

Mathematical and Statistical investigations of Ecological and Evolutionary Systems

The ongoing investigations, supported by during the previous plan by the Ministry of Environments and Forests through a block grant to the Centre for Ecological Sciences, will be continued during the X plan period as well. These would essentially consist of statistical analysis of ecological data, collected during the different studies conducted by the Center for Ecological Sciences, and also by investigators from other organizations with whom there is an active ongoing collaboration. Occasionally, fresh data may also be collected if during analysis it is found that some additional data is likely to considerably improve the scope or depth of the study. The statistical studies will be complemented by investigations of mathematical models that describe ecological dynamics and evolutionary processes and their consequences. Some of the specific questions to be addressed are as follows:

How should one estimate the total number of species in a given region from a small set of samples? Is it necessary to develop special approaches for special habitat types such as wetlands?

How do the neighbouring individuals influence the rate of growth and survival of individual trees in natural ecosystems?

How best should one characterize different habitats/vegetation types based on initial sample surveys to maximize detection rates of new species during subsequent surveys?

Can better ways of estimating standing biomass and annual productivity be devised, making use of the relatively inexpensive imaging devices and image analysis software packages that are recently becoming readily available?

There is a resurgence of interest in some of these old questions, partly due to their relevance to important contemporary issues e.g. Inventorying Biodiversity, exploring approaches to Sustainable Development, Devising Mitigation Measures for Global Warming etc. Additionally, there has been good progress on the theoretical side on topics related to these questions - the infinite series solution to the differential equation describing the species area relationship, and the neutral theory of biodiversity being two specific examples.

The investigations to be undertaken, described above, would be complemented by the development of teaching aids for web-based courses on mathematical modeling and on statistical analysis of ecological data. Extensive use will be made of the widely available and extremely inexpensive open source software (mainly Linux based) for this purpose. Attempt will also be made to make the software packages developed at the centre available to students and researchers in other organizations on CD-ROMS. Responses and feedback from the users will be helpful in regularly improving and periodically updating these packages.

Budget

Budget Head	First Year	Second Year	Third Year	Fourth Year	Fifth Year	Total
Salary	291000	301200	311400	321600	331800	1557,000
Project Assistant -2	180000	186000	192000	198000	204000	960000
Lab Assistant - 1	60000	62400	64800	67200	69600	324000
Field Assistant - 1	51000	52800	54600	56400	58200	273000
Equipment	75000	100000	75000	100000	75000	425000
Travel and other project Costs	75000	80000	85000	90000	95000	425000
Total	441000	481200	471400	511600	501800	2407000

PROJECT 12

(1997-)

1. Title : Mathematical and Statistical Ecology

2. Summary:

While many investigations have been carried out on the patterns of species diversity along a latitudinal, those along an altitudinal gradient have not been studied as extensively. This is particularly true of bird species, the pioneering and comprehensive work of Terborgh being a notable exception. Most of these studies, however, have been in the Neotropics. It is therefore desirable to extend these studies to tropical rainforests in general, and in India, to the Western Ghats in particular, the latter being one of the biodiversity hot spots. One of the first such investigations has been carried out by Mr. T. R. Shankar Raman, a research student at our Center for Ecological Sciences, who has just submitted his thesis under the supervision of Prof. R. Sukumar. As a part of his study of avian communities in the the Kalakad-Mundanthurai Tiger Reserve (KMTR) in southern Western Ghats, Mr. Shankar Raman sampled bird communities at different locations ranging in altitude from about 600 meters to 1300 meters. This report presents an analysis of the pattern of the variation of beta diversity (species turnover) along this elevational gradient in comparison with a set of null models.

The data set comprises of a total of 58 species of rainforest birds, and presence/absence of each species has been recorded at each of the 14 locations, so as to constitute a 58X14 matrix. Species turnover (beta diversity) between every pair of locations was quantified using Jacard index as well as the Euclidean distance. The mean value of beta diversity between all 91 pairs was 0.53 for Jacard index and 18.0 for Euclidean distance, whereas the corresponding values for the 13 pairs of adjacent sites were 0.69 and 10.2 respectively, indicating that the neighbouring sites were much more similar to each other than non-neighbouring sites, i.e., indicating a spatial gradient. To examine the robustness of this pattern, Monte Carlo simulations were carried out using a set of four null models - (i) row totals held constant, (ii) column totals held constant, (iii) both row and column totals held constant (double constraint), and (iv) clustered occurrence (species more likely to be present in adjoining sites) over and above the double constraint. The results indicated that each of the models performed better than the previous one in being closer to the observed values of mean similarity. However, even the double constrained clustered occurrence model underestimated the mean similarity between adjacent sites. A possible interpretation is that the species are not independently distributed of each other, and that they form well-defined assemblages. Further analysis to examine this aspect is in progress.

