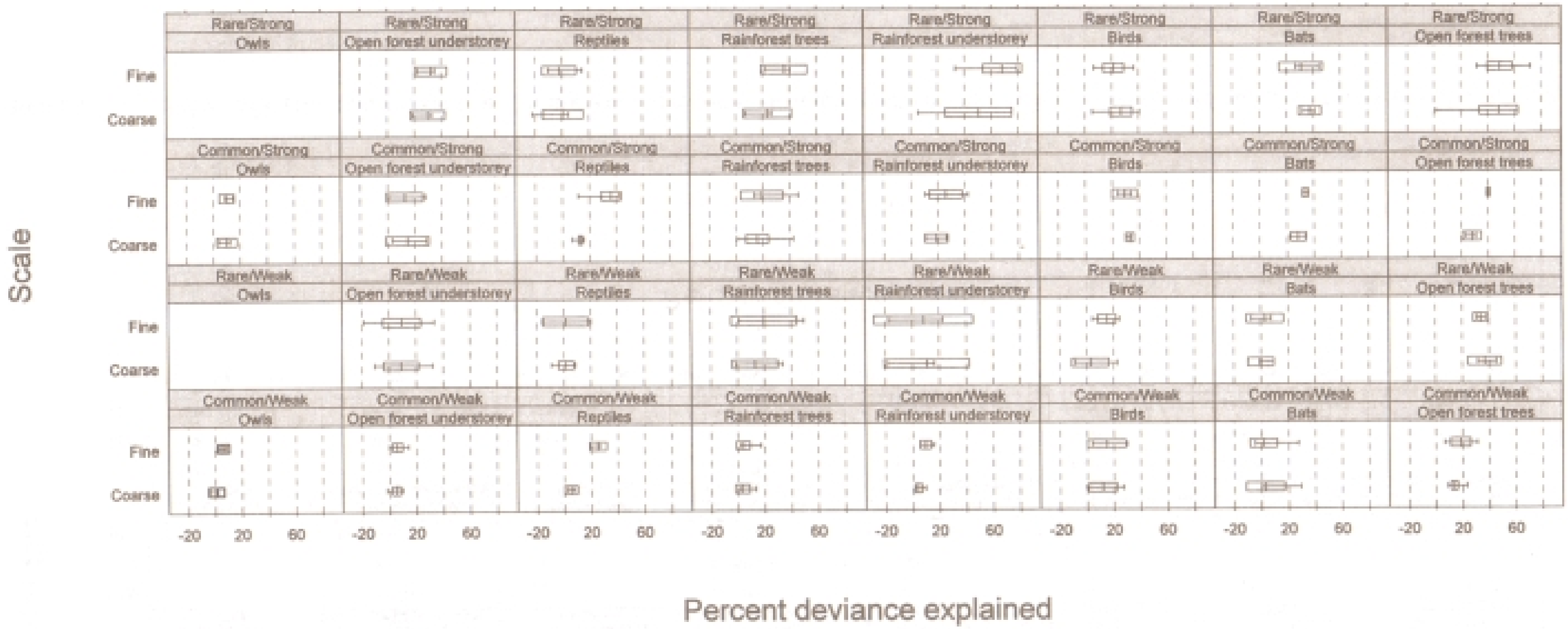
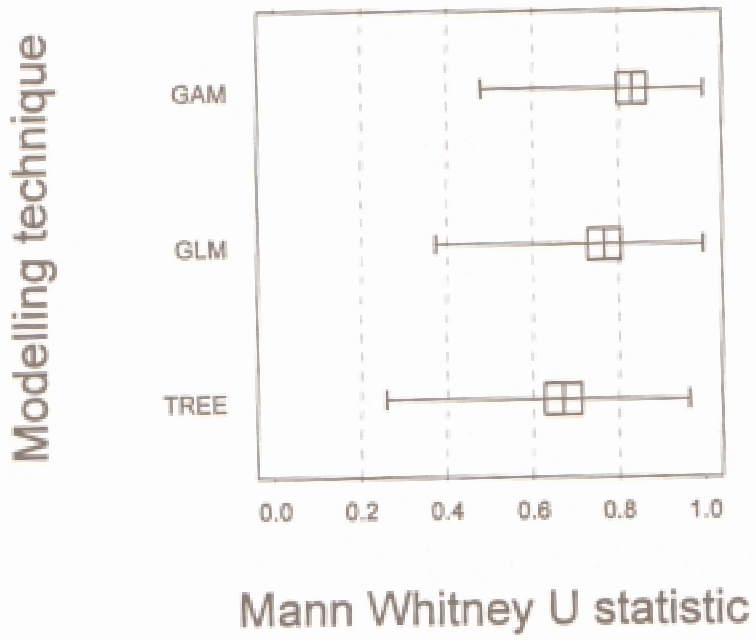


Figure 7.9 Effect of interaction between biological group, species type and scale of environmental data on overall accuracy of presence/absence models.



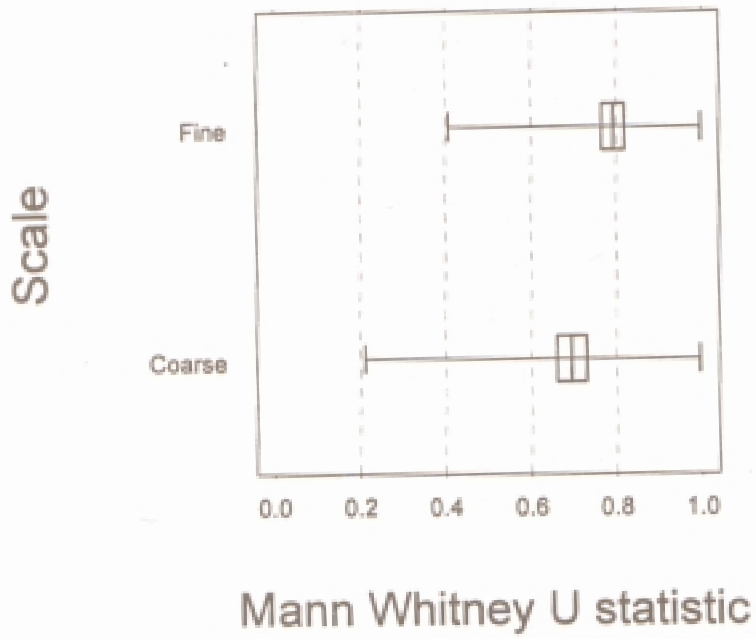
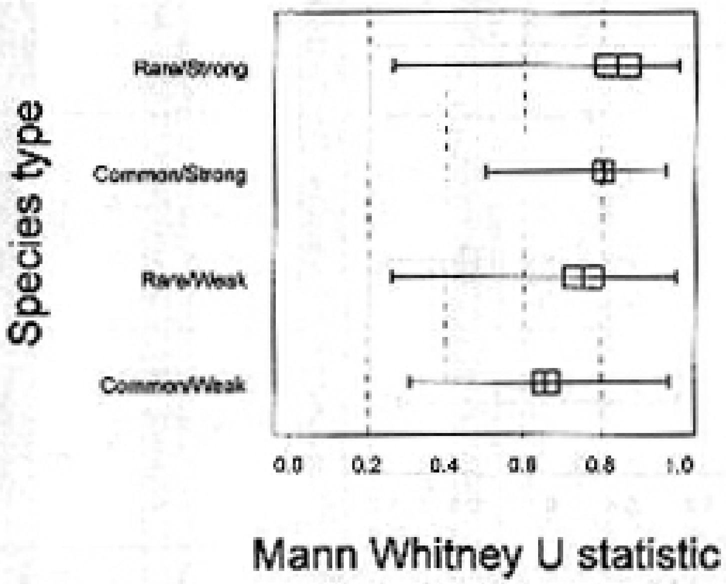


Figure 7.10 Effect of modelling technique and scale of environmental data on discrimination of presence/absence models.



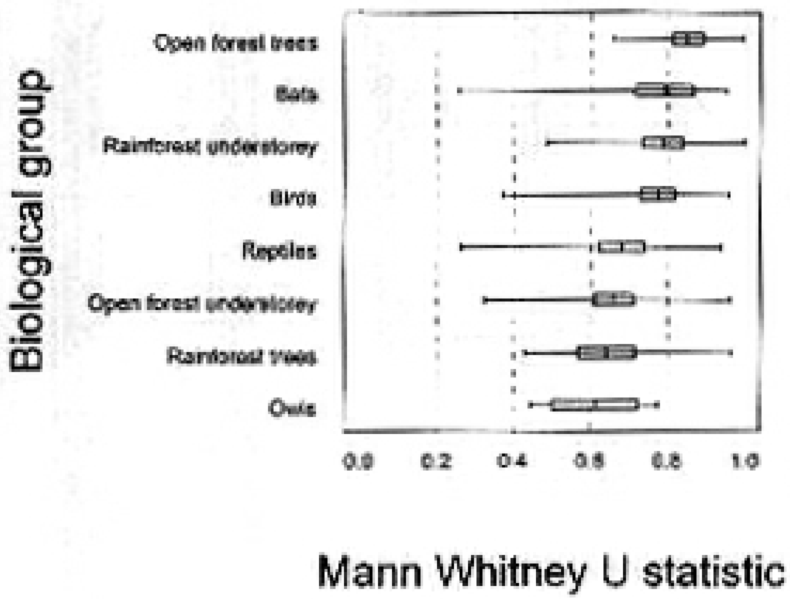


Figure 7.11 Effect of biological group and species type on discrimination of presence/absence models.

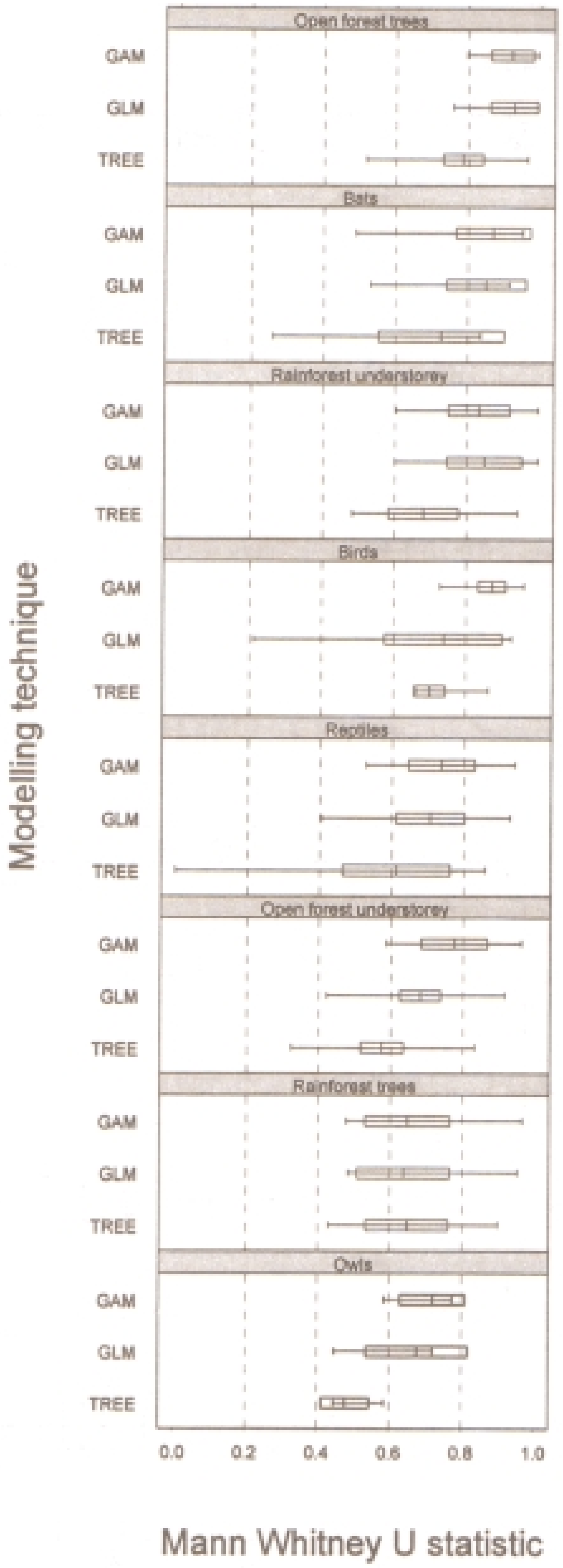


Figure 7.12 Effect of interaction between modelling technique and biological group on discrimination of presence/absence models.

