Factor Analysis of Associations for Fish Genera in Streams of Doon Valley

Deepali Rana1, Ganesh Datt Bhatta2

Researcher, Department of Zoology, D.B.S. (PG) College, Dehradun, Uttarakhand, India1
Lab Supervisor and Faculty, Department of Petroleum Engineering and Earth Sciences, University of Petroleum & Energy Studies, Dehradun, Uttarakhand, India2

ABSTRACT: The study of fish diversity of Eastern and Western Doon Valley streams (2010-2012) viz., Baldi, Song and Suswa in East and Tons and Asan in West has revealed it to be represented by 56 species belonging to 5 Orders, 13 Families and 30 Genera. The present findings are the outcome of Factor Analysis tool applied on 30 genera, fractionated into 2 categories viz., frequent (28 genera) and infrequent genera/species (2 genera + 12 species), with a view to find out determinants of associations amongst various genera. Four Factor groups (represented by 23 frequent genera + 1 set of infrequent genera/species) were obtained through oblique rotation of principal axes, factored from the number of individual’s data matrix of 480 collections, with heavy loadings of ≥ 0.70. The sake of convenience and easy understanding, various Factors have been named on the basis of genus/species showed the highest loading on each Factor. The distribution of fish genera according to their loading values in different Factors showed a specific pattern of association and vis-a-vis highlights meaningful interpretations about the ecological habitat characteristics of various genera.

KEY WORDS: Doon Valley, Factor Analysis, Fish Genera, Principal Component Analysis, Multivariate Analysis.

I. INTRODUCTION

Important treatises on ecology, incorporating a number of researches has been done from time to time [1]. Study covering various aspects of riverine ecology and its continuum concept has been carried out by Vamnote et al., [2]. Multivariate statistical methods including Factor Analysis (FA) have been used successfully which allow deriving hidden information from the data set about the possible influences of the environment [3]. Ecological applications of multivariate statistics have expanded tremendously [4], using various methods to observe the aspects like species association analyses [5], species area relationships [6]. All these analyses are suggested to be done by applying FA, Principal Component Analysis (PCA), multivariate analysis of variance, Deterred Correspondence Analysis (DCA) and Canonical Correspondence Analysis (CCA) [7-9]. As computational upgradation increased in recent decades, there has been an increase in multivariate assessment of distribution of fish species across large geographical areas [10-12]. Distribution of fish species have been linked statistically to individual water quality variables [13]. The research on fish assemblage influenced by environmental factor, variation between geographically defined regions and spatial and temporal characterization of fish assemblages has been aptly worked out by wide variety of workers outside India [14-22], who variously used the statistical tools i.e., FA, PCA, multivariate analysis of variance, DCA and CCA. Use of statistical tools to derive correlation between fish diversity and water quality, a comparatively less quantum of work has been initiated in India, except a few [23-27].FA is a mathematical technique which groups species according to common pattern of abundance and allows objective assessment of species association and their distribution [18]. This was aptly worked out by Rose and Echelle[18] while evaluating the associations of fishes in Little river of Central Texas. FA was also used to determine generalized pattern and their correlation with certain environmental variables by Stevenson et al., [15] while analysing the distributional pattern of fish species in Western and Central Oklahoma. The present study was conducted using FA to determine associations amongst various fish genera, their distribution patterns, genera/species associations, habitat correlations of observed generic groups and to derive meaningful interpretations regarding the distributional pattern of fishes in the riverine ecosystem of Doon Valley, Uttarakhand, India.
II. MATERIAL AND METHODS

STUDY AREA

Doon Valley, part of district Dehradun (latitude - 29°58' and 30°32' N and longitude - 77°35' and 78°20'E) comprises of 2 main river basins, namely, the Ganga and the Yamuna river basin. The present study was carried out on these two river systems comprising of five main rivers i.e., Baldi, Song, Suswa, Tons and Asan (Fig. 1). The climate of the area varies from humid, moist sub-tropical in the southern part to temperate in the northern mountainous region with wide temperature range varying from 4.40-35.10°C during the study period. The mean minimum and maximum temperature recorded were 11.90°C (January, 2011) and 28.0°C (June, 2011), respectively.

