Chapter 2

To assess the pattern of freshwater biodiversity distribution and identify the determinants which influence this distribution

Publications

Munsi, M. and N.A. Aravind. 2016. The pattern of freshwater biodiversity distribution and identify the determinants of its distribution in India. (Under preparation)
Pattern of freshwater biodiversity distribution

Introduction

Freshwater habitats occupy 0.8% of the earth’s surface. Within this small extent, it houses extraordinarily high level of biodiversity (Revenga and Kura, 2003; Naiman et al., 2006; Dudgeon et al., 2006). The freshwater systems support various orders of animals, plants, fungi and algae, contributing to almost 10% of all the species and 35% of all the vertebrates described till date (Strayer and Dudgeon, 2010; Molur et al., 2011; Stenderaet al., 2012). The relative species richness with respect to per unit area (i.e. habitat extent) of freshwater (3.0) is higher than both marine (0.2) and terrestrial systems (2.7) (McAllistore et al., 1997).

The anthropogenic pressures and threats (see Figure 1.2) have resulted in declining freshwater biodiversity at an alarming rate (see Figure 1.3; Dudgeon et al., 2006; WWF, 2014). According to the Living Planet Index, falling by 76%, populations of freshwater species declined more rapidly than marine (declined by 39%) and terrestrial (declined by 39%) populations (WWF, 2014). One of the major gaps in freshwater biodiversity conservation is a lack of data on species distribution, especially in the tropical regions. The conservation efforts for freshwater biodiversity is also far behind the conservation of terrestrial systems (Stenderaet al., 2012).

As the species richness of a region is not distributed homogeneous at a spatial scale, understanding the geographic distribution of species and understanding the main predictor variables that explain the species richness are important. The assessment of biodiversity distribution patterns over space and time, and across environmental gradients is crucial to understand the origin, function and maintenance of biodiversity (Hawksworth, 1995; Heywood and Watson, 1995). Also, knowledge on species distribution pattern is important for developing successful conservation plans (Kerr, 1996; Heino, 2002; Karanthen et al., 2009; Jetzet et al., 2012; Pimmet et al., 2014; Azaeleete et al., 2015). In recent years, a number of studies have been carried out to assess spatial patterns of biodiversity distribution. Even for freshwater systems, efforts have been made to map freshwater resources and biodiversity (Fitoka and Keramitsoglou, 2008; Naiman et al., 2006; Stenderaet al., 2012; http://www.feow.org/). However, data on biodiversity distribution at broad spatial scale is still
lacking (Gaston et al., 1995; Gaston and Williams, 1996; Kerr, 1996; Lamoreux et al., 2006; Karanth et al., 2009; Jetz et al., 2012; Azaele et al., 2015).

There are various proposed explanations for the patterns of biodiversity distribution (Whittaker, 1975; Stevens, 1989; Huston, 1994; Rosenzweig, 1995; Karanth, 2011). In general, much of the geographical variation in species distributions, both in space and time, can be related to either climatic (e.g. mean annual temperature, potential evapo-transpiration) or historical (e.g. ice age glaciations, dispersal) influences (Whittaker, 1975; Stevens, 1989; Currie, 1991; Brown and Lomolino, 1998; Krefl and Jetz, 2007). The debate is still continuing among researchers about the scale at which some of these factors operate (Caley and Schluter, 1997; Scott et al., 2011).

Species distributions and assemblage composition can be studied across several spatial scales (Wiens, 1989; Levin, 1992; Brown, 1995; Maurer, 1999). The focus of large spatial scale analysis is to understand the patterns in large sets of species at broad spatial scales (Brown, 1995; Lawton, 1999). Historically, most of the work relates to the analysis of determinants of species at very small scale like streams, lakes or rivers (Jackson and Harvey, 1993; Allen et al., 1999a; Heino, 2002; Turtureanu et al., 2014). Undoubtedly, such approaches yield important insight on how local habitat conditions influence species assemblage. Nevertheless, given that local communities are also strongly induced by regional species pools (Ricklefs, 1987; Ricklefs and Schluter, 1993), it is important to understand which environmental factors affect species composition at broader spatial scales. Furthermore, broad-scale analyses may generate bolder patterns and more generalizable hypotheses on factors determining species distributions and the limits of regional species pools (Currie, 1991; Brown, 1995; Lawton, 1999; Maurer, 1999).