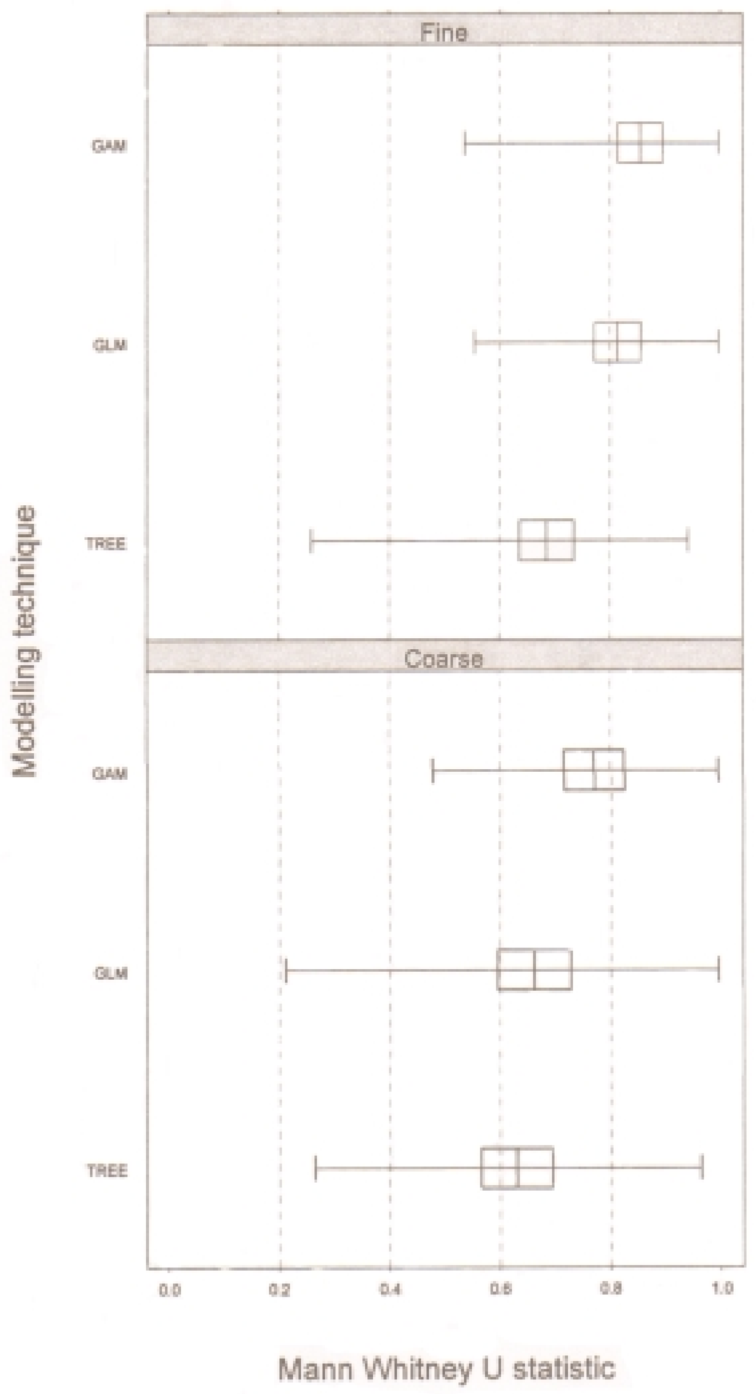


Figure 7.13 Effect of interaction between biological group and scale of environmental data on discrimination of presence/absence models.

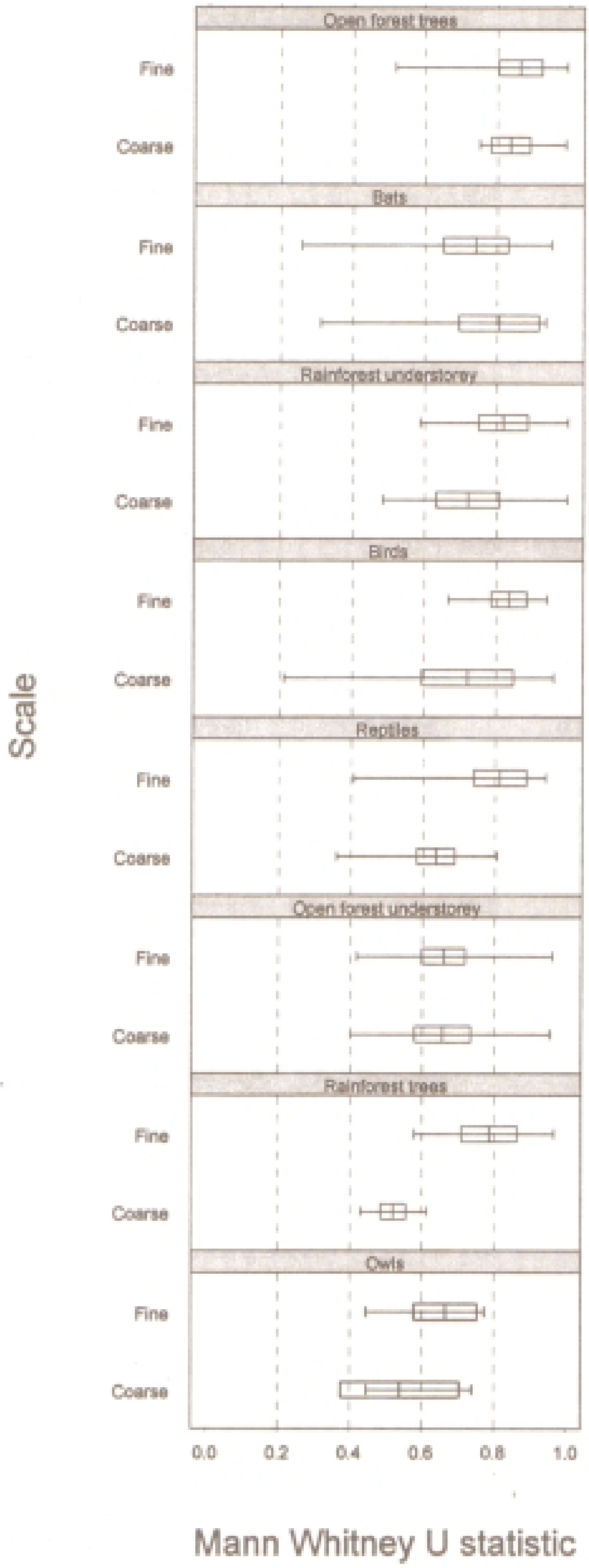


Figure 7.14 Effect of interaction between biological group and scale of environmental data on discrimination of presence/absence models.

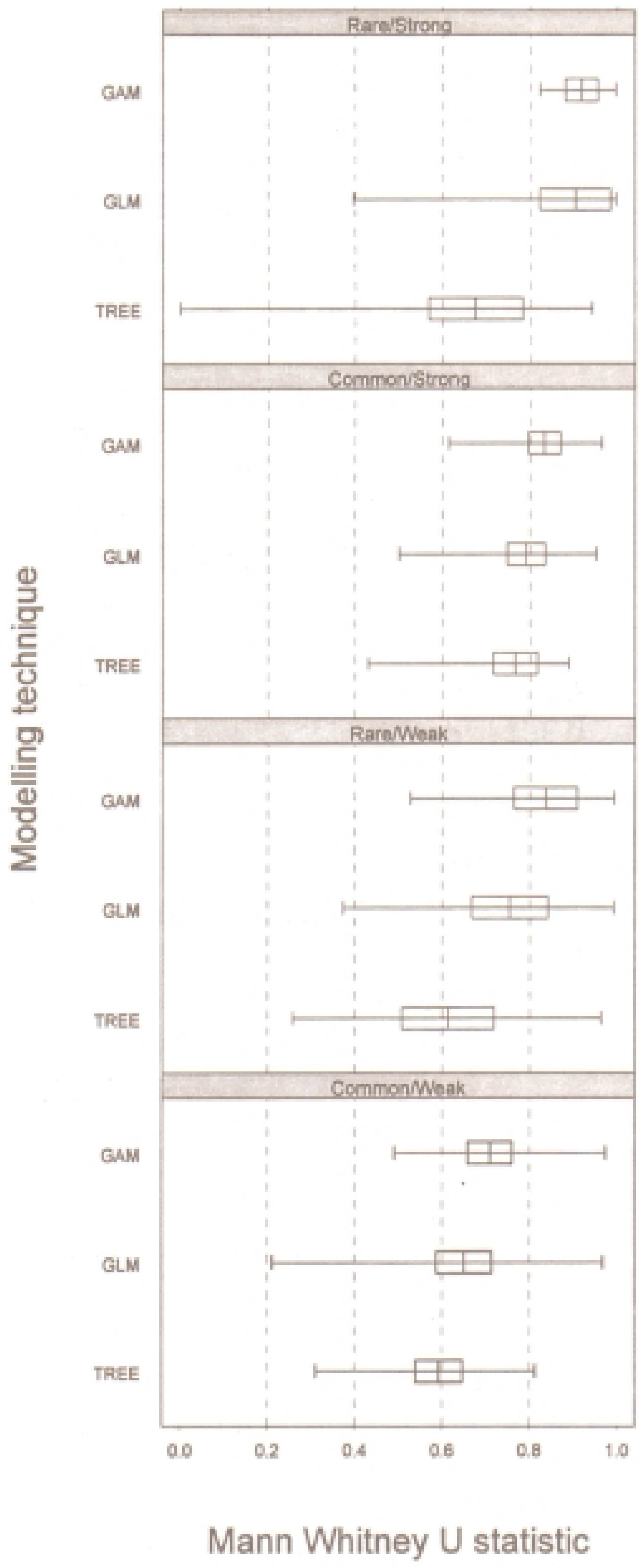


Figure 7.15 Effect of interaction between modelling technique and species type on discrimination of presence/absence models.