The annual rainfall is highly variable and is mainly controlled by the orography, 82-87 % of the annual rainfall occurs under the influence of the South - West monsoon. Maximum rainfall was recorded during June - September and July, 2010 was the wettest month, receiving about 545.80 mm rainfall. Winter rains were prevalent during December - March, accounting for about 8 % of the total annual rainfall. The relative humidity was also variable. November, 2010 was the least humid (18%) whereas June, 2011 was the most humid (63 %) at 071 hours. Sandy loam, sandy clay, coarse sand, silty clay loam, silty loam, sand and clay types are the chief soil types of Doon Valley. Sampling was periodically done for a period of 24 months (March, 2010 - February, 2012) at the 20 sampled stations established along the rivers mentioned above.

Each river was divided into stretches along its length, according to altitudinal variations to adjudge the spatial and temporal interrelationships. Each stretch covering an approximate distance of about 4 - 7 km, was thus established as sampling sites. Fish samples were collected by employing standard gears, using variety of fishing nets of varying mesh sizes – gill nets, cast nets, drag nets with the help of trained fishermen on the sampling in the Eastern and Western part of Doon Valley, respectively (Fig. 2). Fish samples were preserved in 4% formalin and brought to the laboratory for routine identification, meristic and morphometric analyses under the light of available standard literature and revisionary works [28-34].

The study of fish diversity of Eastern and Western Doon Valley streams i.e., Baldi, Song and Suswa in East and Tons and Asan in West has revealed it to be represented by 56 species belonging to 5 Orders, 13 Families and 30 Genera. For the present study, the original genera × locality data matrix was prepared (by pooling the total number of individuals of a species data, within a particular genus) and was finally subjected to FA using STATISTICA, software [35]. Data was reduced and summarized up to the generic level so as to nullify the error arising out due to the randomness observed in the pattern of distribution of certain species. In the data matrix only 28 genera were considered which were represented by 20 or more individuals (Table 1).

Besides, the genera like Cyprinus and Raiamas having one species each and Puntiusterio, Bariliustileo, Bariliusschacra, Lepidocephalichthysannandalei, Glyptothenaraxtelchitta, Clariasgariepinus, Colisalalalia, Colisalabious, Channamarulus and Chanmaharcourtbutleri have been included under serial number 29 as a set of Infrequent genera/species owing to the fact that though they were collected periodically from certain stations but their number was less than 20 (Table 1). Oblique rotation of principal axes, factored from the number of individual’s data matrix of 480 collections was done [35]. The loadings ≥ 0.70 have been put under the highest loadings (bold scores) (Table 1).

For the sake of convenience and easy understanding various Factors have been named according to the name of genus/species having the highest loading on each Factor [18]. The distributions of Factor scores showed general patterns highlighting meaningful interpretation based on ecological characteristics of various Genera (Table 1).
Fig. 1: Location of the study area.
Fig. 2: Sampled stations and distribution of rivers in Doon valley.
III. RESULTS

Two years field data collection and laboratory observations have revealed that a total of 56 species, belonging to 5 Orders, 13 Families and 30 Genera, represent the fish fauna of all the selected 5 streams of Doon Valley. The FA tool was applied on 30 genera, fractionated in two categories i.e., frequent (28 genera) and infrequent genera/species (2 genera + 12 species), with a view to find out determinants of associations amongst various genera. Out of 29 set of variables, four Factors resolved with loading values ≥ 0.70, represented by 24 genera and one set of infrequent genera/species. Four genera are not resolved under any Factor. Total variance extracted by the four Factors is 77.00% (Table 1).

