensis*, an endemic species distributed in Sakarya and Porsuk drainages in Northwestern Anatolia (Turkey), were used to compute a $W_s$ equation by means of Empirical Percentile (EmP) method. The $W_s$ equation obtained was: $\log_{10} W_s = -4.657 + 2.614 \log_{10} TL + 0.127 (\log_{10} TL)^2$ and the total length range of application was 80 - 340 mm. Since EmP $W_s$ equation was not influenced by length variation, the use of this equation to compute the relative weight ($W_r$) for *S. pursakensis* throughout its area of distribution is suggested.

Key words: condition indexes; endemic species, *Squalius*, relative weight.
1. Introduction

Indices of conditions are commonly used by fisheries to evaluate the well-being of fish starting by the assumption that “fatter is fitter” (sensu Glazier, 2000). Then, comparing specimens of the same length, fish with a higher weight are in better condition than those with a lower weight (Anderson and Neumann, 1996; Blackwell et al., 2000). These indexes are important tools for fisheries management (Anderson and Neumann, 1996; Blackwell et al., 2000) and have been used in fisheries research since the beginning of the 20th century (Froese, 2006). According to Fechhelm et al. (1995), the use of indices of condition to monitor the well-being of fish is cost-effective because (being based only on length and weight measurements) they do not require the sacrifice of the specimens and allow to evaluate condition of fish with a minimal mortality.

Relative weight ($W_r$) (Wege and Anderson, 1978) is one of these indices and, over the others available in literature [i.e. Fulton’s (1911) condition factor and Le Cren’s (1951) condition factor], it has the great advantages to be not influenced by changes in body shape and to allow to compare the condition of fish of different lengths and belonging to different populations (Murphy et al., 1991). Relative weight is based on the comparison between the actual weight of a specimen and the standard weight ($W_s$), which is the weight in the same length of an ideal fish of the same species, which is in good physiological condition (Murphy et al., 1991). $W_s$ is predicted by a standard weight equation, that is a length-weight regression typical of the species (Wege and Anderson, 1978). Since its development, the $W_r$ index has been widely used to perform condition analysis of many species (Blackwell et al., 2000). However, its applicability is limited by the lack of species-specific standard weight equations that have to be developed.
using a wide sample of specimens collected throughout the area of distribution of each
species (Bister et al., 2000).

Pursak chub *Squalius pursakensis* (Hankó, 1925) is a cyprinid species endemic of the
Sakarya and Porsuk drainages (Turkey) which flow into the Black Sea (Northwestern
Anatolia) (Özulug and Freyhof, 2011) (Figure 1). The species, originally described as
*Leuciscus orientalis pursakensis* Hankó, 1925, was recently rehabilitated as a valid
species by Özulug and Freyhof (2011). On the basis of molecular differences found in
chubs from Western Anatolia, Durand et al. (2000) suggested that the group of “short
snouted chubs”, before identified as a subspecies of *Squalius cephalus* (Linnaeus,
1758), could be divided into different species. Recently, Özulug and Freyhof (2011)
confirmed this hypothesis and identified the presence of 10 different species of the
genus *Squalius* in the Western and Central Anatolia (4 of them only recently described)
including *S. pursakensis*.

This species is consumed by local people of the basin and used for recreational purposes
(Erk’akan, 1981; Ekmekçi, 1996; Emiroğlu et al., 2001). As both basins (Sakarya and
Porsuk) where *S. pursakensis* occurs have recently been severely polluted and affected
by serious habitat destructions mainly damming and water abstraction, the species has
been suffered from these changes (Bostancı and Polat, 2009; İnnal, 2010).

Two different methods were proposed in the literature for the computation of Wₘ equation: the Regression Line percentile (RLP) method, proposed by Murphy et al.
(1990), and the Empirical Percentile (EmP) method, developed by Gerow et al. (2005).
Until recently, the most widely used method was the RLP (Blackwell et al., 2000).
However, according to Gerow et al. (2004), some of the Wₘ equations developed using
the RLP method showed significant length-related biases and, moreover, the results of
recent studies encouraged the use of the EmP method to develop $W_s$ equations (Rennie and Verdon, 2008; Ogle and Winfield, 2009; Gerow, 2010; Giannetto et al., 2011; 2012a).