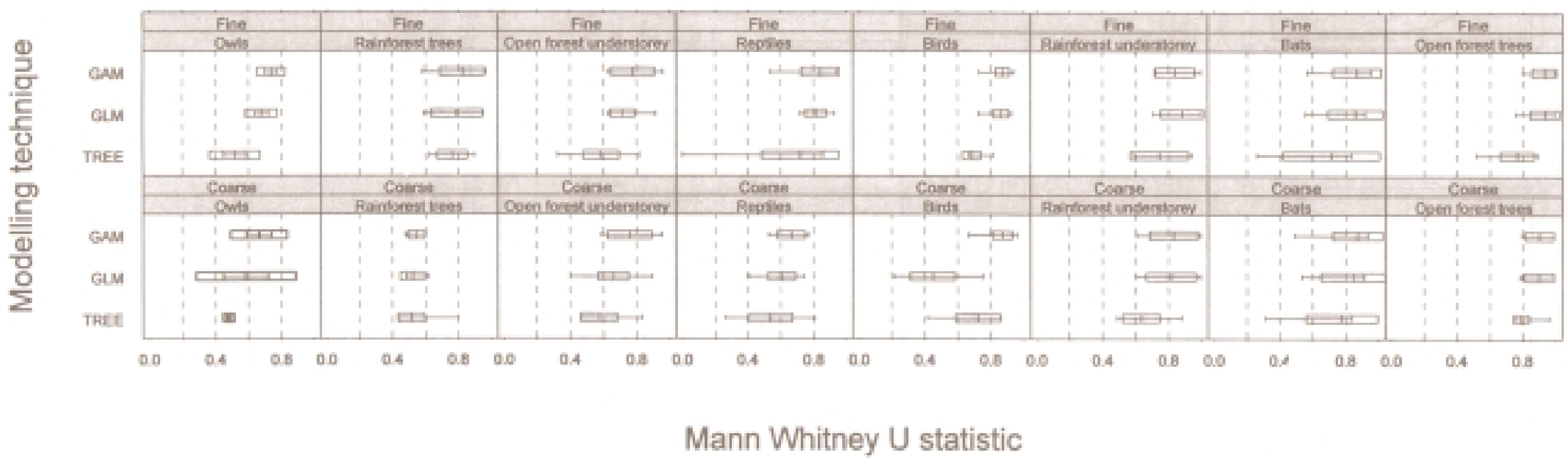


Figure 7.16 Effect of interaction between biological group and scale of environmental data on discrimination of presence/absence models.

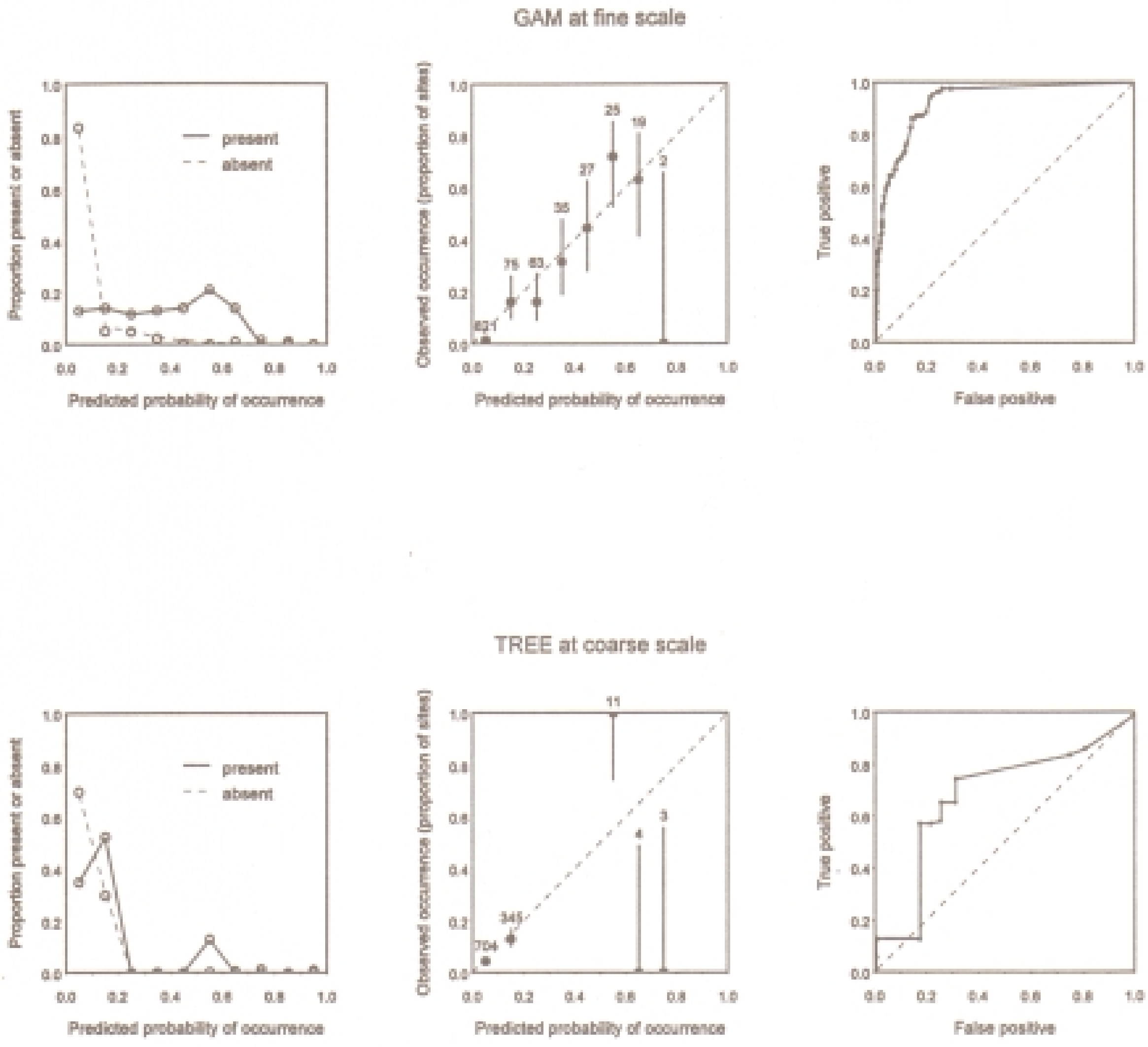


Figure 7.19 An example of the relative performance of a generalised additive model using fine scale environmental data (top graphs) and a decision tree model using coarse scale environmental date (bottom graph) for the burrowing skink *Ophinscincus trancatus.* For further information on the deriviation and interpretation of these graphs see figures 6.1 and 6.2.

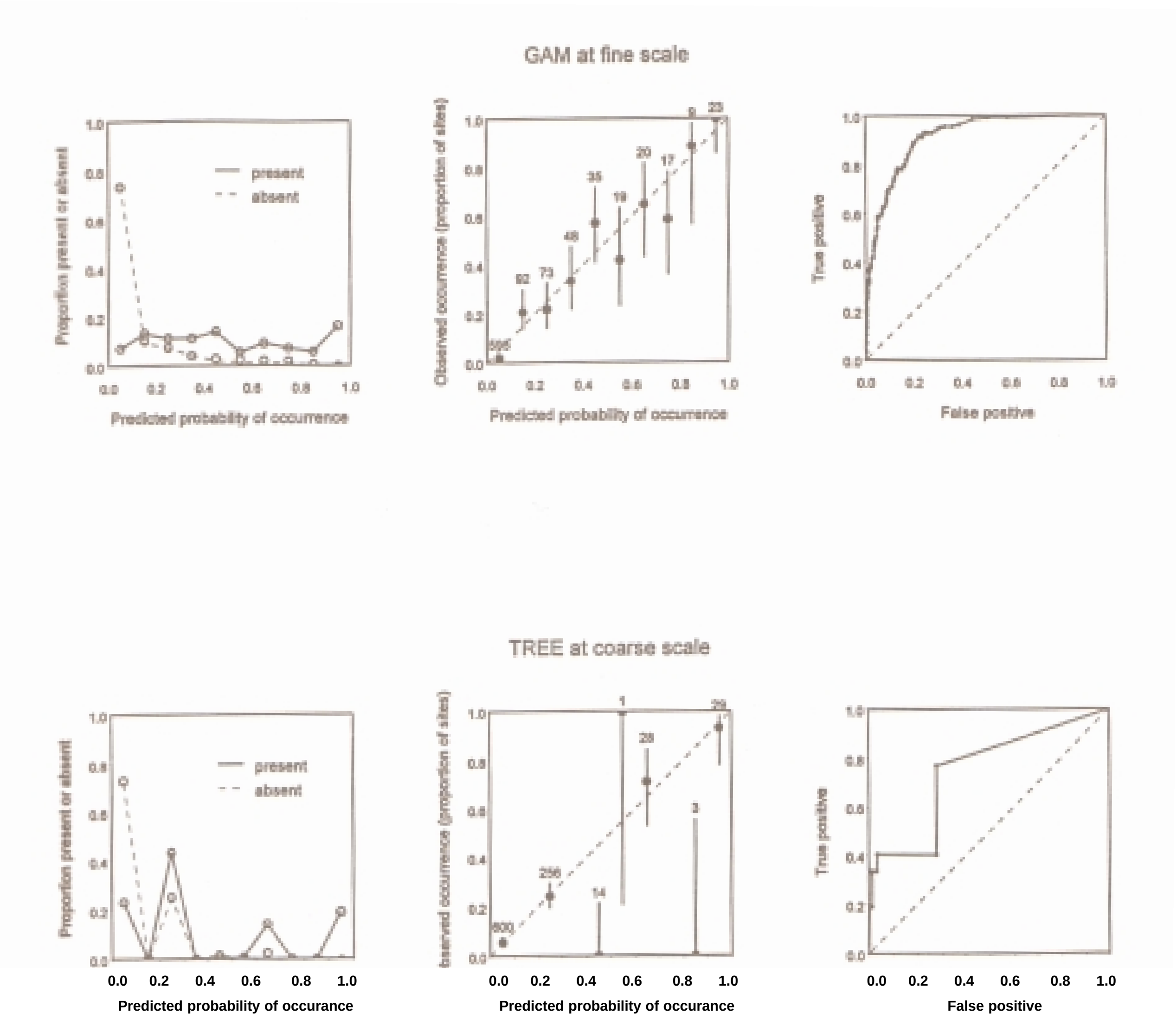


Figure 7.20 An example of the relative performance of a generalised additive model using fine scale environmental data (top graphs) and a decision tree model using coarse scale environmental date (bottom graph) for the tree fern *Dicksonia antarctica.* For further information on the deriviation and interpretation of these graphs see figures 6.1 and 6.2.

rare species tended to perform better than those for common species (within each class of strength of environmental relationship).