### Table 1: Factor analysis of fish genera in Doon Valley.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Fish Genera</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Puntius</td>
<td>0.58</td>
<td>0.15</td>
<td>0.03</td>
<td>0.53</td>
</tr>
<tr>
<td>2.</td>
<td>Chagunius</td>
<td>0.37</td>
<td>0.34</td>
<td>0.52</td>
<td>0.49</td>
</tr>
<tr>
<td>3.</td>
<td>Schizothorax</td>
<td>-0.22</td>
<td>-0.83</td>
<td>0.10</td>
<td>-0.14</td>
</tr>
<tr>
<td>4.</td>
<td>Schizothorachthys</td>
<td>-0.14</td>
<td>-0.70</td>
<td>0.29</td>
<td>-0.19</td>
</tr>
<tr>
<td>5.</td>
<td>Labeo</td>
<td>0.03</td>
<td>-0.08</td>
<td>0.79</td>
<td>-0.06</td>
</tr>
<tr>
<td>6.</td>
<td>Aspidoparia</td>
<td>0.46</td>
<td>0.48</td>
<td>0.31</td>
<td>-0.15</td>
</tr>
<tr>
<td>7.</td>
<td>Barilus</td>
<td>0.64</td>
<td>0.10</td>
<td>0.39</td>
<td>0.53</td>
</tr>
<tr>
<td>8.</td>
<td>Danio</td>
<td>0.85</td>
<td>0.12</td>
<td>0.05</td>
<td>0.19</td>
</tr>
<tr>
<td>9.</td>
<td>Devario</td>
<td>0.80</td>
<td>0.25</td>
<td>0.05</td>
<td>0.37</td>
</tr>
<tr>
<td>10.</td>
<td>Esomus</td>
<td>0.85</td>
<td>0.12</td>
<td>0.03</td>
<td>0.19</td>
</tr>
<tr>
<td>11.</td>
<td>Rasbora</td>
<td>0.80</td>
<td>0.25</td>
<td>0.05</td>
<td>0.37</td>
</tr>
<tr>
<td>12.</td>
<td>Crossocheilus</td>
<td>0.10</td>
<td>0.03</td>
<td>0.76</td>
<td>0.47</td>
</tr>
<tr>
<td>13.</td>
<td>Garra</td>
<td>0.15</td>
<td>-0.20</td>
<td>0.82</td>
<td>0.28</td>
</tr>
<tr>
<td>14.</td>
<td>Tor</td>
<td>0.26</td>
<td>-0.29</td>
<td>0.78</td>
<td>0.03</td>
</tr>
<tr>
<td>15.</td>
<td>Lepidocephalichthys</td>
<td>0.77</td>
<td>-0.05</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>16.</td>
<td>Acanthocobitis</td>
<td>0.77</td>
<td>-0.02</td>
<td>0.16</td>
<td>0.28</td>
</tr>
<tr>
<td>17.</td>
<td>Schistura</td>
<td>-0.18</td>
<td>-0.76</td>
<td>-0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>18.</td>
<td>Amblyceps</td>
<td>0.88</td>
<td>0.18</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>19.</td>
<td>Glyptothorax</td>
<td>-0.13</td>
<td>-0.73</td>
<td>0.40</td>
<td>-0.13</td>
</tr>
<tr>
<td>20.</td>
<td>Claris</td>
<td>0.87</td>
<td>0.14</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>21.</td>
<td>Heteropneustes</td>
<td>0.90</td>
<td>0.13</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>22.</td>
<td>Mystus</td>
<td>0.74</td>
<td>0.27</td>
<td>0.18</td>
<td>-0.30</td>
</tr>
<tr>
<td>23.</td>
<td>Xenentodon</td>
<td>0.82</td>
<td>0.25</td>
<td>0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>24.</td>
<td>Macroganthurus</td>
<td>0.85</td>
<td>0.17</td>
<td>0.17</td>
<td>0.22</td>
</tr>
<tr>
<td>25.</td>
<td>Mastacembelus</td>
<td>0.79</td>
<td>0.18</td>
<td>0.12</td>
<td>0.36</td>
</tr>
<tr>
<td>26.</td>
<td>Badis</td>
<td>0.83</td>
<td>0.22</td>
<td>0.06</td>
<td>0.37</td>
</tr>
<tr>
<td>27.</td>
<td>Colisafasciatus</td>
<td>0.29</td>
<td>0.06</td>
<td>0.12</td>
<td>0.81</td>
</tr>
<tr>
<td>28.</td>
<td>Channa</td>
<td>0.78</td>
<td>0.29</td>
<td>0.17</td>
<td>0.40</td>
</tr>
<tr>
<td>29.</td>
<td>Infrequent genera/species</td>
<td>0.50</td>
<td>0.15</td>
<td>0.24</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Explained Variation</td>
<td>11.79</td>
<td>3.39</td>
<td>3.58</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>Total Proportion</td>
<td>0.41</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Bold blue colour indicates the highest loadings values ≥ 0.70’