3. Objectives :

The composition of avian communities at any location is governed by a combination of deterministic and stochastic factors. While unambiguous demonstration of deterministic influences like antagonistic or synergistic interactions between species is difficult, it is relatively easy to estimate the contribution of some of the stochastic factors. Here too, one needs to judiciously specify the constraints under which the stochastic factors will operate. The objective of the present investigation is to examine the patterns of species turnover along an elevational gradient vis-a-vis those obtained from four hierarchical null models incorporating increasingly tighter constraints. In particular, the objective is to examine the extent to which the observed high level of similarity between the avian communities at adjacent sites can be accounted for by constructing random assemblages wherein the species are distributed independently of each other.

Towards achieving this overall objective, an attempt has been made to examine the effect of imposing one or more constraints so as to make the random assemblages conform to four of the simpler aspects of the observed patterns: (i) the number of species observed at a site being kept constant, (ii) the number of sites at which a species is seen being kept constant (iii) a double constraint

wherein both the previous constraints are imposed and finally (iv) the observed level of contiguity of distribution of species being maintained, in addition to imposing the double constraint mentioned earlier.

4. Methodology:

Sampling:

The data used in this analysis was collected by Mr. T. R. Shankar Raman mostly during February to May 1998 (data at four sites being collected in March 1999) in the Kalakad-Mundanthurai Tiger Reserve at the forest camps located near Kannikatti (740 m asl, 8°37' N and 77°16' E), Sengaltheri (1040 m asl, 8°31' N and 77°26' E) and Kakachi (1220 m asl, 8°33' N and 77°24' E), the details have been described in his thesis as well as in several publication. Briefly, avian communities at 14 different locations were sampled using the fixed radius point count method. The surveys were carried out during the first three hours after sunrise. The count duration was five minutes, and all birds seen, heard, or noticed flying under the canopy within a radial distance of 50 meters were recorded. In 13 of the 14 sites, 25 point count surveys were carried out, while at the remaining site, only 18 counts were sampled. To ensure uniform effort at each site, using rarefaction technique, standardized estimates of species richness based on 17 counts were obtained for each site.

Data Analysis:

A total of 58 species were seen at the 14 sites. This was tabulated in the form of a matrix with 58 rows and 14 columns, wherein 1 or 0 was used to indicate the presence or absence of a particular species at a particular site. The matrix so obtained had 812 elements, 393 being equal to 1 and the rest 419 being equal to zero. The species turnover between any pair of sites was quantified using Jacard index and Euclidean distance. For a pair of sites 'i' and 'k', if 'a' denotes the number of species seen in both of them, 'b' denotes the number of species seen in site 'i' but not in site 'k', and 'c' denotes the number of species seen in site 'k' but not in site 'i', then J_{ik} , the Jacard index of similarity between the two sites is equal to $a/(a+b+c)$, and E_{ik} , the Euclidean distance between the two is equal to $b+c$. Using these formulae, values for these indices for every pair of sites (91 pairs in all) were computed. The mean value for the 13 pairs of adjacent sites was also computed separately to quantify the effect of spatial proximity.

A suit of fortran programs was developed to explore the four null models described above using Monte Carlo simulations. At the beginning of a typical simulation, all the 812 elements of the 58X14 matrix were set equal to zero. Next, one out of 58 species was randomly selected (all with equal probability), and one out of 14 sites was randomly selected (again with equal probability). The corresponding element of the matrix was then made equal to 1. This process was continued till all the 393 elements (as present in the observed matrix) were made equal to 1. This simulated matrix was then used for computing the various indices. Each of the four null models was simulated 5000 times.

Simulations for the four null models differed from each other in the ~~exact~~ way in which the 393 elements were chosen. Thus, for the saturation model (wherein the number of species observed at each site was kept constant), once the total of a specific column became equal to that of the observed matrix, that particular column (site) was not chosen for the placement of subsequent 1s. Similarly, for the distribution model, once the row total became equal to the observed row total, that species was omitted from subsequent selections. The double constraint model was also simulated in an analogous manner.

To examine the effect of spatial contiguity, the enhancement in the probability of occurrence of a species next to a filled site compared to that next to an empty site was computed from the observed data. Since the 14 sites are arranged along an altitudinal gradient, the sites at the lowest and the highest elevation have only one neighbouring site, while the other 12 sites have two neighbours each. The two extreme sites thus have only one index of enhancement (probability of the site

harbouring a species given that the neighbouring site also has the same species, divided by the probability of the species being present given that it is NOT present in the neighbouring site). The intermediate sites have two indices of enhancement, depending on whether the species occurs at only one or both the neighbouring sites. In simulating the spatial contiguity model, for the first occurrence of any species, all sites were considered to be equally likely. For subsequent placement of a species, the sites adjacent to filled sites were chosen with a higher probability as given by the indices obtained from the observed data.

It is possible that a simulated matrix may be the same as the observed matrix, except for a permutation of the serial number of species and/or sites. To guard against this possibility, use was made of the fact that the spectrum of inter-site ^{Euclidean} Euclidean distances is invariant with respect to permutation of row and column indices. The ^{Euclidean} Euclidean distance between any pair of sites in the present situation can take any integer value between 0 and 58. The frequency distribution of the 91 pairwise distances was computed for the observed matrix as well as for each of the simulated matrix. Any simulated matrix with distribution identical to the observed one was discarded. As indicated above, the process continued till 5000 simulated matrices were obtained.

5. Results:

For the set of 91 pairwise distances between the 14 sites, the mean value of the Jacard index was 0.5289, and the standard deviation was 0.1403. The corresponding values for the ^{Euclidean} Euclidean distance were 17.9890 and 7.1438. For the 13 pairs of adjacent sites, the mean and standard deviation of Jacard index were 0.6934 and 0.1116 respectively, while for the ^{Euclidean} Euclidean distance, the values were 10.2308 and 4.1664.

Table 1 contains statistical summary of the results of the four models for Jacard's index. It is seen that the saturation model gives a rather poor fit to the data on the mean Jacard index of 91 pairwise values, while the distribution model does considerably better. The remaining two models do equally well. On the other hand, the first three models are unable to reproduce the observed high value of similarity between adjacent sites; in fact, as expected, these models do not make a distinction between adjacent and non-adjacent sites.

The contiguity model incorporates another feature the observed distribution in that in the simulated distribution of a species, adjacent sites are more likely to be occupied. From the data, it is seen that if a next-to-extreme site is occupied by a species, the probability that a species is found at the adjoining extreme site is 3.6 times higher compared to the case when the next-to-extreme site is empty. For the sites in the middle, the enhancement (compared to when both the neighbouring sites are empty) in the probability of occurrence is by a factor of 4.5 is when only one of the neighbouring sites is occupied, and by a factor of 9, when both are occupied. As seen from table 1, the mean value of the Jacard index for the 13 pairs of neighbouring sites obtained from the spatial contiguity model is much closer to the observed value than the other three models.

Table 2 shows the results for the ^{Euclidean} Euclidean distances obtained from the four models. It is interesting to note that the overall mean (of 91 pairs) is an invariant whenever the row totals are kept same as the observed values (as in the distribution model). Here too the spatial contiguity model gives a better fit to the data on mean of 13 pairs of neighbouring sites.

6. Discussion:

The observed pattern of distribution of bird communities along the elevational gradient clearly demonstrates that the neighbouring assemblages are significantly more similar than non-neighbouring ones. None of the four models was able to satisfactorily reproduce the observed similarity between neighbouring assemblages. This suggests that the species are not distributed independently of each other, and that occurrence of one species is likely to positively or negatively affect the occurrence of

other species. This possibility can be explored by extending the present investigation to the patterns of association between species. Another possible refinement would be to modify the contiguity model so as to include the actual distance separating the neighbouring (and non-neighbouring sites) instead of the simpler dichotomous distinction used in the present study. These investigations are currently in progress.

7. Tables and Charts :

Table 1. Mean values of Jacard index between all the 91 pairs of sites, as well as between the 13 pairs of adjacent sites for the four null models described in the text. The observed values are also shown in the table.

	Mean	Std. Dvn.	Maximum	Minimum
All 91 pairwise distances				
Observed mean = 0.5289				
Saturation Model	0.3204	0.0062	0.3448	0.3027
Distribution Model	0.5164	0.0009	0.52	0.513
Doubly Constrained Model	0.5165	0.0007	0.5198	0.5147
Spatial Contiguity Model	0.5181	0.0009	0.5222	0.5156
The 13 pairs of adjacent sites				
Observed mean = 0.6934				
Saturation Model	0.3164	0.0164	0.3783	0.2514
Distribution Model	0.5167	0.0142	0.5702	0.4628
Doubly Constrained Model	0.5148	0.0143	0.5728	0.4676
Spatial Contiguity Model	0.6218	0.0161	0.6793	0.5687

Table 2. Mean values of the Euclidean distance between all the 91 pairs of sites, as well as between the 13 pairs of adjacent sites for the four null models described in the text. The observed values are also shown in the table.

	Mean	Std. Dvn.	Maximum	Minimum
All 91 pairwise distances				
Observed mean = 17.9890				
Saturation Model	28.9914	0.3954	30.1648	27.4396
Distribution Model	17.989	0	17.989	0
Doubly Constrained Model	17.989	0	17.989	17.989
Spatial Contiguity Model	17.989	0	17.989	17.989
The 13 pairs of adjacent sites				
Observed mean = 10.2308				
Saturation Model	28.8853	1.0459	33.1538	25.1538
Distribution Model	17.9752	0.6869	20.4615	15.4615
Doubly Constrained Model	17.8414	0.6859	20.2308	15.1538
Spatial Contiguity Model	13.0264	0.6728	15.3077	10.6923

8. Recommendations : N. A.

9. Publications : Nil

10. Degrees awarded : Nil (The work presented above forms part of one chapter of the thesis submitted by Mr. T. R. Shankar Raman in November 2001 for the award of the Ph.D. degree)