Figures 7.19 and 7.20 provide examples of the relative performance of two different modelling approaches (GAM using fine environmental data and TREE using coarse environmental data) applied to two species, using the graphical diagnostic techniques described in Sections 6.4.1 and 6.4.2.

7.5.2 Performance of presence-only modelling

As detailed in Section 6.4 the performance of presence-only models could be measured only in terms of discrimination ability, not in terms of overall accuracy. Table 7.14 presents raw results obtained using the Mann-Whitney measure of discrimination for all evaluated combinations of modelling technique, species and scale of environmental data.

The results from Table 7.14 were subjected to an ANOVA to test the effect of different factors on the performance of presence-only models (Table 7.15). The variables used in this ANOVA were almost identical to those employed in the analysis of presence/absence models in Section 7.5.1. The only difference was the addition of an extra class, BIOCLIM, to the modelling technique variable.

Table 7.15 indicates that modelling technique had a highly significant ($p < 0.001$) effect on the performance of models. The effects of biological group and species type were also significant ($p < 0.01$). In contrast to the analysis of presence/absence models, scale of environmental data did not have a significant ($p > 0.05$) effect on the performance of presence-only models. A number of interaction terms were significant (some at the $p < 0.001$ level) suggesting that the effects of modelling technique, biological group and species type were not consistent across all combinations of these variables.

The main effects are depicted graphically in Figures 7.21 and 7.22. Models derived using GAM and GLM performed significantly better than models derived using BIOCLIM and TREE. As was the case for presence/absence modelling, species with strong environmental relationships performed better than species with weak relationships and rare species performed better than common species. The relative importance, however, of these differences was not the same as for presence/absence modelling. The rare/common dichotomy appears to have had a greater effect on performance than the weak/strong dichotomy.

Interaction effects significant at the $p < 0.001$ level are illustrated in Figures 7.23 to 7.25.

7.5.3 Comparison of presence/absence and presence-only modelling

The final analysis conducted on the modelling results compared the discriminatory performance of presence/absence and presence-only models. The analysis used GAM, GLM and TREE models for all species.

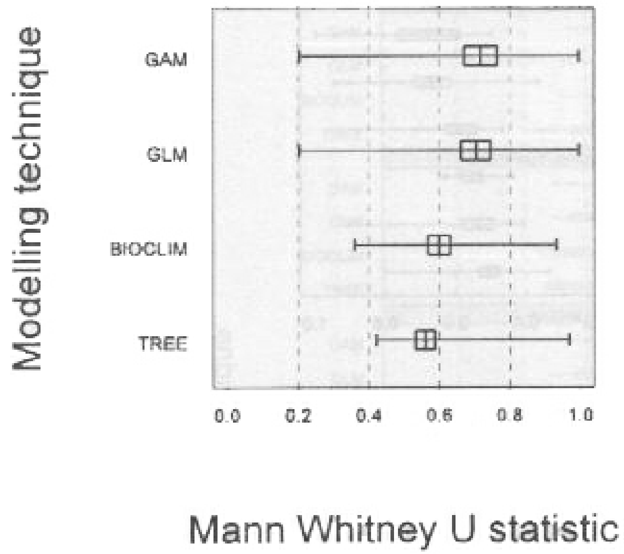
The ANOVA table for the analysis is presented in Table 7.17. Note that a new variable, data type, was included in the ANOVA. This variable has two classes; presence/absence and presence-only. The effect of data type on model performance was highly significant ($p < 0.001$). Presence/absence models performed significantly better than presence-only models (Figure 7.28). All other main effects (modelling technique, scale, biological group, species type) were also significant. These effects are depicted graphically in Figures 7.28 and 7.29. A large number of interaction terms were significant, but only those involving data type are presented graphically in Figures 7.31 to 7.34.

Table 7.14 Discrimination (Mann-Whitney statistic) achieved by presence-only models, derived using different modelling techniques and coarse and fine scale environmental data.

Species	BIOCLIM coarse	BIOCLIM fine	GAM coarse	GAM fine	GLM coarse	GLM fine	TREE coarse	TREE fine
Bats								
Chalinolobus gouldii	0.581	0.573	0.500	0.500	0.551	0.520	0.502	0.520
Falsistrellus tasmaniensis	0.670	0.699	0.705	0.811	0.688	0.706	0.514	0.674
Kerivoula papuensis	0.463	0.481	0.557	0.519	0.622	0.556	0.500	0.500
Nyctophilus bifax	0.796	0.811	0.867	0.922	0.883	0.917	0.775	0.858
Nyctophilus gouldi	0.558	0.552	0.500	0.500	0.507	0.500	0.481	0.511
Vespadelus darlingtoni	0.657	0.682	0.710	0.728	0.710	0.719	0.553	0.581
Vespadelus vulturnus	0.511	0.509	0.593	0.509	0.734	0.810	0.500	0.507
Nocturnal birds								
Powerful owl	0.546	0.694	0.458	0.453	0.459	0.436	0.539	0.543
Sooty owl	0.557	0.630	0.579	0.572	0.576	0.565	0.530	0.570
Diurnal birds								
Rose-crowned fruit dove	0.726	0.655	0.634	0.843	0.816	0.658	0.805	0.856
Glossy-black cockatoo	0.647	0.625	0.550	0.617	0.566	0.597	0.525	0.681
Superb lyrebird	0.769	0.743	0.835	0.852	0.789	0.795	0.672	0.727
Varied triller	0.742	0.720	0.804	0.803	0.737	0.813	0.695	0.557
Spectacled monarch	0.613	0.596	0.615	0.725	0.644	0.714	0.709	0.582
Olive whistler	0.910	0.780	0.923	0.921	0.927	0.935	0.762	0.964
Scarlet honeyeater	0.650	0.611	0.610	0.643	0.955	0.625	0.619	0.541
Forest raven	0.582	0.453	0.813	0.781	0.754	0.791	0.560	0.624
[illegible] reptiles								
Cacophis krefftii	0.611	0.555	0.812	0.870	0.790	0.814	0.661	0.758
Calyptotis scutirostrum	0.507	0.533	0.762	0.570	0.570	0.559	0.673	0.504
Coeranoscincus reticulatus	0.521	0.904	0.915	0.937	0.936	0.935	0.774	0.926
Egernia mcpheei	0.449	0.662	0.743	0.669	0.671	0.538	0.601	0.717
Eulamprus murrayi	0.533	0.686	0.800	0.608	0.725	0.508	0.655	0.528
Ophioscincus truncatus	0.597	0.559	0.923	0.841	0.901	0.868	0.795	0.562
Pseudechis porphyriacus	0.391	0.503	0.360	0.636	0.480	0.552	0.510	0.500
Saltuarius swaini	0.496	0.761	0.656	0.621	0.613	0.527	0.729	0.725
Open-forest trees								
Eucalyptus fastigata	0.712	0.704	0.978	0.983	0.582	0.963	0.549	0.695
Angophora costata	0.788	0.811	0.917	0.903	0.844	0.828	0.644	0.963
Corymbia gummifera	0.614	0.527	0.742	0.640	0.701	0.708	0.444	0.505
Eucalyptus carnea	0.765	0.672	0.657	0.642	0.640	0.659	0.597	0.537
Eucalyptus campanulata	0.559	0.506	0.551	0.821	0.532	0.731	0.545	0.543
Corymbia intermedia	0.739	0.719	0.800	0.805	0.792	0.763	0.620	0.619
Eucalyptus blakelyi	0.660	0.525	0.957	0.972	0.971	0.953	0.655	0.589
Eucalyptus nova-anglica	0.895	0.770	0.981	0.968	0.974	0.973	0.738	0.794
Open-forest understorey								
Cassinia quinquefaria	0.662	0.602	0.824	0.624	0.902	0.824	0.464	0.423
Poa sieberiana	0.667	0.559	0.732	0.776	0.731	0.720	0.570	0.519
Lepidosperma laterale	0.512	0.557	0.500	0.500	0.475	0.403	0.512	0.481
Glycine clandestina	0.556	0.527	0.512	0.518	0.520	0.516	0.540	0.511
Eustrephus latifolius	0.175	0.502	0.500	0.505	0.500	0.500	0.481	0.545
Imperata cylindrica	0.510	0.528	0.500	0.500	0.500	0.500	0.560	0.500
Marsdenia liisae	0.808	0.601	0.792	0.792	0.792	0.792	0.661	0.755
Acrotriche aggregata	0.729	0.700	0.759	0.887	0.405	0.435	0.610	0.585
Rainforest trees								
Cupaniopsis anacardioides	0.495	0.492	0.204	0.204	0.204	0.204	0.485	0.494
Diploglottis australis	0.510	0.527	0.500	0.555	0.483	0.445	0.539	0.514
Heritiera actinophylla	0.498	0.509	0.500	0.500	0.495	0.500	0.500	0.530
Schizomeria ovata	0.173	0.542	0.500	0.500	0.627	0.627	0.521	0.615
Syzygium luehmannii	0.434	0.490	0.335	0.755	0.715	0.758	0.480	0.451
Toona ciliata	0.501	0.496	0.553	0.513	0.535	0.648	0.500	0.530
Alectryon subdentatus	0.730	0.781	0.794	0.798	0.792	0.768	0.731	0.819
Rainforest understorey								
Desmodium acanthocladum	0.801	0.922	0.967	0.774	0.992	0.932	0.772	0.903
Alocasia brisbanensis	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.505
Elatostema reticulatum	0.525	0.496	0.574	0.580	0.590	0.550	0.500	0.510
Cyathea leichhardtiana	0.585	0.591	0.667	0.654	0.716	0.695	0.525	0.505
Dicksonia antarctica	0.555	0.535	0.869	0.738	0.714	0.698	0.551	0.554
Tasmannia purpurascens	0.845	0.581	0.955	0.795	0.992	0.957	0.589	0.569
Corokia whiteana	0.836	0.744	0.961	0.756	0.970	0.969	0.739	0.732

Table 7.15 ANOVA results from analysis testing effect of modelling technique (GLM, GAM, TREE, BIOCLIM), scale of environmental data, biological group and species type on discrimination (Mann-Whitney statistic) of presence-only models.

	df	SS	MS	F.value	P.value
Error: Species					
Biol. group	7	2.137176	0.3053108	4.530877	0.00222542
Species type	3	**1.262183**	**0.4207277**	**6.243689**	**0.00259404**
Biol. group * Species type	19	2.720920	0.1432063	2.125212	0.03914577
Residuals	25	1.684612	0.0673845		
Error: Within					
Modelling technique	3	0.633070	0.2110233	46.69531	0.0000000
Scale	1	**0.009754**	**0.0097541**	**2.15839**	**0.1429884**
Modelling technique * Biol. group	21	0.500908	0.0238528	5.27815	0.0000000
Modelling technique * Biol. group * Scale	31	0.310739	0.0100239	2.21808	0.0004015
Modelling technique * Biol. group * Species type	66	**0.813707**	**0.0123289**	**2.72814**	**0.0000000**
Residuals	263	1.188538	0.0045192		



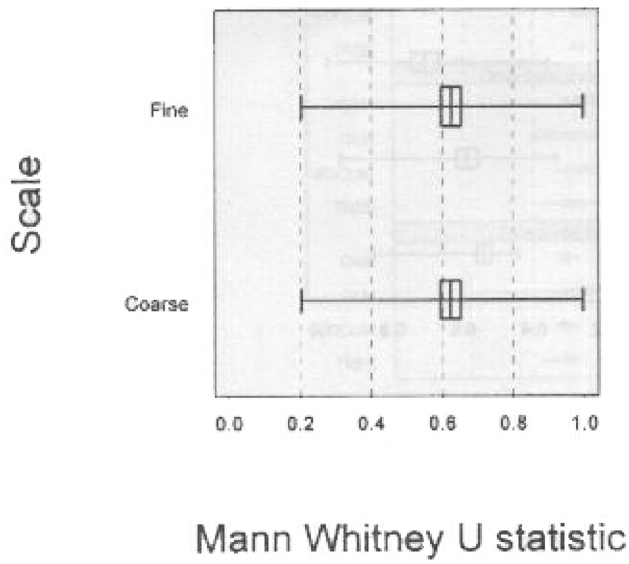


Figure 7.21 Effect of modelling technique and scale of environmental data on discrimination of presence-only models.