**Factor 1 (Fossils Group)** - 15 genera has the highest positive loadings on Factor 1 with scores ranging from 0.74-0.90. As per their relative scores, these 15 genera can safely be placed in three categories according to the overlapping of altitudinal ranges.

1. Category - I: 0.90-0.87 (300-500 m)
2. Category - II: 0.85-0.78 (300-500 m)
Category I:
Three genera (*Heteropneustes, Amblyceps* and *Clarias*) appeared most highly correlated with Factor 1. Since, *Heteropneustes* exhibits the highest scores; for the sake of convenience and understanding Factor 1 is labeled as Fossilis Group owing to the fact that swampy, slow moving, plain-type conditions provided the dwelling and feeding ground for such fishes.

Category II:
Comparatively next in order are seven genera of highest scores (0.85-0.78) i.e., *Esomus* (0.85), *Macrognathus* (0.85), *Danio* (0.85), *Badis* (0.83), *Xenentodon* (0.82), *Devario* (0.80) and *Rasbora* (0.80) are more frequent than the earlier three genera and inhabit stretches with feeble flow.

Category III:
The highest positive loadings on Factor 1 are of *Mastacembelus* (0.79), *Channa* (0.78), *Lepidocephalichthys* (0.77), *Acanthocobitis* (0.77) and *Mystus* (0.74). The fish genera falling under Category III indicates less habitat specific and venture the other areas. Simultaneously the score revealing the habitat of Factor 1.

**Factor 2 (Snow - trout Group):** Only four genera had their negative loadings on Factor 2. *Schizothorax* maintained the highest score (-0.83) as compared to the other three i.e., *Schistura* (-0.76), *Glyptothorax* (-0.73) and *Schizothorachthys* (-0.70). The habitat and altitude evaluation of these genera, it becomes evident that this group constitutes the fishes dwelling in high altitude ranges of which *Schizothorax* and *Schistura* are the exclusive cold water, fast running stream dwelling fishes (altitude range from 800 - 1000 m) with adaptations i.e., lips and labial folds modified for adhesion (*Schizothorax*), adhesive apparatus on the chest (*Glyptothorax*), long narrow band-shaped caudal peduncle, reduced eyes, highly placed and reduced external gill openings etc. As far as *Glyptothorax* and *Schizothorachthys* are concerned, though their distributional ranges (500 - 800 m) overlaps with that of *Schizothorax* but their downstream range of distribution may be beyond the lower limit of *Schizothorax*.

**Factor 3 (Garra Group):** Four genera had their highest positive loadings Factor 3 of which *Garra* showed the highest score (0.82). As regards the other three, their scores are not too far different (Labeo 0.79, Tor 0.78 and *Crossocheilus* 0.76) from *Garra* meaning thereby closeness of forming Factor 3, owing mainly to the fact that they have dwelling range from torrential streams of higher altitudes (800 m) through modest stretches (500 - 800 m) to feeble to fast flowing stretches (300-500 m). Characteristics showed all the genera have rheophilic morphological adaptations ranging from sectorial disc at the lower lip of *Garra* to highly indented upper lip, horny covering inside the jaws, highly protractile upper jaw and lips continuous at the angles of mouth, long narrow, band-shaped caudal peduncle (*Crossocheilus, Labeo* and *Tor*).