Only a few freshwater studies have addressed the concordance of community structure and species richness among different taxa (Jackson and Harvey, 1993; Allen et al., 1999b; Kilgour and Barton, 1999; Paszkowski and Tonn, 2000; Heino, 2002). In general, different taxa seem to show broadly concordant community patterns among local systems (Jackson and Harvey, 1993; Paszkowski and Tonn, 2000), although the environmental determinants of these patterns may vary among taxa (Jackson and Harvey, 1993; Allen et al., 1999a). However, the degree of concordance in species richness may be low among some taxonomic
groups (see Allen et al., 1999b), and almost certainly varies with scale (Gaston and David, 1994; Reid, 1998). The study of concordance of distributional patterns amongst different biological groups can help determine the factors involved in shaping patterns in distribution, which in turn will help in determining surrogates for conservation planning. The question still unanswered is whether different groups of organisms respond to similar environmental gradients across different spatial scales? And what historical factors have contributed to present-day distributional patterns (Heino, 2002). Thus, the cross-taxon analysis allows us to understand broad distributional patterns and their determinants.

There are no studies in India that have compared patterns and determinants in the geographical distribution of freshwater taxa at the national scale. However, only the geographical distribution patterns in freshwater fishes, (Dahanukar et al., 2004), amphibians (Daniels, 1992; Sarma, 2012), freshwater molluscs (Aravind et al., 2011) and odonates (Subramanian and Shivaramakrishnan, 2002) have been studied at the Western Ghats scale but, the factors that determine their distribution patterns are not known. Hence, this chapter addresses the following questions: 1) What is the distribution pattern of selected freshwater taxa (aquatic birds, freshwater molluscs, anurans and freshwater fishes) in India? 2) What are the variables that are the best determinants of species distribution in India? 3) Is there a concordance between different freshwater taxa at a large spatial scale?

**Materials and Methods**

**Study Area**

The study was carried out at all-India scale. The total freshwater area in India is 3,14,400 km². Freshwater habitats in India houses approximately 1,000 species of freshwater fishes (out of which almost 25% are endemic), 220 species of freshwater molluscs (out of which almost 40% are endemic), 380 species of amphibians and 242 species of wetland dependent birds (Gururaja and Aravind, 2006; Molur et al., 2011). The World Conservation Monitoring Centre in 1999 had designated the Western Ghats biodiversity hotspot as the “Freshwater hotspot for fishes, molluscs and crabs” (Thomson, 1999; Molur et al., 2011).

For the present study, river sub-basins of India were used. The GIS layer was downloaded from [http://hydrosheds.cr.usgs.gov/datadownload.php?reqdata=30bass](http://hydrosheds.cr.usgs.gov/datadownload.php?reqdata=30bass) (Annexure 2.1). A
total of 2,254 sub-basins in India were used for analysis. The number of sub-basins with respect to the size of sub-basins represents normal distribution (Annexure 2.2) and shows that numbers of sub-basins are highest under medium size ranges, whereas sub-basins with very small or large areas are fewer in number.

Data compilation

Biodiversity data

Distribution of four freshwater taxa, i.e. aquatic birds, freshwater molluscs, anurans and freshwater fishes were used in this study. The distribution data (Annexure 2.3) were obtained from primary field surveys (freshwater molluscs and anurans) as well as from secondary sources like published reports, journals articles and existing data points from Zoological Survey of India (Kolkata; for freshwater mollusces only). The data were then geo-coded using Google Earth. This geo-coded information was then converted into shapefiles for spatial representation of biodiversity distribution. A total of 851 data points were obtained for 145 freshwater mollusc species (Annexure 2.3a); 6,040 data points for 800 species of freshwater fishes (Annexure 2.3b), 13,488 data points for 284 species of aquatic birds (Annexure 2.3c); and 3,972 data points for 229 species of anurans (Annexure2.3d) (Table 2.1). The datasets were corrected by removing those points, which did not have location information or those records, which had very vague or very broad location information such as “India”, “South India”, “Malabar”, etc. The list of threatened species was obtained from the IUCN Red List of Threatened Species (http://www.iucnredlist.org/). This information was linked with distribution data to obtain the spatial distribution of threatened species in India (Annexure 2.3f). The details of the datasets used are provided in table 2 and 3. The analysis was done using Google Earth, DIVA-GIS (Ver. 7.5.0), ArcGIS (Ver. 10.3), and MySQL platforms implemented in Windows 7.