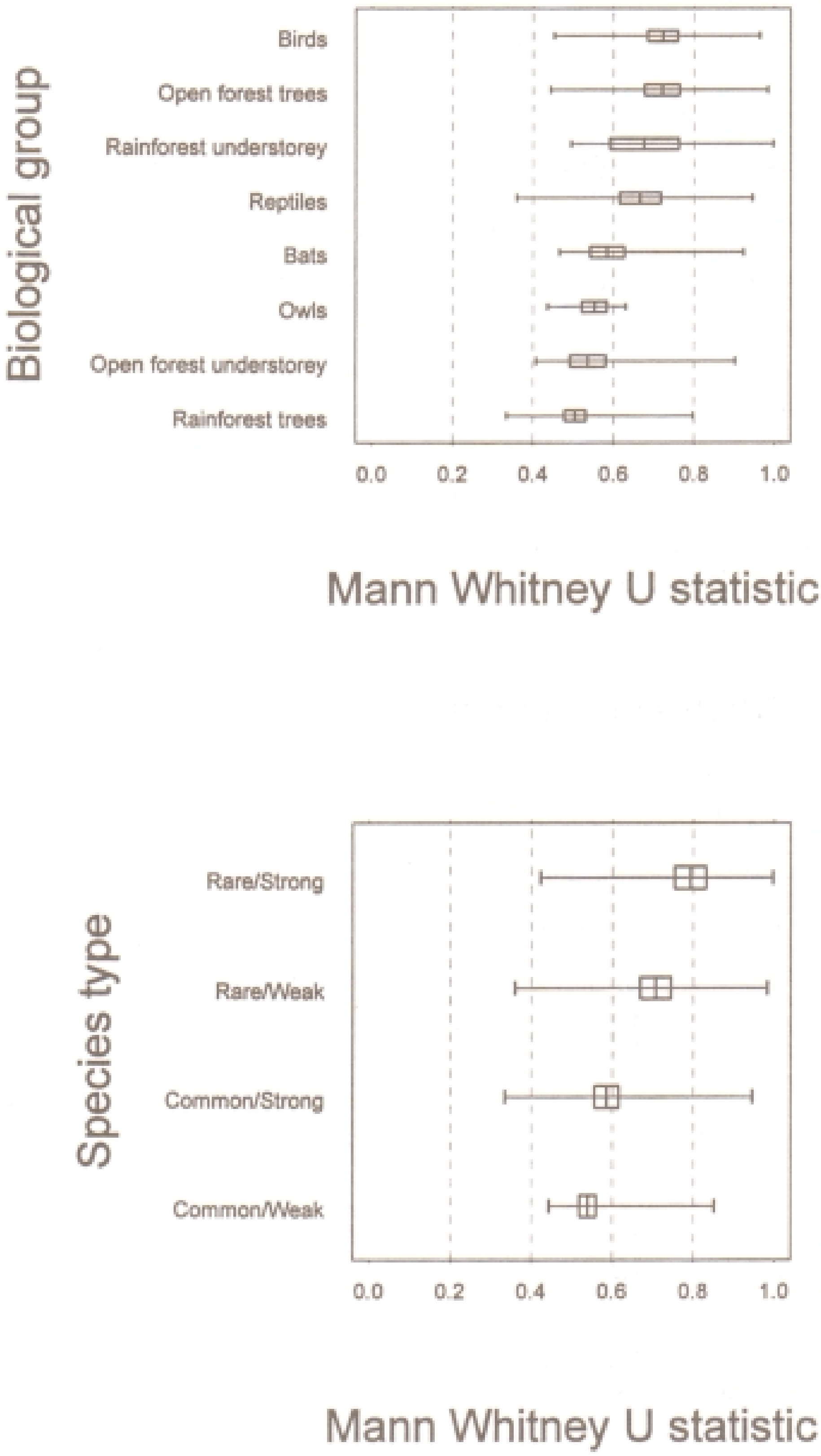


Figure 7.22 Effect of biological group and species type on discrimination of presence only models.

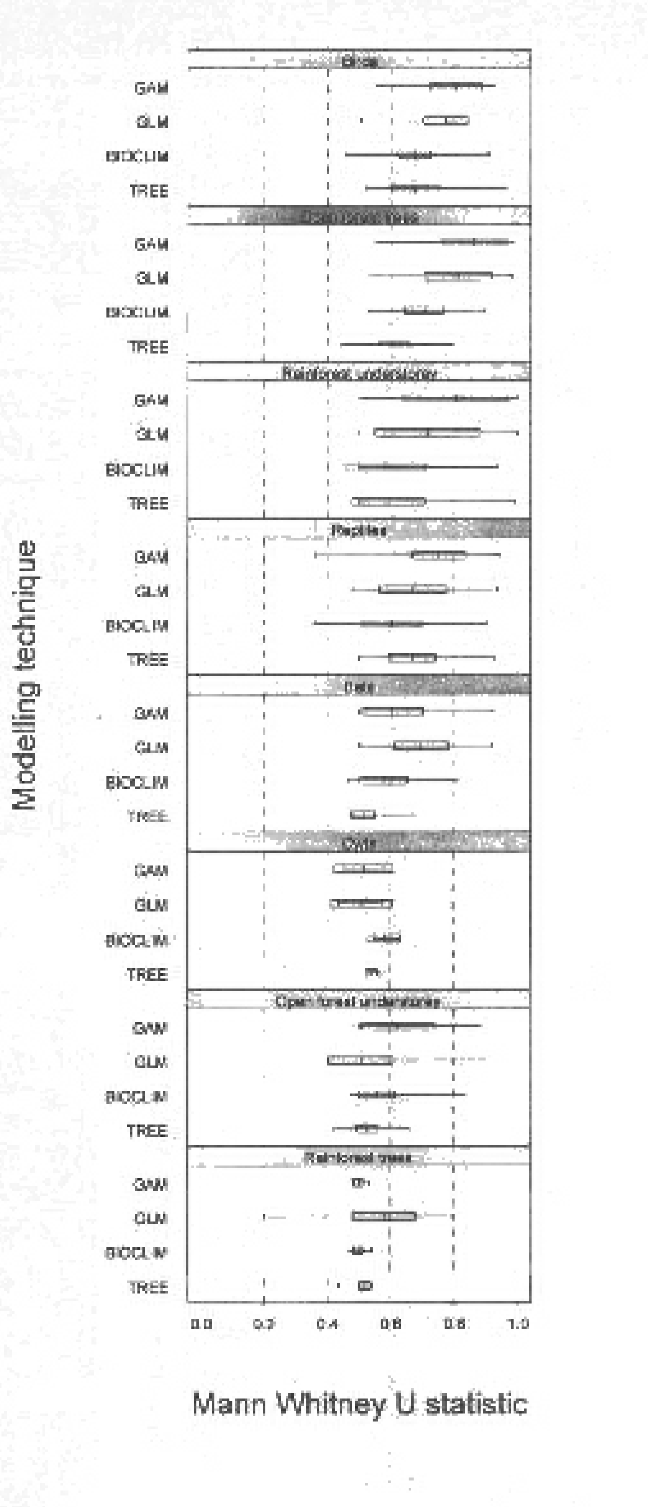


Figure 7.23 Effect of interaction between modelling technique and biological group on discrimination of presence only models.

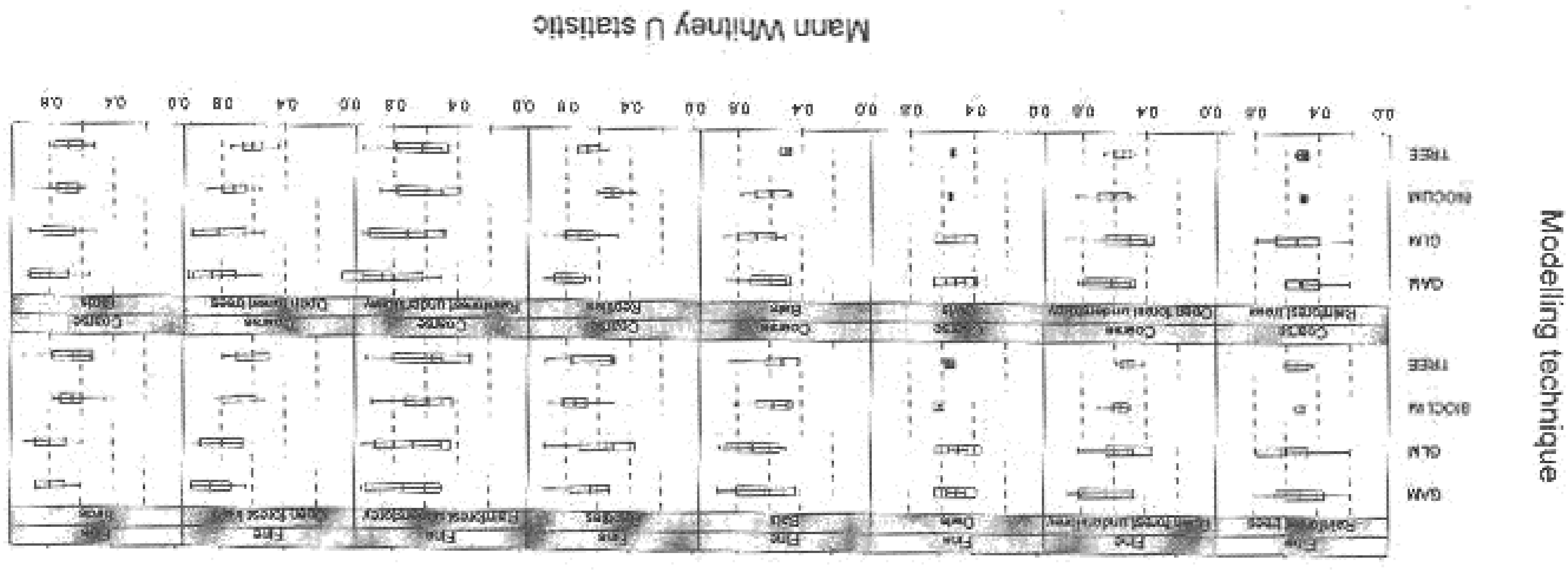


Figure 7.24 Effect of interaction between modelling technique, biological group and scale of environmental data on discrimination of presence-only models.