**Factor 4 (Gourami Group):** Only one Genus *Colisa* (one species *fasciatus*) had the highest loadings (0.81) on the Factor 4, the other highest loadings (0.73) coming on account of the set of genera (*Cyprinus* from *S*; and *Raiamas* from *S*) and species (*Puntiusterio, Bariliustileo, Bariliusschacra, Lepidocephalichthysannandalei, Glyptothoraxtelchita, Clariasgariepinus, Colisalalia, Colisalabiosus, Channamarulius and Channaharcourtbutleri) assigned Infrequent status (Table 1) owing to the fact that they were infrequent in collections (less than 20 individuals and restricted to particular areas of specific sampling stations i.e., majority of them (*Puntiusterio, Bariliustileo, Bariliusschacra, Lepidocephalichthysannandalei, Glyptothoraxtelchita, Clariasgariepinus, Colisalalia, Colisalabiosus, Colisamarulius and Channaharcourtbutleri) are the inhabitants of *S*11 and *S*12 stretches of feeble flowing and marshy Suswa. Of the rest, *Cyprinascarpio* from *S*1 and *Clariasgariepinus* from *S*13 are the two accidental exotics escaped to the natural habitats, but the linking of dwelling into pool-type vegetated habitats. Other contributors to the Factor 4 loadings have been *Bariliustileo* and *Raiamas bola* from the forested, flowing stretch of Song (*S*7 only) and *Bariliusschacra* from similarly characterized stretch of Asan (*S*20). It is also worth mentioning the loadings of Factor 4 (Infrequent set of genera/species) mainly contributed by the new records from the valley i.e., *Puntiusterio* (*S*12), *Bariliustileo* (*S*12), *Bariliusschacra* (*S*12), *Lepidocephalichthysannandalei* (*S*12), *Glyptothoraxtelchita* (*S*12), *Clariasgariepinus* (*S*12), *Colisalalia* (*S*12), *Colisalabiosus* (*S*12), *Channamarulius* (*S*12) and *Channaharcourtbutleri* (*S*12).
Obviously, as per the location of the sampling stretches, all these new records relate to downstream sections of river Suswa characterized partly by protected, forested areas and partly by heavy infestation of aquatic plants along sides and amidst the river.

**IV. DISCUSSION**

The abundance data for fish genera/species to FA has categorically produced Factors which group genera/species according to common pattern of abundance and allows objective assessment of genera/species associations and their distributions [18]. Almost similar to observations by Stevenson et al., [15], the Factors extracted out of FA of fish genera/species showed the Factor groups of various fish distribution and environmental variables and reflect general patterns about the biology of various fish form. The present study has attempted to name the Factors arbitrarily using the common name/specific name of the species showed the highest loading on a Factor [15 - 16, 18]. The high (+ve) correlations have indicated frequent genera/species co-occurrences at particular sampled stations in Eastern Doon valley (Danio, Devario, Esomus, Rasbora, Amblyceps, Clarias, Heteropeustes, Mystus, Macrognathus) from Factor 1; Schizothorax, Schizothorachys, Schistura and Glyptothenorax using Factor 3 and Colisafasciatus and set of infrequent genera/species using Factor 4. High (-ve) correlations indicated genera/species that do not commonly co-exist with (+ve) correlation indicated genera/species (Labeo, Crossocheilus, Garra and Tor) from Factor 2 has also been contemplated by Cantu and Winemiller [22] while describing assemblage patterns on the basis of PCA. Rose and Echelle [18] found the different species with (+ve) loadings on the same Factor tend to occur together and they tend not to occur in low abundance, with species having (-ve) loadings on the Factor as substantiated here by the (-ve) loadings of genera under Factor 3 for Eastern Doon and Factor 2 for Western Doon. The +ve and -ve loadings are also accounted due to the availability of fish in different seasons. Cantu and Winemiller [22] while analyzing fish assemblages concluded that species ordinate according to seasonal abundance with (+ve) values indicating high abundances during fall and winter and (-ve) values indicating high abundance during summer.