The aim of this research was to develop standard weight equation by means of the EmP method for *S. pursakensis*, which could allow estimating the relative weight throughout the entire area of distribution of this species. Given the species under threat of habitat disturbance and biotic degradations by non-native species, which are commonly increasing in the area, proposed standard weight equation would serve an important tool for conservation of this endemic species in its restricted distribution area.

2. Materials and Methods

2.1. Dataset selection

Specimens of *S. pursakensis* were collected throughout the area of distribution of the species (Figure 2) by means of electrofishing (SAMUS 725G) between 2009 and 2012. After collection, each fish was measured for length (TL, SL, FL) to the nearest mm and wet weight ($W$) to the nearest 0.1 g in situ, with fish being returned to the water.

According to Giannetto et al. (2011, 2012b) the following methods were used to validate the total dataset. Firstly, by the TL-$W$ regression of the total sample, all specimens that were large outliers were excluded, as they were probably derived from wrong measurements. Then, the dataset was separated into statistical populations on the basis of the data and location of collection (Ogle and Winfield, 2009). In this regard, a stock (population) was regarded as different according to sampling years and periods from the same location because of emigration/immigration, recruitment and mortality (Begg and Waldman, 1999). Populations with less than 10 specimens were removed.
from the dataset (Ogle and Winfield, 2009; Lorenzoni et al., 2012). Hence, in order to identify all anomalous values, data were further validated by plotting the $\log_{10} TL - \log_{10} W$ regression for each population separately; all values that diverged more than double the expected value by the regression curve were removed (Bister et al., 2000). These equations were then analyzed and all populations showing an $R^2$ value less than 0.90 or for which a value of the slope ($b$) fell outside the range of 2.5 - 3.5 were excluded from further analyses (Froese 2006).

The last step for the validation was to plot the slopes ($b$) of all populations against all intercepts ($\log_{10} a$) (Pope et al., 1995) to identify the outliers of this regression represented by populations composed by few fish or samples with a narrow length-range (Froese, 2006).

2.2. Determination of the applicable total length range for the $W_s$ equation

Once the dataset was validated, the next step was to determine a suitable length-range for the application of the $W_s$ equation. A minimum length is requested because small fishes show a high variance due to the differences in growth forms that arise in the juvenile stages and to the potential error associated with their measurement in the field (Murphy et al., 1990; Rypel and Richter, 2008). In accordance with Willis et al. (1991), the minimum total length was determinate by plotting the variance/mean ratio for $\log_{10} W$ on 10-mm total length intervals and the minimum length was determined as the size at which the value of the ratio was less than 0.01 (Murphy et al., 1990).

Since the EmP method was utilized to develop the $W_s$ equation, a maximum length for the application of that equation was also required (Gerow et al., 2005). This length was determined as the length-class for which at least three fish populations were present in
the dataset because three represents the smallest sample size that allows estimation of quartiles (Gerow et al., 2005). All fish outside the suitable length range were not utilized for further analyses.

Next the dataset, thus validated, was divided into two sets: a larger development dataset (used to compute the $W_s$ equation) and a smaller validation dataset (used to assess potential length-related biases in the $W_s$ equation) (Rypel and Richter 2008; Ogle and Winfield, 2009; Giannetto et al., 2012b).

### 2.3. Development of $W_s$ equation

The EmP method proposed by Gerow et al. (2005) was used to develop the $W_s$ equation for *S. pursakensis*. The mean empirical $W$ for each 10-mm length-group was calculated by the log$_{10}$ transformed TL and $W$ of each population of the development dataset; the $3^{rd}$ quartiles of these mean empirical $W$ for each length-group were then regressed against TL to develop the EmP $W_s$ equation by using a weighted quadratic model (Gerow et al., 2005).

Then, using the $W_s$ equation obtained, the $W_r$ of each specimen from each population was calculated by means of the equation provided by Wege and Anderson (1978): $W_r = 100 \left(\frac{W}{W_s}\right)$ where $W$ is the weight of an individual in grams and $W_s$ is the standard weight predicted by the $W_s$ equation.

### 2.4. Validation of the EmP $W_s$ equation

The principal attribute of a good condition index is that, in order to enable accurate comparisons of condition assessments of fish of different size, the measure of the condition should be free from length-related biases (Murphy et al., 1991; Anderson and...
To this aim, three different methods were used to detect potential length-related biases and to validate the EmP Wₚ equation calculated for *S. pursakensis*: the analysis of the residuals of the Wₚ equation (to investigate whether the distribution of residuals exhibits evident patterns) (Ogle and Winfield, 2009; Giannetto et al., 2012a); the Willis method (Willis et al., 1991) (whereby a chi-square test was used to determine whether the proportion of populations with a significant positive slope in the TL-Wᵣ equation is equal to the proportion of those with a significant negative slope); the Empirical Quartiles (EmpQ) method (Gerow et al., 2004) as modified by Ogle and Winfield (2009) - applied to the validation dataset by using the FSA package (Ogle, 2009) of R Software - (to determine if the slope of the quadratic regression of the 3rd quartile of the mean W standardized by Wₛ against length intervals of 10-mm had a value of zero) (Ogle and Winfield, 2009; Giannetto et al., 2011; 2012a, 2012b).

3. Results

3.1. Dataset selection

A total of 2590 *S. pursakensis* were collected during the research. The total length of specimens collected ranged from a minimum of 26.000 mm to a maximum of 445.000 mm (mean length 152.562 ± 51.348 mm) and the weight from a minimum of 0.200 g to a maximum of 1263 g (mean weight 62.174 ± 82.197 g). A new maximum total length (445 mm) for this species was recognized.

The log-transformed TL - W equation calculated on the total sample was:

\[
\log_{10}(W) = -5.304 + 3.176 \log_{10} TL (\text{mm}) \ (R^2 = 0.985, p < 0.001; n = 2590)
\]

The SL-TL and FL-TL equations resulted:
TL = 5.724 + 1.157 SL (R² = 0.997; n = 1684)

and

TL = 3.225 + 1.053 FL (R² = 0.998; n = 1684)

The total dataset was then divided into 78 statistical populations. By the linear plot between log_{10} (a) - b, no population were identified as outlier and, for all 78, the value of R² was > 0.95 and the b value resulted within the range 2.5 - 3.5.

3.2. Determination of the applicable total length range for the Wₚ equation

The minimum TL for the application of the Wₚ equation was determined as 80 mm (Figure 3). The maximum TL was assigned to be 340 mm because, despite the presence in the dataset of fish longer than this size, 340 mm was the length at which at least 3 values were present (Table). According to this the length-range judged to be suitable for the EmP Wₚ equation was 80 - 340 mm.

The total sample was, then, divided into two dataset: a development data set made up of 2210 specimens and a validation data set of 380 fish.

3.3. Development of Wₚ equation

By applying the EmP method to the development dataset, the Wₚ equation for S. pursakensis resulted:

\[
\log_{10} W_p = -4.657 + 2.614 \log_{10} TL + 0.127 (\log_{10} TL)^2 \quad (R^2 = 0.999, p < 0.001).
\]

3.4. Validation of the EmP Wₚ equation

The residuals values of the EmP Wₚ equation showed a random distribution and did not exhibit evident patterns (Figure 4). Applying the Willis method, the number of TL-W
relationships with significant positive slopes (12) resulted not significantly different from those with significant negative slopes (5) to $\chi^2$ test ($\chi^2 = 2.882, p = 0.089$) showing and independence of the $W_r$ values calculated by means of the EmP $W_s$ equation developed by the size of the fish.

Applying the EmpQ method, the value of the slope was not significantly different from zero for both terms of the equation ($p_{\text{quadratic}} = 0.508, p_{\text{linear}} = 0.763$) indicating that the EmP $W_s$ equation was not influenced by fish length (Figure 4).

4. Discussion

In this study, a standard weight equation was developed for *S. pursakensis*, an endemic species of the Sakarya and Porsuk drainages in Northwestern Anatolia (Turkey). Moreover, a new maximum total length (445 mm) for this species was caught from Seydi Suyu in Sakarya Basin. Therefore, it can be argued that the representativeness of the dataset used in the present study has contributed to strengthen the validity of the findings. Maximum total length was previously reported as 392 mm from Sarıyar Reservoir by Ekmekçi (1996). Although it is well known that individuals in populations exposed to high level of fisheries pressure reach relatively smaller lengths (Emiroğlu et al., 2012), it is not the case for *S. pursakensis* as there is no high fisheries pressure on the species except than insignificant effect of local anglers in its distribution range. The finding of large specimens of a species could be incidental or attributed to the inadequate current knowledge on the species. However, given congeneric species of *S. pursakensis* may reach 600 mm SL (*S. cephalus*) (Kottelat and Freyhof, 2007), it can grow above the maximum length found in the present study.
The relative weight is currently widely used to perform condition analysis of many species but its applicability is often limited by the lack of species-specific standard weight equations. With regard to Turkey, the only $W_s$ equations available are those for European perch *Perca fluviatilis* (Giannetto et al., 2012a) and Aegean chub *Squalius fellowesii* (Giannetto et al., 2012b).

Previous studies on *S. pursakensis*, which was cited as mainly *Leuciscus cephalus*, focused on some basic reproduction and growth features including condition indices (Erk’akan, 1981; Ekmekçi, 1996; Emiroğlu et al., 2001; Bostancı and Polat, 2009; Innal, 2010). Interestingly, most of the previous studies were conducted in artificial water bodies (reservoirs and ponds) (e.g. Ekmekçi, 1996; Bostancı and Polat, 2009; Yiğit et al., 2008; Innal, 2010) with the exception of two studies, which were from Sakarya (Erk’akan, 1981) and Porsuk (Emiroğlu et al., 2001) rivers. Yiğit et al. (2008) analyzed different condition indices with the aim of finding best fitting condition factor for *S. pursakensis* and found that the one based on height of fish along with length and weight proposed by Jones et al. (1999) was the best fitting condition factor. However, this factor is not useful for broad comparisons due to the lack of available data (i.e. height is usually not measured) and its uncommon use (i.e. this condition factor should be available as site-specific). Indeed, this was the case for Yiğit et al. (2008), as they could use only Fulton condition factor ($K$) for comparisons with other studies though it had least coefficient factor among all analyzed condition indices. In this context, relative weight is the best alternative to condition factors commonly used in fisheries biology because it provides a single species-specific equation for the species, which allow reliable comparisons among different locations and specimens of different lengths.
Further research is encouraged to extend the use of this methodology to other species with particular attention to those endemic. Turkey is an important area for fish biodiversity being characterized by a high number of endemic species (78 species). Notably, almost half of the endemic species in Turkey are classified as Critically Endangered and 32% as Endangered (Fricke et al., 2007). These species have, often, a very small area of distribution (like S. pursakensis), are rarely studied and currently they are threatened by the presence of an increasing number of non-native species introduced by stocking practices that have become very common in Turkish waters to increase fish production and sport fishing (Innal and Erkakan, 2006; Aydı̇n et al., 2011; Önsoy et al., 2011, Tarkan et al., 2012). For these reasons, all management tools that can assist in conserving the populations of these endemic species such as S. pursakensis would be advantageous to assess the population-level responses to ecosystem disturbance.