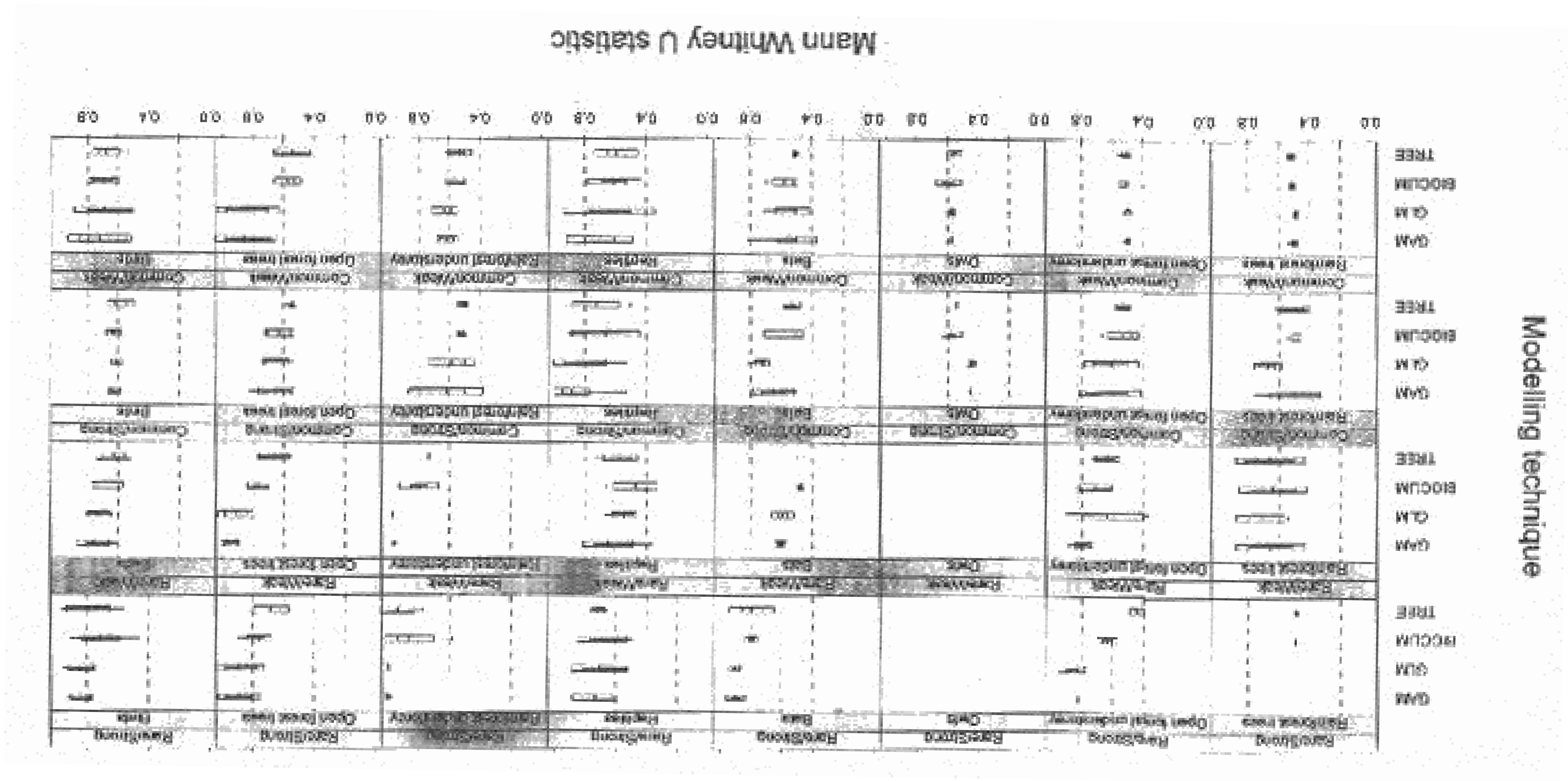


Figure 7.25 Effect of interaction between modelling technique, biological group and species type on discrimination of presence-only models.

Table 7.17 ANOVA results from analysis testing effect of data type (presence-only, presence/absence), modelling technique (GLM, GAM, TREE), scale of environmental data, biological group and species type on discrimination (Mann-Whitney statistic) of models.

	df	SS	MS	F.value	P.value
Error: Species					
Biol. group	7	3.061369	0.4373384	4.738471	0.00169216
Species type	3	1.491397	0.4971325	5.386328	0.00533475
Biol. group * Species type	19	3.271439	0.1721810	1.865546	0.07204039
Residuals	25	2.307381	0.0922952		
Error: Within					
Modelling technique	2	1.605321	0.8026606	125.3174	0.00000000
Data type	1	0.642578	0.6425783	100.3242	0.00000000
Scale	1	0.293255	0.2932550	45.7852	0.00000000
Modelling technique * Data type	2	0.102995	0.0514973	8.0401	0.00037014
Modelling technique * Biol. group	14	0.462707	0.0330505	5.1601	0.00000000
Data type * Biol. group	7	0.456669	0.0652384	10.1855	0.00000000
Modelling technique * Scale	2	0.050807	0.0254035	3.9662	0.01960292
Data type * Scale	1	0.183765	0.1837647	28.6907	0.00000013
Biol. group * Scale	7	0.284155	0.0405936	6.3378	0.00000040
Modelling technique * Species type	6	0.183136	0.0305227	4.7654	0.00010015
Data type * Species type	3	0.454311	0.1514369	23.6435	0.00000000
Modelling technique * Data type * Biol. group	14	0.247075	0.0176482	2.7554	0.00061670
Modelling technique * Data type * Scale	2	0.076658	0.0383291	5.9842	0.00272024
Modelling technique * Biol. group * Scale	14	0.191877	0.0137055	2.1398	0.00924219
Data type * Biol. group * Scale	7	0.289823	0.0414033	6.4642	0.00000028
Modelling technique * Data type * Species type	6	0.153087	0.0255145	3.9835	0.00067374
Modelling technique * Biol. group * Species type	38	0.663905	0.0174712	2.7277	0.00000047
Data type * Biol. group * Species type	19	0.688235	0.0362229	5.6554	0.00000000
Data type * Scale * Species type	3	0.052600	0.0175335	2.7375	0.04304603
Residuals	456	2.920689	0.0064050		

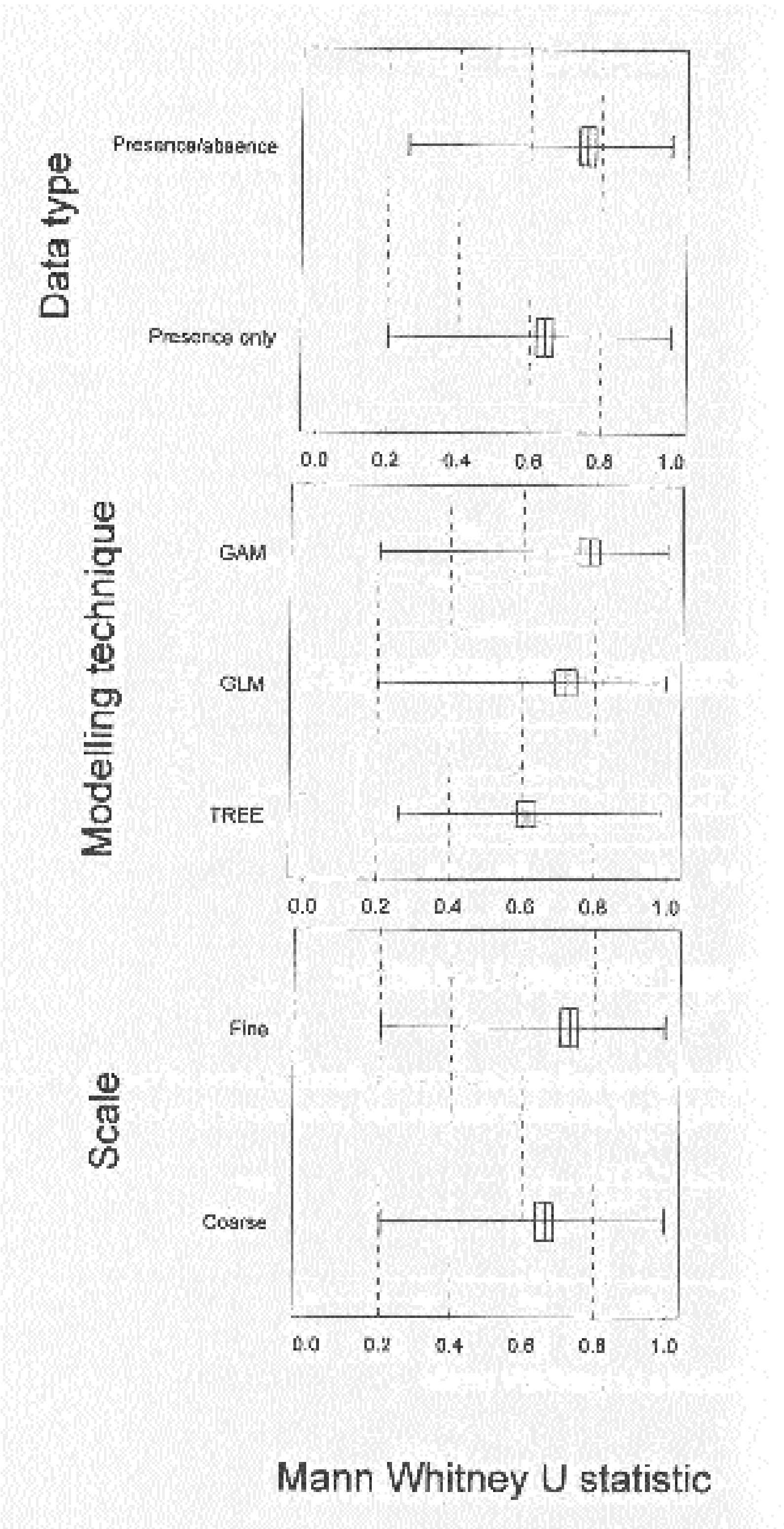


Figure 7.28 Effect of data type, modelling technique and scale of environmental data on discrimination of models (based on combined analysis of presence/absence and presence-only models).