The former part in winter is well applicable to the seasonal fish abundance pattern in Doon valley. While scrutinizing the results of FA it has been found that seven genera including Puntius, Chagunius, Barilius, Xenentodon, Mastacembelus, Badisand Puntius in Eastern Doon and four species including Clarias, Mystus, Mastacembelus and Colisa in the Western Doon were not found resolved under any Factor ($r = \leq 0.70$), though they were quite frequent in sample collections. Rose and Echelle [18] regarded the species not resolved under any Factor as to be those having independent distribution. Smith and Fisher [14] and Stevenson et al., [15] also talked about independency of distribution. In the present observations such not resolving genera/species, though have been found close to some Factor, but their uniform and frequent availability throughout the year are not majorly affected by the changing associations, leading to their independency in occurrence. The results of FA have potentially confirmed that species form similar associations even in distantly separated streams of different nature [15] as exemplified by the associations of typical hill stream/ high altitude dwelling (Schizothorax, Schizothorachys, Schistura and Glyptothenorax) and pool/marshy dwelling fishes (Danio, Devario, Esomus, Rasbora, Amblyceps, Clarias, Heteropeustes, Mystus, Macroganathus).

The observations of the present study, temporal variability pattern, characterization of fishes into pool-dwellers and hillstream forms, their preferential substrate affinity was also observed from the results of FA of fishes from both Eastern and Western Doon valley. Abundant evidences are accumulated which suggests that measures of total community behaviour and structure [17] respond similar in different stream system and they are primarily functions of physical and climatic aspects of the stream environment. The results have shown that fish genera/species can be placed along a longitudinal environmental gradient [20, 21] since the physical variables within a river system followed a continuous gradient of physical conditions [2]. The study has revealed the increment in number of species from upstream to downstream sections [19]. The analysis of Factor 1 and Factor 4 from East and Factor 1 from West has clearly elucidated that due to heterogeneity in the downstream stretches, fish communities are stable and diverse [1]. The distribution of fish genera according to their loading values in different Factors showed a specific pattern of association and **vis-a-vis** highlights meaningful interpretations about the ecological characteristics of various genera, the major determinants of association among them.
The results of the present study i.e., factor groupings resolved after FA separately, for Eastern and Western Doon valley are indicative of meaningful interpretation, based on ecological preferences by various genera. In general, FA also supported the observations on distributional pattern that the factors identify the associations from upstream to downstream sections. Through the results of FA, temporal variability pattern, characterization of fishes into pool dwellers and their preferential substrate affinity have become evident. The results have clearly shown that the fish genera/species can be placed along a longitudinal environmental gradient since the physical variables within a river system followed a continuous gradient of physical conditions. FA has also elucidated that due to heterogeneity in downstream stretches, fish communities are stable and diverse.

VI. ACKNOWLEDGEMENT

The study was funded by Uttarakhhand State Council for Science and Technology, Dehradun, Uttarakhand. Authors gratefully acknowledge the support and encouragement received from the Zoological Survey of India, Northern Regional Station, Dehradun, Uttarakhand, India for fish faunal identification and library facility during the study in Doon valley. The authors are also thankful to Dr. Raman Nautiyal, Scientist, Indian Council of Forestry Research and Education, Dehradun, Uttarakhand, India, for statistical data analysis.

BIOGRAPHY

Dr. Deepali Rana is presently working as Assistant Professor at Department of Zoology, Baba Farid Group of Institutions, Dehradun, Uttarakhand, India. Before joining Baba Farid Group of Institutions, she was also Assistant Professor at Uttaranchal College of Science and Technology, Dehradun, Uttarakhand, India. She has two year experience in teaching and five year experience in research. She is M.Sc. and Ph.D. in Zoology with specialization in Fish Taxonomy, Fish Ecology, Aquatic Biodiversity, and Environmental Biology. She has worked as a Junior Research Fellow in a Major Research Project Funded by Uttarakhhand State Council for Science and Technology, Department of Science and Technology, Dehradun, Uttarakhand, India.

REFERENCES