The relative weight, together with other population metrics (e.g., age and growth), could allow researchers to increase the basic knowledge on population ecology of S. pursakensis that will be useful to improve its management and conservation status (Murphy et al., 1991; Blackwell et al., 2000). Low values of relative weight will permit, by comparison of data of different periods, to detect any decline in condition of specimens probably due to environmental alterations or biological disturbance (i.e. presence of non-native species) (Giannetto et al., 2012c). The results of previous studies reported clear habitat-based variations in age and growth features for other chubs species such as Aegean chub (Balık et al., 2004; Dirican and Barlas, 2007). These variations are more prominent between populations inhabiting lentic and lotic areas (Şaşı̇ and Balık, 2003; Torcu et al., 2007). Because of the positive correlation between fish growth and environmental quality (Bister et al., 2000), relative weight could be an
easy and powerful tool to identify ecological changes just as the incidence of
phenomena of intra-species competition (Johnson, 1992), impact of non-native species;
Giannetto et al., 2012c) or environmental disturbances (Gabelhouse, 1991; Hubert et al.,
1994; Liao et al., 1995). Considering S. pursakensis is commonly found in reservoirs,
which are considered as unfavourable and variable environments for fish species
(Tarkan, 2007; Emiroğlu et al., 2012), monitoring ecological changes via simple and
effective tools such as relative weight is useful.
The EmP equation developed for S. pursakensis is not biased by fish length
according to all methods used for the validation (Willis, EmpQ and analysis of the
residuals of the equation). The results highlighted the reliability of the EmP method to
develop standard weight equation confirming those obtained by previous studies
(Rennie and Verdon, 2008; Ogle and Winfield, 2009; Giannetto et al., 2011; 2012a,
2012b; Lorenzoni et al; 2012). On the basis of these results, the use of the EmP equation
to determine W, for S. pursakensis throughout its area of distribution is suggested.

Acknowledgements
The study was jointly funded by Anadolu University Research Fund (Project no:
1101F011) and Osmangazi University Research Fund (Project no: 2012/19A113).
The statistical analyses of data of this research were conducted within the 2216 -
Yabancı Uyrukluar için Araştırma Burs Program supported by TÜRKİYE BİLİMSEL
ve TEKNOLOJİK ARAŞTIRMA KURUMU (TÜBITAK).
References


Ogle DH (2009). Data to support Fish Stock Assessment package: Package FSA. GPL version 2 or newer URL (http://www.rforge.net/FSA).


<table>
<thead>
<tr>
<th>TL (mm)</th>
<th>Populations</th>
<th>Individuals</th>
<th>TL (mm)</th>
<th>Populations</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>31</td>
<td>62</td>
<td>270</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>90</td>
<td>35</td>
<td>70</td>
<td>280</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>100</td>
<td>42</td>
<td>105</td>
<td>290</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>110</td>
<td>52</td>
<td>153</td>
<td>300</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>120</td>
<td>67</td>
<td>203</td>
<td>310</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>130</td>
<td>63</td>
<td>209</td>
<td>320</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>140</td>
<td>70</td>
<td>257</td>
<td>330</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>150</td>
<td>59</td>
<td>189</td>
<td>340</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>160</td>
<td>62</td>
<td>184</td>
<td>350</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>170</td>
<td>54</td>
<td>133</td>
<td>360</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>180</td>
<td>49</td>
<td>121</td>
<td>370</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>190</td>
<td>41</td>
<td>93</td>
<td>380</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>200</td>
<td>40</td>
<td>71</td>
<td>390</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>210</td>
<td>39</td>
<td>60</td>
<td>400</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>220</td>
<td>38</td>
<td>68</td>
<td>410</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>230</td>
<td>26</td>
<td>41</td>
<td>420</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>240</td>
<td>20</td>
<td>26</td>
<td>430</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>250</td>
<td>24</td>
<td>27</td>
<td>440</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>260</td>
<td>21</td>
<td>24</td>
<td>450</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table:** Number of populations and individuals for each length class of 10 mm in the development dataset of *Squalius pursakensis.* (The length classes in bold were those removed from the dataset because had less than 3 populations.)
Figure 1. *Squalius pursakensis* from Seydi Suyu in Sakarya Basin (445 mm TL).
Figure 2: Area of distribution of *Squalius pursakensis* (dark grey area).
**Figure 3**: Relationship between variance/mean for log$_{10}$ of weight (W) and total length at 10-mm intervals for the determination of the minimum total length of *Squalius pursakensis*. The dotted line indicates the value of 0.01.
Figure 4: Plots showing the distribution of the residuals (a) and the results of the application of the Empirical Quartiles (EmpQ) method (b) used to investigate potential length-bias in the standard mass ($W_s$) equation for *Squalius pursakensis*. (Residuals = standardized residuals of the regression; Fitted values = values obtained by the model fit; Standardized 75° percentile mean $W = $ standardized 75° percentile mean weights calculated by $W_s$ equation).