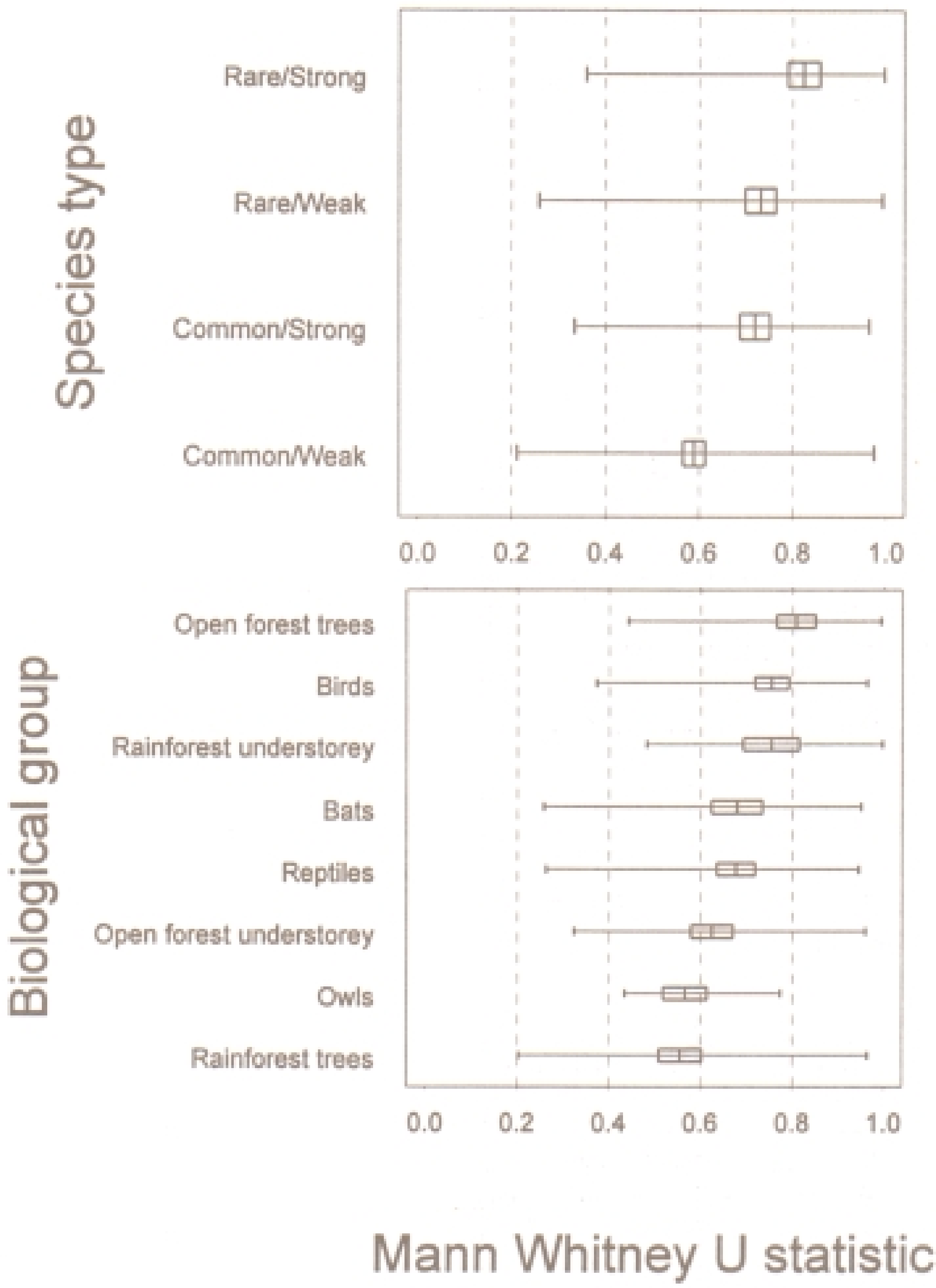


Figure 7.29 Effect of biological group and species type on discrimination of models (based on combined analysis of presence/absence and presence-only models).

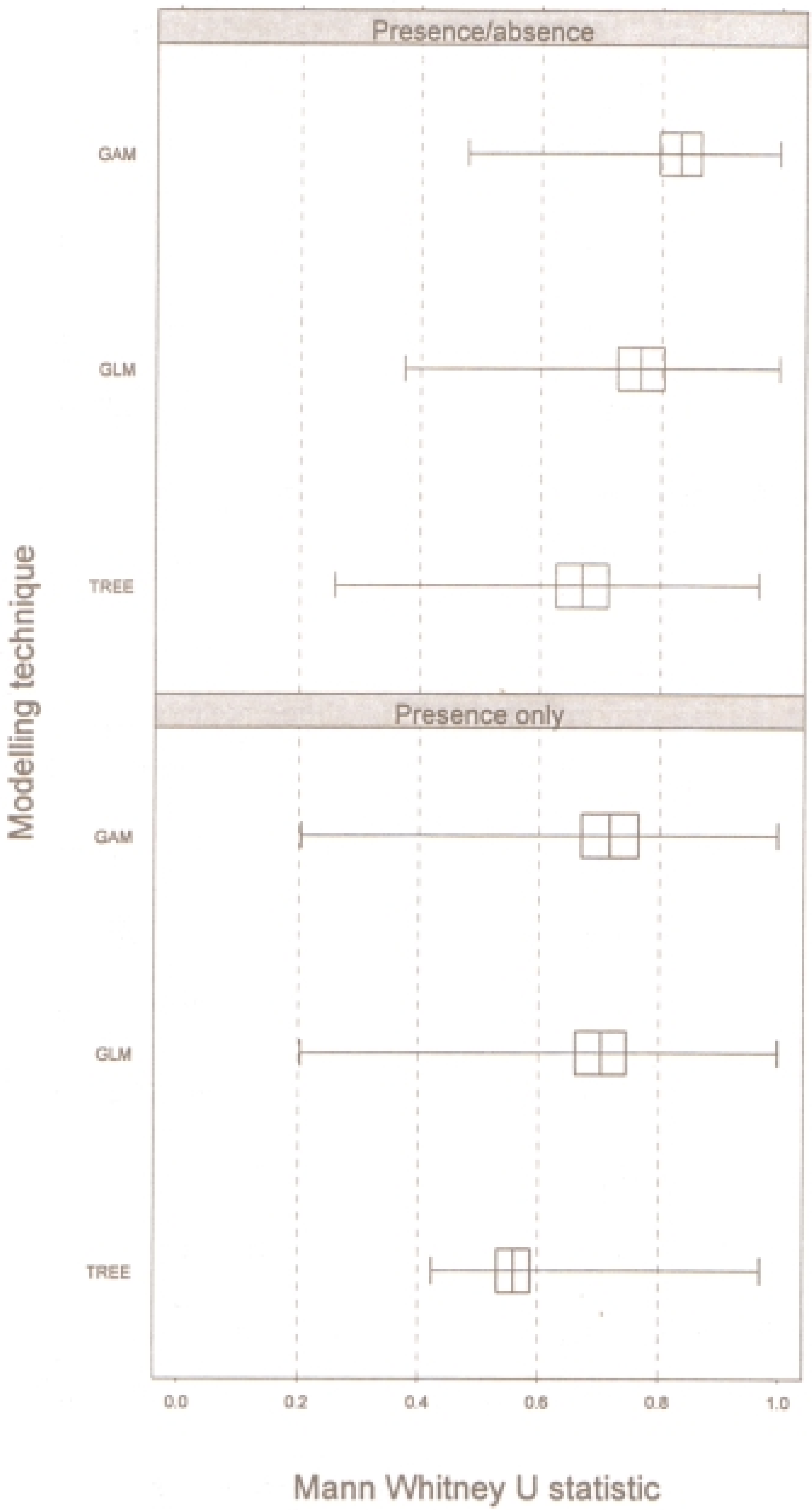


Figure 7.31 Effect of interaction between data type and modelling technique on discrimination of presence only models.

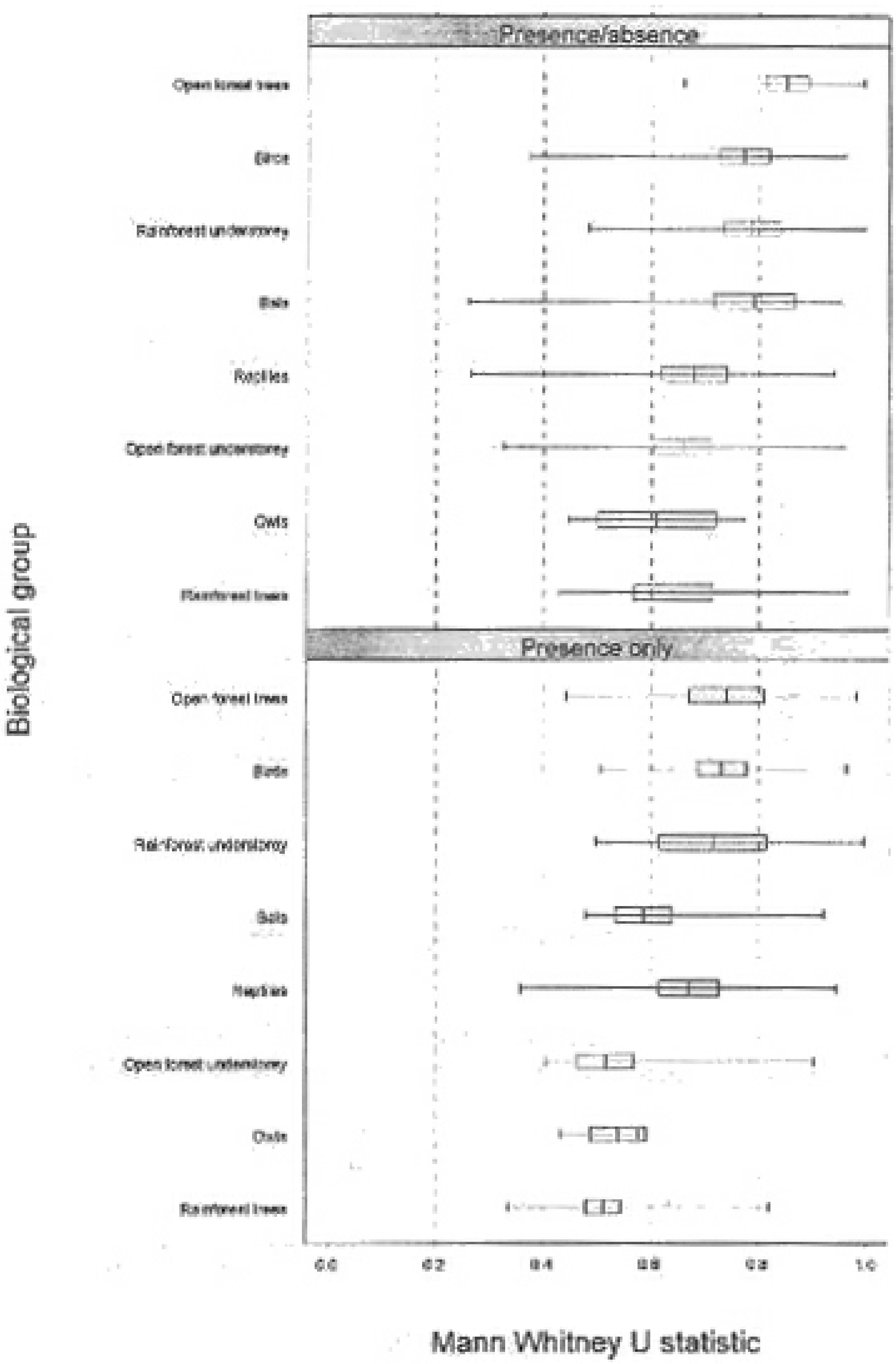


Figure 7.32 Effect of interaction between data type and biological group on discrimination of models.