Table 2.1: Biodiversity dataset used in this study

<table>
<thead>
<tr>
<th>Biodiversity datasets</th>
<th>Type of data</th>
<th>Data points</th>
<th>Number of Species</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshwater molluscs</td>
<td>Primary data</td>
<td>851</td>
<td>145</td>
<td>Field work</td>
</tr>
<tr>
<td></td>
<td>Secondary data</td>
<td></td>
<td></td>
<td>Literature review; ZSI, Kolkata;</td>
</tr>
</tbody>
</table>
Variables affecting the distribution of freshwater taxa

Data on threats to freshwater systems, physical factors and resource use were analysed to understand their roles in determining the distribution of freshwater taxa. Table 2.2 provides the list of these variables.

Data for threats to freshwater systems: The list of 5,000 large dams was obtained from Central Water Commission, Government of India. A large dam is defined as any dam with a height of 10 meters or more. The large dams were then geo-coded and then digitised in Google Earth to obtain the spatial distribution of the dams (Annexure 2.4d).

Data for physical factors: The data on temperature seasonality (Annexure 2.5a), temperature (Annexure 2.5b), elevation (Annexure 2.5c) and rainfall distribution (Annexure 2.5d) were obtained from the worldclimdatabase(http://www.worldclim.org/current). The resolutions of these data are 30 arc seconds (~1 km²). The river network data was obtained from http://hydrosheds.cr.usgs.gov/hydro.php (Annexure 2.5e). Data on water chemistry such as turbidity, pH, salinity, etc., are not available for India, hence not used for the analysis.

Data for resource use: The Land Use Land Cover (LULC) map for India was clipped from GlobCover 2009 map derived from MODIS (Moderate Resolution Imaging Spectroradiometer) flown on two NASA spacecraft (Terra and Aqua) with the resolution of 300 meters (Annexure 2.6). The numbers of pixels under each land use/cover (habitat types) were counted to calculate the area of each habitat and number of habitats under each sub-basin.

Table 2.2: List of Variables used and the sources from where it has been downloaded
<table>
<thead>
<tr>
<th>Categories</th>
<th>Sub layers</th>
<th>Unit</th>
<th>Sources of data</th>
<th>Resolution</th>
<th>Platform used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threats</td>
<td>Dams</td>
<td>km</td>
<td>Central Water Commission</td>
<td>5000 large dams</td>
<td>Google earth; ArcGIS 10.3; DIVA-GIS 7.5.0; MySQL</td>
</tr>
<tr>
<td>Physical factors</td>
<td>Rainfall</td>
<td>mm</td>
<td><a href="http://worldclim.org/current">http://worldclim.org/current</a></td>
<td>30 arc second</td>
<td>ArcGIS 10.3; MySQL</td>
</tr>
<tr>
<td></td>
<td>Altitude</td>
<td>m asl</td>
<td><a href="http://worldclim.org/current">http://worldclim.org/current</a></td>
<td>30 arc second</td>
<td>MySQL</td>
</tr>
<tr>
<td></td>
<td>Mean Temperature</td>
<td>°C</td>
<td><a href="http://worldclim.org/current">http://worldclim.org/current</a></td>
<td>30 arc second</td>
<td>MySQL</td>
</tr>
<tr>
<td></td>
<td>Temperature seasonality</td>
<td>°C</td>
<td><a href="http://worldclim.org/current">http://worldclim.org/current</a></td>
<td>30 arc second</td>
<td>MySQL</td>
</tr>
<tr>
<td></td>
<td>River length</td>
<td>km</td>
<td><a href="http://hydrosheds.cr.usgs.gov/hydro.php">http://hydrosheds.cr.usgs.gov/hydro.php</a></td>
<td>15 arc second</td>
<td></td>
</tr>
<tr>
<td>Resource use</td>
<td>LULC Map</td>
<td></td>
<td>GlobCover 2009 (MODIS): <a href="http://due.esrin.esa.int/page_globcover.php">http://due.esrin.esa.int/page_globcover.php</a></td>
<td>300 m</td>
<td>ArcGIS 10.3</td>
</tr>
</tbody>
</table>

**Analysis**

*Spatial Join:* The shapefiles representing the distribution of data points of each taxon were joined to the sub-basin layer of India using “spatial join” tool in ArcGIS (ver 10.3). This tool helps in joining attributes from one feature to another feature on the basis of spatial relationships like latitude and longitude. Once the distribution points were joined to the sub-basins of India, the total number of species and threatened species were counted for each sub-basin.

*Circular Neighbourhood Analysis:* The distribution of data points of each taxon were used for circular neighbourhood analysis. The calculations are made based on the observation found within a circle of specified radius with its centre in the middle of each grid cell. The advantage of this method is that it produces a smoother surface (Hijmans et al., 2012). The
The grid size used for the analysis was 1 km². The analysis was done to obtain species richness maps for individual taxa as well as for combined dataset.

**Statistical analysis**: Non-parametric Spearman rank correlation was used to remove the auto-correlated variables (Annexure 2.7). Since mean temperature and temperature seasonality were highly correlated, the temperature was removed. Variation in temperature seasonality is known to affect species richness (Clarke and Gaston, 2006) hence this variable was chosen over temperature. The variables selected for further analyses are mean rainfall (mm), elevation (coefficient of variation) (m), temperature seasonality (standard deviation) (°C), river length (m), dams (km²) and land use land cover heterogeneity. Generalised Linear Model (GLM) was used to identify the combination of independent variables that are the best determinant of the freshwater taxa distribution. Best four models are identified for distribution of all four taxa together, as well as for individual taxon distribution. Table 2.4

---

**Results**

**Species distribution patterns**

The species richness map for different sub-basin across India for all taxa combined is shown in Figure 2.1. The map shows there is a huge gap in the data, especially in central and western parts of India. As expected, the number of studies and thus number of species recorded in the Western Ghats region is comparatively higher than other parts of India. Urban regions were found to have high richness for birds than other regions.

For anurans (Figure 2.1c), the highest diversity is seen in the Western Ghats region followed by north-east India. When compared to other taxa, the data availability for aquatic birds (Figure 2.1d) is much better. For freshwater fishes, the Western Ghats and the Eastern Ghats are species rich and also certain regions of north-eastern India. For freshwater molluscs, the central Western Ghats and parts of eastern India and north-east India show high diversity.
**Distribution of Threatened freshwater taxa**

The sub-basins with a highest number of threatened species for all taxa are reported from the Western Ghats. Few sub-basins in north-east India also have a few threatened species (Figure 2.1f). For individual taxa also the same pattern is seen, however, with variation within the biogeographic regions.

**Basin area and species richness**

The species richness (Figure 2.2a) and threatened species were plotted against an area of sub-basin (Figure 2.2b) to understand the influence of sub-basin size on the distribution of species. The result shows that there is a weak positive correlation between sub-basin area and species richness for all taxa ($R^2 = 0.013$, $P<0.05$) and between sub-basin area and threatened species richness ($R^2 = 0.012$, $P<0.05$). The same pattern is seen for individual taxa as well.

**Determinants of species richness**

The six variables selected from Spearman rank correlation were used for GLM to identify the combination of variables that best describes the distribution pattern of freshwater taxa in India. When the distribution of all the taxa was considered together, the result shows that mean rainfall, CV of elevation, and land use heterogeneity are the most important variables that are positively correlated with distribution. The temperature seasonality shows a negative correlation with species richness. The influence of river length is weak. For anurans, the result shows that mean rainfall, CV of elevation, and land use heterogeneity were the most important variables that positively correlated with distribution. However, for freshwater molluscs, CV of elevation shows negative correlation and mean rainfall and land use heterogeneity were positively correlated with distribution and influence of river length is weak. For aquatic birds, mean rainfall, CV of elevation, and land use heterogeneity are the most important variables that were positively correlated with distribution. Similarly, for freshwater fishes, mean rainfall, CV of elevation and land use heterogeneity are the most important variables that were positively correlated with distribution. For all taxa and for combined data, the temperature seasonality is negatively correlated and influence of river length is weak (Table 2.3).

**Concordance in species richness patterns**
Among different taxa, a strong positive correction was found between anurans with molluscs and birds and a slightly weak relation between birds and molluscs. However, all the correlations are highly significant ($P<0.001$). Concordance among combined data for all species and individual species was also relatively high (Table 2.4).
Figure 2.1: Distribution map of a) freshwater molluscs, b) freshwater fishes, c) anurans, d) aquatic birds, e) all species and d) threatened species in India
Figure 2.2: Correlation between a) molluscs and sub-basin size, b) freshwater fishes and sub-basin size, c) frogs and sub-basin size, d) aquatic birds and sub-basin size, e) all taxa and sub-basin size, and f) number of threatened species and sub-basin size
Table 2.3: Best four GLM models for all taxa and combined data for all species

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Model</th>
<th>AIC</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>$R + E - TSs + RI + LU$</td>
<td>21622.86</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + D + RI + LU$</td>
<td>21624.25</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + LU$</td>
<td>21634.92</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs - D + LU$</td>
<td>21636.31</td>
<td>0.176</td>
</tr>
<tr>
<td>Anurans</td>
<td>$R - TSs + D + LU$</td>
<td>11438.96</td>
<td>0.212</td>
</tr>
<tr>
<td></td>
<td>$R - TSs + D + RI + LU$</td>
<td>11439.69</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + D + LU$</td>
<td>11440.64</td>
<td>0.212</td>
</tr>
<tr>
<td></td>
<td>$R - TSs + RI + LU$</td>
<td>11441.42</td>
<td>0.211</td>
</tr>
<tr>
<td>Freshwater</td>
<td>$R - E - TSs + RI + LU$</td>
<td>5218.01</td>
<td>0.072</td>
</tr>
<tr>
<td>Molluscs</td>
<td>$R - E - TSs + D + RI + LU$</td>
<td>5218.86</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>$R - E - TSs + LU$</td>
<td>5219.68</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>$R - TSs + LU$</td>
<td>5221.19</td>
<td>0.066</td>
</tr>
<tr>
<td>Aquatic Birds</td>
<td>$R + E - TSs + RI + LU$</td>
<td>16640.41</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + LU$</td>
<td>16642.02</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + D + RI + LU$</td>
<td>16642.18</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs - D + LU$</td>
<td>16643.94</td>
<td>0.033</td>
</tr>
<tr>
<td>Freshwater</td>
<td>$R + E - TSs + RI + LU$</td>
<td>10893.79</td>
<td>0.088</td>
</tr>
<tr>
<td>Fishes</td>
<td>$R + E - TSs + D + RI + LU$</td>
<td>10895.78</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs + LU$</td>
<td>10902.77</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>$R + E - TSs - D + LU$</td>
<td>10903.76</td>
<td>0.08</td>
</tr>
</tbody>
</table>

(R: mean rain; E: elevation coefficient of variation; TSs: temperature seasonality standard deviation; RI: river length; D: dams; LU: land use land cover heterogeneity)

Table 2.4: Concordance between different freshwater biodiversity elements used in this study

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Anurans</th>
<th>Aquatic Birds</th>
<th>Freshwater fish</th>
<th>Freshwater molluscs</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anurans</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aquatic Birds</td>
<td>0.424*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freshwater fish</td>
<td>0.409*</td>
<td>0.394</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freshwater molluscs</td>
<td>0.449*</td>
<td>0.346</td>
<td>0.427*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0.632**</td>
<td>0.860**</td>
<td>0.651**</td>
<td>0.525*</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*P<0.05, **P<0.01
Discussion

The determinants of species distribution and the concordance between different freshwater taxa were carried out at a pan-India scale for the first time for selected freshwater taxa. The results of this study show that there is a huge gap in species distribution data of freshwater taxa in India, especially in the central and western parts. There is a high spatial bias in the data used for all taxa. This is inevitable and the problem is not specific to this study (Williams et al., 2002). It is interesting to note that certain urban regions have high richness for birds than other regions. This is mainly because of a large number of birdwatchers are concentrated in these urban areas. However, with limited data, a broad conclusion can be drawn (Currie, 1991; Brown, 1995; Lawton, 1999; Maurer, 1999). Much of the freshwater biodiversity, like terrestrial, is concentrated in the Western Ghats and the Eastern Himalaya, the two of the four-biodiversity hotspots of India and to smaller extent in the Eastern Ghats. Even though North-eastern India is also known for its unique biodiversity and has been included within the Eastern Himalayan biodiversity hotspot, the region was not seen highlighted in the species richness map, because the region is still largely under-sampled due to difficult terrain and lack of expertise in that area (Bhuyan et al., 2003; Chatterjee et al., 2006; Aiyadurai, 2011).

Current distribution patterns in freshwater biodiversity in India might have been influenced by historical factors as shown earlier in several studies (Tonnet et al., 1990; Gaston 2000; Heino, 2002). There are several pieces of evidences to show how historical factors have influenced the distribution of some of the Indian freshwater taxa. For example, the freshwater mollusc genus Paracrostoma show “into India” migration from SE Asia (Köhler and Glaubrecht, 2007), the toad family Bufonidae dispersed into India and diversified during the Late Oligocene to Early Miocene (van Bocxlaer et al., 2009). The current distribution is influenced by past and ongoing changes in the environment mainly due to anthropogenic activities, such as land use and land cover change, pollution and introduction of exotic species. All these factors influence species interaction and distribution. Like elsewhere, there is a strong latitudinal distribution pattern seen in India, more specifically in the Western Ghats. For example, distribution of fishes, molluscs, amphibians, odonates and aquatic plants decreases significantly as one move from south to north (Subramanian and Shivaramakrishan, 2002; Dahanukar et al., 2004; Aravind et al., 2011; Sarma, 2012). This decrease in the richness with an increase in latitude might be due to climatic factors such as temperature and
rainfall. In India, as one moves north, the rainfall and number of rainy day decrease and mean temperature increase with high seasonality. It is clear from this study that the temperature seasonality (SD) was negatively correlated to species distribution of freshwater taxa in India. Variation in temperature has been regarded as major correlates of species richness at broad spatial scales (Currie, 1991; Wright et al., 1993). In the tropics, subtropics, and warm temperate zones, water variables were shown to be the strongest predictors of species distribution (Hawkins et al., 2003). In this study, a major proportion of the variation in species richness was explained by rainfall, elevation, land-use (habitat heterogeneity) and temperature seasonality for all taxa as well as for combined data. Land use land cover heterogeneity seems to be the most important variable that determines the distribution patterns. Thus, the type of habitat in which the freshwater ecosystem is found plays an important role in determining species richness. Diverse habitats, especially areas with natural land cover provide more niches and thus, more number of species can adapt and contribute to higher biodiversity richness. Mean rainfall and variation in elevation also are positively correlated (Kathuria and Ganeshiah, 2002). River length and area under dams show positive but weak correlation. Variation in temperature seasonality is, as expected, negatively correlated with species distribution. At individual taxon level, the freshwater fish diversity and distribution is influenced by climate as shown by McAllister et al. (1986) and Tonn (1990). Species’ distributions in the other freshwater taxa have also shown to be influenced by broad-scale climatic factors (e.g. Abell et al., 2000). At a global scale, the study by Kreft and Jetz (2007) shows that potential evapotranspiration, the number of wet days per year, and measurements of topographical and habitat heterogeneity emerge as core predictors of vascular plants species richness, whereas the study by Lobo et al. (2001) at a smaller scale (Iberian Peninsula) showed that maximum altitude was the major predictor. At a much smaller scale, individualspecies and functional groups might respond differently to different habitat and environmental variables (Faith and Norris, 1989; Bhat, 2004; Leathwicket al., 2005; Hoeinghaus et al., 2007). Studies in Europe and China have shown that temperature seasonality is the major influencing factor for aquatic macro-invertebrates especially aquatic insects (Tonkin et al., 2015; Shah et al., 2015). In this study also, temperature seasonality emerged as one of the most influencing environmental variable for all taxa. This suggests that despite the large gap in the data, the broad pattern presented here conform to the pattern seen elsewhere.

**Concordance among different freshwater taxa**
At the global scale, the patterns of richness are highly correlated among amphibians, reptiles, birds and mammals and endemism patterns (Lamoreux et al., 2006). In this study, the concordance in the species richness of fishes, anurans, molluscs and aquatic birds and for combined data across all taxa was very high. However, global distribution patterns for one species is not necessarily representative of those for other freshwater taxa as shown by Heino et al. (2005). For example, the distribution of amphibians (http://www.iucnredlist.org/initiatives/amphibians) and freshwater turtles (Buhlmann et al., 2009) indicate that while there is considerable overlap between fish, amphibian and turtles but there are some areas which are unique for amphibians and turtles but not for fishes (http://www.iucnredlist.org/initiatives/amphibians; http://www.feow.org/; Buhlmann et al., 2009). This has been clearly shown in studies conducted in Central America, north and eastern coastal Australia, Upper Brahmaputra freshwater ecoregion and the North American drainages. All these areas are not very rich for fishes but are extremely rich for freshwater turtles (Abell et al., 2008; Buhlmann et al., 2009; http://www.feow.org/). A study in Andean-Amazon basin shows high congruence among five freshwater vertebrates (birds, mammals, reptiles, amphibians and fishes) groups (Lessmann et al., 2016). It is worth testing if a similar pattern is prevalent when one considers other taxa, especially freshwater invertebrates.

The degree of concordance in species richness among different taxa varies with the scale of the study as well (Gaston and David, 1994; Prendergast and Eversham, 1997), and thus, one should use this with caution (Prendergast et al., 1993; Gaston and Williams, 1996; Heino, 2010). It has been postulated that the concordant patterns can result from biotic interactions among different taxa, which may not be strong enough in this study due to a huge data gap. Heino et al., (2003) also did not find any pattern of concordance among insect groups in Finland. In addition, most of the studies that found strong patterns of concordance considered organisms of similar size (Allen et al., 1999a; Bilton et al., 2006; Bini et al., 2008). Similarly sized organisms can exhibit similar responses to environmental variation as they tend to utilize similar environmental and habitat conditions (Allen et al., 1999a; Heino, 2010). In the present study, the size difference between the taxa was highly variable (with birds being large size and freshwater mollusc being small). It is also suggested that the habitat-use is a phylogenetically conserved trait (Webb et al., 2002), hence, it is reasonable to predict a stronger concordance between phylogenetically related taxa than between less related ones (Beccaloni and Gaston, 1995).
Gaston and Williams (1996) has suggested that the mechanisms responsible for strong cross-taxon congruence can be divided into five general groups. According to them strong congruence may be generated by (a) random draw of species from the regional species pool; (b) similar responses of different taxonomic groups to the same environmental conditions or gradient(s); (c) responses to different, but correlated environmental gradients; (d) biotic interactions and (e) inconsistent sampling effort, with some sites being sampled more efficiently for multiple organism groups, may affect congruence, species richness patterns in particular (Gaston, 2000). The last one is particularly true in this study as many areas in central, eastern and western India is less sampled compared to the Western Ghats.

**Caveats in the study**

The biggest problem faced was the data gap. The collection of new data on biodiversity was not feasible given time, cost and area of study hence, existing data that are available from various sources were used in this study. Another biggest caveat of this study is a lack of extensive basin wise water chemistry information. Species richness estimates of different sub-basin are generally unreliable due to lack of means to determinethe completeness of inventory data (Gaston, 1996; Lobo et al., 2001). In spite of the lack of sampling effort measurements, an attempt can be made to identify and understand geographical patterns of species distribution using information based on regional and national atlases (Lobo et al., 2001).

Out of the 2,254 sub-basins in India, 285 sub-basins didn’t have any data, and 444 sub-basins have <5 data points. Only 1,207 sub-basins (53.54 %) have more than 10 data points. Even common and widespread species are not well represented in most sampling units in India as shown in an earlier study for frogs of the Western Ghats (Roshmi, 2012). The data obtained is also not uniformly distributed. For example, the data for bird distribution is highly concentrated near cities largely due to the number of bird watchers in these regions (Figure 2.1d). Hence, an extensive survey is required to address this data deficiency. With more data, the clear patterns in distribution will emerge and will improve prioritization exercises. Even though it is well known that dams threaten the downstream biodiversity, the areas under dams are showing a positive correlation because the reservoir areas generally harbour a higher number of migratory birds and introduced and common fishes. The downstream impact of dams couldn’t be highlighted for doing the analysis due to lack of data.
Unfortunately, describing geographical patterns of species number and identifying the important factors are complicated by two principal methodological difficulties: (i) to consider all the available historical and present biological and geographical species data, while taking into account the unevenness of sampling effort in different areas; and (ii) to differentiate the patterns from processes, detecting causal relationships underlying multivariate environmental and spatial correlations (Lobo et al., 2001). Understanding broad-scale biodiversity distribution pattern and factor responsible for their distribution is essential for understanding the origins, maintenance, conservation and management of biological diversity at a large ecosystem scale (Markovic et al., 2014; Shah et al., 2015).


References


Whittaker, R.H. 1975. Communities and ecosystems. 2nd ed. Macmillan, New York


Annexure

Annexure 2.1: Sub-basins in India

Annexure 2.2: Distribution of sub-basins under different size range
Annexure 2.3: Distribution of freshwater mollusc (a), freshwater fishes (b), aquatic birds (c), anurans (d), all taxa (e), and threatened taxa (f)
Annexure 2.4: Threats to freshwater systems: (a) Population data, (b) Railway network, (c) Road network, and (d) distribution of large dams
Annexure 2.5: Physical factors: (a) Temperature seasonality, (b) Temperature, (c) Elevation, (d) Rainfall, and (e) River network
Annexure 2.6: Land Use Land Cover (LULC) map of India
### Annexure 2.7: Spearman correlation analysis between the independent variables considered in the study

<table>
<thead>
<tr>
<th></th>
<th>SpC</th>
<th>R</th>
<th>E</th>
<th>Ts</th>
<th>Tcv</th>
<th>TSs</th>
<th>TScv</th>
<th>RI</th>
<th>D</th>
<th>Ro</th>
<th>Ra</th>
<th>P</th>
<th>LU</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpC</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.20*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.15*</td>
<td>0.42</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ts</td>
<td>-0.16*</td>
<td>0.21</td>
<td>0.35</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tcv</td>
<td>-0.18*</td>
<td>0.35</td>
<td>0.41</td>
<td>0.86</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSs</td>
<td>-0.20*</td>
<td>-0.02</td>
<td>0.16</td>
<td>0.62*</td>
<td>0.52*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TScv</td>
<td>-0.09</td>
<td>0.24</td>
<td>0.40</td>
<td>0.63</td>
<td>0.62</td>
<td>0.78</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rl</td>
<td>0.15*</td>
<td>-0.17</td>
<td>-0.11</td>
<td>0.39</td>
<td>0.36</td>
<td>0.53</td>
<td>0.25</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.04</td>
<td>-0.11</td>
<td>0.00</td>
<td>0.25</td>
<td>0.32</td>
<td>0.29</td>
<td>0.31</td>
<td>0.41</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ro</td>
<td>0.15*</td>
<td>-0.16</td>
<td>-0.12</td>
<td>0.27</td>
<td>0.30</td>
<td>0.41</td>
<td>0.22</td>
<td>0.87</td>
<td>0.44</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ra</td>
<td>0.10*</td>
<td>-0.17</td>
<td>-0.08</td>
<td>-0.17</td>
<td>-0.06</td>
<td>0.00</td>
<td>-0.09</td>
<td>0.41</td>
<td>0.19</td>
<td>0.49</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.09</td>
<td>0.11</td>
<td>0.00</td>
<td>-0.49</td>
<td>-0.34</td>
<td>-0.42</td>
<td>-0.27</td>
<td>-0.24</td>
<td>-0.24</td>
<td>-0.07</td>
<td>0.33</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>LU</td>
<td>0.15*</td>
<td>0.34</td>
<td>0.41</td>
<td>0.41</td>
<td>0.46</td>
<td>0.30</td>
<td>0.47</td>
<td>-0.08</td>
<td>0.14</td>
<td>-0.11</td>
<td>-0.25</td>
<td>-0.33</td>
<td>1.00</td>
</tr>
</tbody>
</table>

1. (SpC: species richness; R: mean rainfall; E: elevation coefficient of variation; Ts: temperature standard deviation; Tcv: temperature coefficient of variation; TSs: temperature seasonality standard deviation; TScv: temperature seasonality coefficient of variation; RI: river length; D: dams; Ro: road; Ra: rail; P: population; LU: land use land cover heterogeneity; *p<0.05, df = 2252)