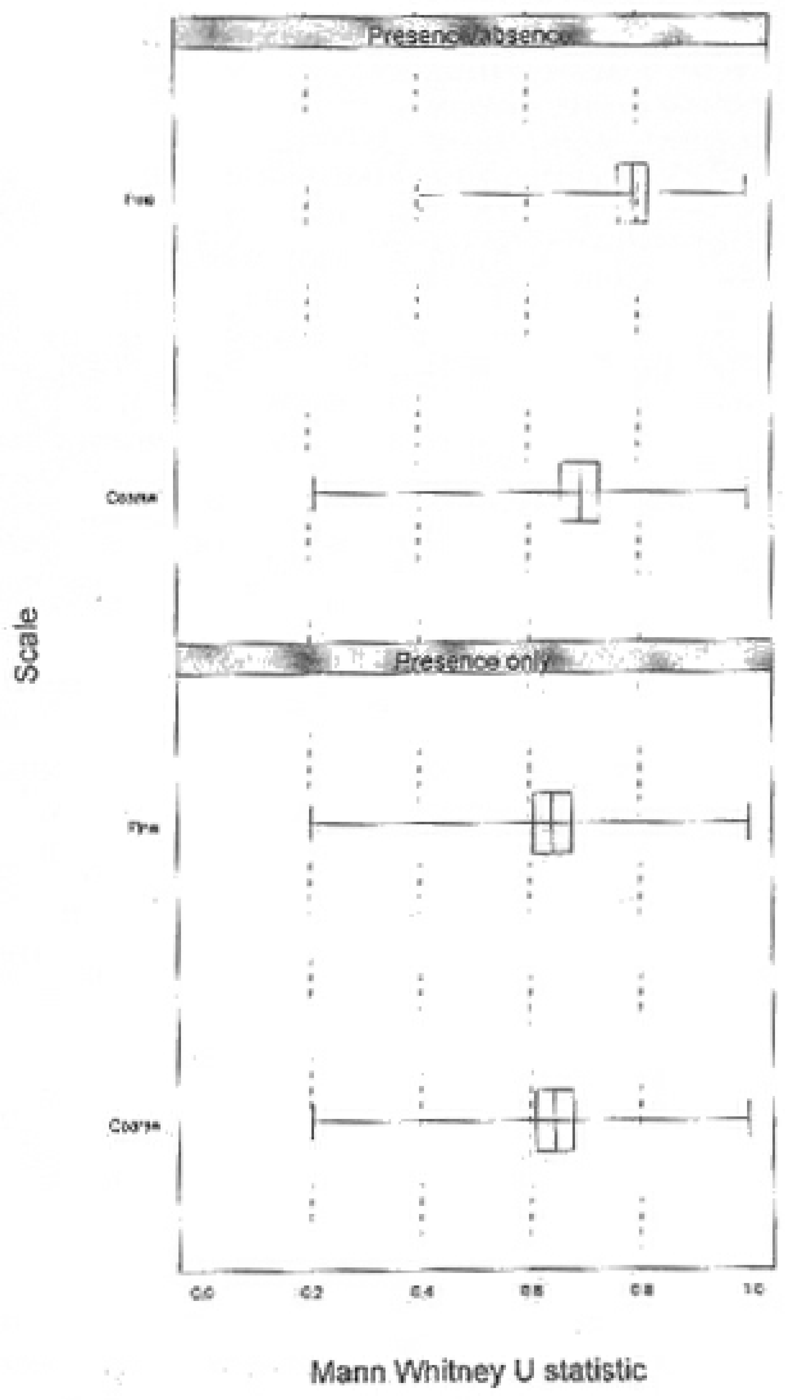


Figure 7.33 Effect of interaction between data type and scale of environmental data on discrimination of models.

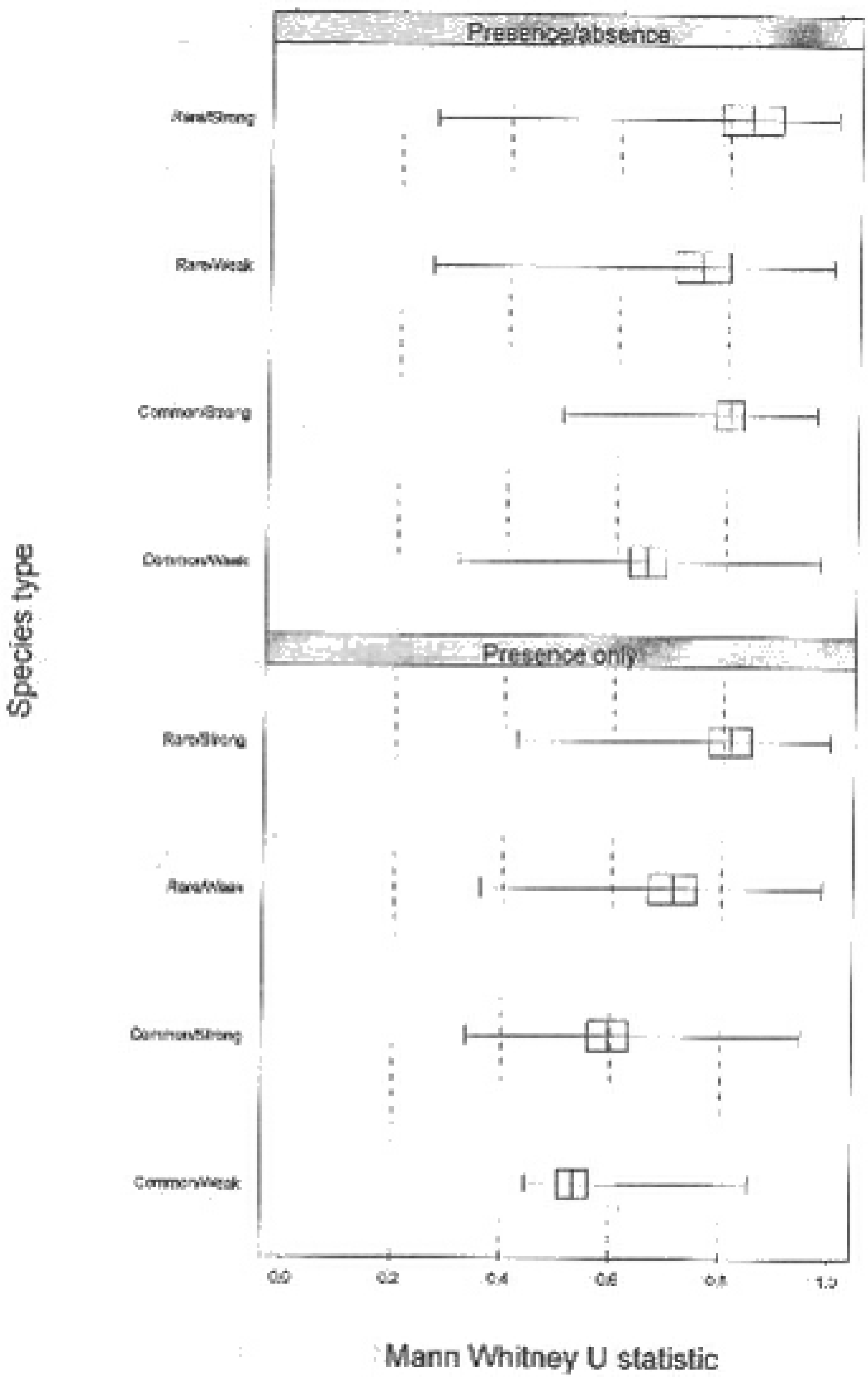


Figure 7.34 Effect of interaction between data type and species type on discrimination of models.

7.6 Discussion

The relative performance of GAM, GLM and TREE in modelling presence/absence data from north east NSW closely matches results obtained by Austin *et al*. (1995), who evaluated these techniques using simulated ecological data. The results presented in this chapter indicate, however, that choice of modelling technique is only one of a number of factors affecting the performance of presence/absence models. These factors include the spatial resolution of environmental predictors and characteristics of the species being modelled. Models based on fine scale environmental data perform better than those based on coarse data. Common species do not necessarily produce better performing models than rare species. What seems to be more important is the strength of relationship between a species and the environmental predictors employed. Related consultancies performed by Austin *et al*. (1995) and Pearce and Ferrier (1996) provide further information on the effects of variation in quality and quantity of environmental and biological data on the performance of presence/absence models.

The evaluation of presence-only modelling in north east NSW produced some interesting differences compared to presence/absence modelling. GAM did not perform significantly better than GLM using presence-only data, but both GAM and GLM outperformed TREE and BIOCLIM. A possible reason for GAM outperforming GLM when using presence/absence data but not when using presence-only data is that the nonparametric curve fitting procedures employed in GAM are more likely to fit spurious response functions to poor quality (biased) data such as that contained in presence-only datasets than to high quality data generated by presence/absence surveys. This is consistent with the findings of Pearce and Ferrier (1996).

The scale of environmental data used to derive presence-only models did not have a significant effect on the performance of those models. Presence-only models derived from coarse environmental data (5km grid) performed as well as models derived from fine data (200m grid). This is probably because presence-only data derived from sources other than systematic field surveys are likely to be less accurately georeferenced than systematic presence/absence data. Improving the resolution of environmental data will not necessarily improve the performance of presence-only models if biological data are not also georeferenced at a comparably fine resolution.

Another interesting result from the evaluation of presence-only models was that models for rare species generally performed better than those for common species. The most likely explanation for this is that records in presence-only datasets are biased towards rarer species. There is a natural tendency for observers to record opportunistic detections of rarer and/or more interesting species than detections of common species. A close inspection of Tables 7.1 and 7.2 confirms that common species were indeed under-represented in the presence-only datasets collated for north east NSW.

It is hardly surprising that models derived from presence/absence data performed better than those derived from pesence-only data. Better biological survey data will obviously produce better models. It should be noted however that presence/absence modelling offers far more benefits than simply an improved level of discrimination (the only criterion employed in this evaluation). For example, presence/absence models predict species occurrence in terms of probability rather than an index of relative likelihood. Presence/absence modelling also provides greater scope for estimating the uncertainty associated with predictions (e.g. by calculation of confidence limits). The downside is that the benefits of presence/absence modelling can only be realised if appropriate systematic survey data are available. Such data are costly to collect. Austin (1994) provides a useful discussion of the availability and modelling potential of presence-only and presence/absence datasets in Australia.

Caution should be exercised in generalising the results presented in this chapter to other regions and modelling techniques, for the following reasons: