2 Methodology **(Eds A Bryant, D Steller and D Connolly)**

This study adopts methods that are being increasingly used to standardise data gathered for biological diversity and its component groups across New South Wales. DEC brought its expertise and experience in the assessment of biological diversity from similar regional conservation assessments elsewhere in New South Wales.

Unlike many of those studies however, which generally were aimed primarily at the regulation of a single land use such as forestry activities, land use in the Georges River catchment includes both urban and rural zones under a range of tenures. These include industrial, residential, rural residential, agricultural and special uses (e.g. drinking water, army reserve etc.) areas. The range of land tenures in the Georges River catchment had a direct effect on the specific methods chosen and the broad approach adopted.

Given the limited funding available and the size of the region it was decided that the study would focus on representative elements of the biological diversity of the catchment, rather than being a more comprehensive assessment of total biological diversity. The significant increase in our understanding of biological diversity in the catchment resulting from such an approach was considered to outweigh the inherent flaws in adopting a representative approach.

Two factors were fundamental premises in developing methods for collation and management of appropriate data on biological diversity in the Georges River catchment:

- it would be used as a basis for future land use planning decisions; and
- the system chosen would need to be both accessible and flexible in its application.

The graphic nature of the majority of the data collected during this study required the adoption of a information system capable of handling large quantities of data in a graphical way. ArcView was chosen as the preferred Geographic Information System (GIS), as it is the primary GIS facility of the NPWS and also found in relatively common usage amongst local governments and other agencies. Examples of the successful application of GIS to the task of regional biological diversity assessment include the Comprehensive Regional Assessments in the north and south of New South Wales and numerous international examples, notably those in Canada, the UK and New Zealand.

Once the data management mechanisms had been established, the information that would be required to provide a regional perspective on the status of biological diversity in the catchment was identified. The following logic sequence resulted from an acknowledgment of the large study area and the fragmented nature of its natural areas:

- What do we already know about vegetation, flora and fauna in the area?
- In what ways is this information insufficient for the purposes of land use planning, and therefore what further vegetation, flora and fauna information is required?
- How are natural habitats (e.g. vegetation communities) distributed across the area and what is their condition?
- What key threatened species of flora and fauna are known to occur in the area?
- What is the status of potential habitat for the key threatened species, and consequently what is their potential distribution?

- As a result of this increase in knowledge about biological diversity in the catchment, what are the areas of high conservation value and what are the conservation priorities in the catchment?

This sequence was broken into the following steps:

- review existing information;
- translate vegetation information from aerial photograph interpretation into GIS layers as a basis for habitat modelling;
- survey sites throughout the catchment for vegetation generally, with targeted surveys for selected flora species;
- survey sites throughout the catchment for fauna generally, with targeted surveys for selected fauna species; and
- analyse the data to identify areas of relatively high conservation value.

These steps are detailed in sections 2.1 through 2.5.

NOTE: This study did not attempt to assess total biological diversity, relying instead upon key elements and broad-scale vegetation communities to represent the biological diversity of the region. In particular, the study did not attempt to survey invertebrates, fish or marine vegetation. The methodologies involved in assessing invertebrate communities were considered to be too labour intensive for a project of this scale. Fish and marine vegetation were also considered to be outside the scope of this study, however they were identified for a complementary second study involving NSW Fisheries.

2.1 Literature and data review

The data review examined data sources to identify records of systematic survey data or data on key threatened species. These selected species were determined after consultation within DEC Central Directorate (Comprehensive Regional Assessment Unit, Threatened Species Unit, Natural Heritage Unit and Environmental Planning Unit), and were finally endorsed by DIPNR.

The review aimed to gather data on sightings of target species—number of individuals, date, time, location and conditions of sighting, name of observer and the methodology of survey or observation.

The fourteen councils with administrative responsibilities in the Georges River catchment were contacted and asked to contribute relevant reports. DIPNR contacted environmental officers and requested their assistance in providing access for DEC to flora and fauna information held by councils. Most councils responded with a list of relevant documents and access was arranged for DEC staff to view the documents. The reports, which are listed in Appendix A, were individually assessed and relevant information extracted.

During the second phase the DEC contacted recognised ecological experts and both NSW and commonwealth government agencies that had undertaken work in the area. A number of high-quality data sets were obtained through this component of the project, in particular information provided by the Department of Defence (Holsworthy Military Area), Mr Arthur White (Biosphere Pty Ltd) and Mr Gary Daly (GAIA Pty Ltd).

Unfortunately many other data sources were deemed inappropriate for use as systematic data due to a lack of information on the method and effort applied, as well as having a focus on less common species to fill EIS requirements. This lack of consistency in methodology meant that there was little useful survey-based information obtained during the data review.

Consultants were used to gather field survey data. These individuals provided significant expertise and valuable information to the data pool.

Data on flora species were extracted from the Atlas of New South Wales Wildlife, along with other species-orientated documents, and reviewed for inconsistencies and unlikely records. These were then combined with recent records arising from the systematic plot-based flora surveys for the Western Sydney Vegetation Mappping Project and Campbelltown Mapping Project.

2.2 Vegetation—methods for assessing ecological communities

Throughout this report vegetation communities have been adopted as an imperfect surrogate for ecological communities, unless otherwise stated. This approximation of ecological communities is extremely crude, but has been adopted in the absence of more detailed ecological data on the areas in question. The lead for this approach has been taken from the NSW Scientific Committee, which has, to date, defined Endangered Ecological Communities (see the *Threatened Species Conservation Act 1995*) in terms of their floristic content. Data resulting from this study have been assessed to approximate faunal assemblages and ecosystem assemblages in order to provide a more sophisticated examination of the ecological communities in the region.

Endangered Ecological Communities on the Cumberland Plain have been defined by their floristic composition, and the survey work described in this report is based on the description of floristic assemblages (groups of co-occurring plant species). The description of floristic assemblages in a landscape involves the recognition of consistent and distinct patterns in species composition. This is a complex and time-consuming task involving the assessment of the distribution patterns of hundreds of individual species, many of which are small and difficult to identify. Methods used in the description of floristic assemblages vary in the degree of rigour with which this task is approached. Assemblages can be described by subjectively grouping sites according to structure and floristics. The degree to which the resulting classification reflects patterns in full floristic composition will depend on the skill of the observer, but the pathway to the result will be difficult to describe explicitly. The result may reflect a bias of the observer toward easily recognisable species (such as trees) and could be difficult to replicate by an independent observer.

The classification of vegetation communities in the Cumberland Plain used prior to this current study is largely the product of the subjective interpretation of floristic patterns (Benson 1992). For the purposes for which these data have been used, this classification system is adequate and has provided the basis for the listing of several vegetation communities in the TSC Act. Although systematic plot-based sampling has been conducted in some areas of the Cumberland Plain (e.g. Cohn & Hastings 1993) this has been insufficient for quantitative descriptions of communities throughout the whole of the Cumberland Plain.

With continuing pressure for development on the highly populated Cumberland Plain, the composition of vegetation communities has become the focus of intense scrutiny. The absence of quantitative data describing the relative abundance of species within communities has made it difficult to settle competing claims with respect to the classification of particular stands of vegetation. In addition, more recent studies have proposed the existence of new communities (NPWS 1997) for consideration for listing as Endangered Ecological Communities under the TSC Act (1995). The lack of a formal quantitative basis for Benson's (1992) original classification has meant that the assessment of new communities is problematic. The inherent variability within existing communities has not been quantified, and the position of any proposed communities within the floristic hierarchy can not be assessed. Since the communities are defined subjectively then disagreement amongst botanical experts is likely.

The current study aimed to avoid these problems associated with subjective classification by using a systematic method. Systematic techniques are designed to minimise observer bias and ensure that the pathway to the final classification is both explicit and repeatable. A formal sampling strategy was employed here to ensure data were collected in a systematic way and mathematical clustering techniques were used to display natural groupings based on the full floristic data. Quantitative data used to define clusters can be analysed to derive species lists that assist to differentiate between communities. Classifications derived using systematic methods are more readily compared than those derived subjectively. The use of systematic techniques has the advantage that analyses can be repeated when additional survey data become available. The floristic composition of new sites can thus be directly compared to sites in the original classification.

2.2.1 Sample stratification

i) Cumberland Plain

Survey sites on the Cumberland Plain were stratified using factorial combinations of geological substrate, temperature and rainfall across the study area (see Table 2). Some combinations were poorly sampled due to the small area of remnant vegetation present in those strata.

ii) Woronora Plateau

Survey sites on the Woronora Plateau were stratified using factorial combinations of three variables—sampling variation in substrate, slope/aspect and rainfall (see Table 3). Note that slope and aspect were included in place of temperature in recognition of the potential influence of the more complex topography of the Woronora Plateau.

The locations of survey sites from previous studies were evaluated to identify inadequately represented strata, which were then targeted for further survey.

Table 2 Cumberland Plain sample stratification

Feature	Classes
Geological substrate	• Wianamatta Shale • Holocene Alluvium (draining shale soils) • Tertiary Alluvium • Estuarine Sediments • Aeolian Deposits (Bannerman & Hazelton 1990)
Annual rainfall	• 701—800 mm • 801—900 mm • 901—1000 mm • 1001—1100 mm
Maximum temperature January (˚C)	• 26.1–27.0 ˚C • 27.1–28.0 ˚C • 28.1–29.0 ˚C

Table 3 Woronora Plateau sample stratification

Feature	Classes
Geological substrate	• Wianamatta Shale • Mittagong Formation • Holocene Alluvium (draining sandstone soils) • Hawkesbury Sandstone
Slope and aspect	• Slope ≤ 5° • Slope > 5° AND Aspect ≤ 45° OR Aspect > 315° • Slope > 5° AND 45° < Aspect ≤ 135° • Slope > 5° AND 135° < Aspect ≤ 225° • Slope > 5° AND 225° < Aspect ≤ 315°
Annual rainfall	• < 954 mm • 954—1183 mm • >1183 mm

iii) Campbelltown LGA

Sampling was concentrated in a narrow zone along the geological boundary between Wianamatta Shale and Hawkesbury Sandstone. Samples were stratified into five classes based on the composition of the overstorey as interpreted from colour aerial photographs (1998, 1:16 000) as part of the Western Sydney Vegetation Mapping project (Table 4). These classes covered the transition in vegetation from pure shale to sandstone floristic assemblages. Samples were arranged in a series of replicate transects laid out to traverse gullies.

Table 4 Overstorey composition of the API polygons used for the stratification of additional vegetation surveys in the Campbelltown LGA

Class	API Code	Species
1	**2**	*Eucalyptus moluccana*, *E. tereticornis*
	3	*E. fibrosa*, *E. crebra*, *E. moluccana*, *E. tereticornis*
2	**13**	*E. punctata*, *E. resinifera*, *A. bakeri*, *E. eugenioides*
3	**22a**	*E. pilularis*, *E. punctata*, *A. costata*, *S. glomulifera*
4	**22**	*E. pilularis*, *E. gummifera*, *E. punctata*
5	**23**	*E. sclerophylla*, *E. haemastoma*
	27	*E. gummifera*, *E. oblonga*, *E. sieberi*, *E. piperita*

2.2.2 Mapping of extant vegetation—aerial photograph interpretation

The area of extant native vegetation was estimated for the Cumberland Plain using aerial photograph interpretation (API). Aerial photographs taken between November 1997 and March 1998 were interpreted at a scale of 1:16 000 using a stereoscope. Remnants were classified into six classes according to remnant size and the density of eucalyptus tree cover (Table 5). The floristic composition of the overstorey was estimated for Classes A, B and C. Class C polygons included remnants with a non-eucalyptus tree stratum and remnants with no tree stratum (e.g. shrublands). Descriptions of the understorey were mainly qualitative (e.g. presence/absence of shrubs, weeds or vines), but dominance by particular genera was noted where possible (e.g. casuarina, melaleuca, olea). Class B polygons of area less than 5 ha were mapped as class Tx (scattered trees). Areas of scattered trees where agricultural activities were evident (e.g. heavily grazed areas, mustering yards, cropped land) were mapped as Txr. Areas of scattered trees with building structures present were mapped as Txu.

Table 5 Condition classifications for vegetation polygons using aerial photograph interpretation (API)

Class	Description	Eucalyptus tree species present	Minimum area mapped (ha)	Overstorey species recorded	Understorey characteristics recorded
A	Crown cover >10% (CCPD*)	Yes	0.5	Yes	Yes
B	Crown cover <10% (CCPD*)	Yes	5.0	Yes	Yes
C	Non-eucalyptus spp.	No	0.5	Yes	Yes
Tx	Scattered tree cover	Yes	0.5	No	No
Txr	Scattered tree cover, rural	Yes	0.5	No	No
Txu	Scattered tree cover, urban	Yes	0.5	No	No
S	Condition not assessed				

* CCPD—crown cover projected density (%)

i) Cumberland Plain

A map of extant native vegetation for the Cumberland Plain was produced by combining the modelled distribution of communities with a map of vegetation remnants interpreted from aerial photographs taken between November 1997 and March 1998. Structural and floristic patterns were interpreted using stereoscopic pairs at a scale of 1:16 000. By comparison, the resolution of the modelled distribution was unknown, but was at worst 1:100 000, being the scale at which the soil landscape coverages used in the model were mapped. In order to preserve the resolution obtained by using large-scale aerial photography, overstorey and understorey features recorded using API were used to assign Map Units to some polygons. Diagnostic species identified from the survey data were used for this purpose. The decision framework for allocation of mapped polygons to Map Units of specific vegetation communities is set out in Table 6.

The reliability with which overstorey and understorey species were described using API was variable. Expert knowledge obtained from field notes, site data and additional field traverses was used to modify the classification of polygons following the above procedure.

Table 6 Allocation of mapped areas (polygons) to vegetation community classes (Map Units)

Polygon description	Map Units
Polygons centred within the modelled distribution of Map Units 1, 2, 9, 10, 11, 12, 15 and 152 were assigned to Map Units according to the model. Exceptions to this rule were: • If Eucalyptus punctata was identified as an overstorey species (API) and the modelled class was Shale/Sandstone Transition Forest—Low Sandstone Influence (Map Unit 1) then the polygon was assigned to Shale/Sandstone Transition Forest—High Sandstone Influence (MU2) • If Eucalyptus fibrosa was identified as an overstorey species (API), and the modelled class was Shale Plains Woodland (MU10), and survey sites representative of Shale/Gravel Transition Forest (MU103) were present nearby, then the polygon was assigned to MU103	1, 2, 9, 10, 11, 12*, 15, 152
Polygons which intersected the modelled distribution of Alluvial Woodland (MU11) and contained casuarina spp. or melaleuca spp. (API)	11*
Polygons which intersected the modelled distribution of Western Sydney Dry Rainforest (MU13) and contained an understorey with mesic or vine species (API)	13
The extant coverage of Moist Shale Woodland (MU14) was mapped by cutting the modelled distribution directly to the API coverage	14
Polygons centred within the modelled distribution of Castlereagh Ironbark Forest (MU3) were assigned to MU3 if *Eucalyptus fibrosa* and Melaleuca spp. were identified, or if shrubland without overstorey was identified (API) • Polygons not meeting these criteria were assessed on an individual basis and assigned to Map Units based on overstorey composition, field notes and other published sources.	3*
Polygons containing sites representative of Castlereagh Swamp Woodland (MU4), and adjacent polygons containing melaleuca understorey (API)	4
Polygons centred within the modelled distribution of Castlereagh Scribbly Gum Woodland (MU6), and adjacent polygons containing *Eucalyptus parramattensis* or *E. sclerophylla* API)	6
Polygons intersecting the modelled distribution of Agnes Banks Woodland (MU8), and adjacent polygons containing *Banksia serrata* (API)	8
Polygons in which wetlands were identified were assigned to Freshwater Wetlands (MU36), although freshwater wetlands were not sampled in this survey	36*
Polygons in which saltmarsh or mangroves were identified were assigned to Mangrove/Saltmarsh Complex (MU34)	34*

* Polygons of Class C (Table 5) that did not meet specific criteria relating to MU3, MU11, MU12, MU34 and MU36 were not assigned a Map Unit because it was considered that there was insufficient information to differentiate these areas from Class X polygons (no native tree cover).

ii) Woronora Plateau

The extent of native vegetation on the Woronora Plateau was estimated using Landsat imagery. Forest structure was delineated through manual interpretation based on the degree of observable reflectance. The minimum polygon size was 10ha, although smaller remnants were mapped these were easily discernible. The results were updated using a 1994 Spot Image to determine the extent of clearing since 1990. Vegetation classes were assigned to this layer from the modelled distribution of ecological communities. The coverage layer for vegetation was finally derived by combining the API and Landsat layers.

2.2.3 Spatial data layers

Spatial data layers used in modelling the distribution of vegetation communities are listed in Appendix B. Parent geology was derived from soil landscape units (Chapman & Murphy 1989, Hazelton *et al*. 1990, Bannerman & Hazelton 1990). The soil landscape coverage was used to derive variables representing gradients in soil characteristics. These included the calculation of the shortest distance to sandstone-derived soils (which was calculated twice, with soils derived from the Mittagong Formation alternately included and excluded), the shortest distance to soils derived from shale and the shortest distance to Tertiary Alluvium. Distance to stream was calculated using all streams included in a data coverage supplied by the NSW Land Information Centre (LIC). This coverage ranked streams by size from one (smallest) to six (largest, e.g. Nepean River), allowing distance to larger streams to be calculated separately. Terrain variables were derived from a Digital Elevation Model supplied as a 25 m grid by LIC. Two indices of localised variation in elevation were calculated using cells contained within a square cell neighbourhood. Ruggedness was calculated as the standard deviation in elevation in neighbourhood cells. Terrain was calculated by subtracting the average elevation across neighbourhood cells from the central cell value. Neighbourhoods of 100, 300, 500, 700 and 900 m were employed. Climatic variables (rainfall, temperature, solar radiation) were derived from the digital elevation model using ESOCLIM (Hutchinson 1989).

2.2.4 Field sampling

Field sampling was carried out between October 1998 and June 1999. At each survey site the vegetation within a 0.04 ha quadrat was described. Quadrats were marked out using tape measures in an area representative of the surrounding vegetation and as far as possible away from areas of weed infestation or soil disturbance (highly disturbed or weed infested areas were not sampled). Where possible, several quadrats were used to sample local variation in slope, aspect and landform (e.g. gullies and ridges). Quadrats were square in shape except where a rectangular configuration was required to ensure homogeneity of terrain and soils across the plot. The sites surveyed are illustrated in Figure 2.

All angiosperm species rooted within the quadrat were recorded and assigned a cover/abundance score using a modified Braun–Blanquet scale (Poore 1947) (Table 7).

The height range and projected foliage cover were estimated for each of four structural strata (tree, small tree, shrub or forb) where recognisable at the site. A clinometer was used to measure the slope at the centre of the quadrat, whilst a compass was used to determine aspect at the centre. The two instruments together were used to determine horizon azimuths at compass bearings of 0, 45, 90, 135, 180, 225, 270 and 315° The location and elevation of the site were determined in the field using 1:25 000 topographic maps and/or a geographic positioning system (GPS). The soil type was determined by hand-texturing. Evidence of rock outcropping, erosion, weed invasion, logging, soil disturbance or recent fire was noted.

Table 7 Modified Braun–Blanquet (Poore 1947) scale used to evaluate the cover/abundance of angiosperms in each quadrat

Rank	Abundance/distribution	Cover, %
1	Rare, few individuals present	< 5
2	Uncommon	< 5
3	Common	< 5
4	Very Abundant	0–19
5	Very Abundant	20–49
6	Very Abundant	50–74
7	Very Abundant	75–100

Plant species that could not be identified in the field were collected for later identification. Where necessary, collections were compared to specimens held at the NSW Herbarium to confirm their identity. Specimens that could not be identified to species level were not included in the analysis. Nomenclature was standardised to follow Harden (1990). In many cases species recognised at the subspecies level were identified only to the species level. Published species lists (Benson *et al.* 1996; James *et al.* 1999) were consulted to determine the number of subspecies recorded for the study area. If only one subspecies had been recorded then the subspecific epithet was adopted. If two or more subspecies had been recorded then subspecies were pooled for analysis. The species name used in the community descriptions indicates the taxonomic level used in the analyses. Exotic species were recorded but excluded from analysis.

THIS PAGE HAS BEEN INTENTIONALLY LEFT BLANK FOR THE INSERTION OF FIGURE 2

Figure 2 Vegetation survey field sampling sites

2.2.5 Vegetation classification

Data analysis was performed on the raw cover/abundance scores using the PATN package (Belbin 1991). Initially, only data collected in the present survey were included in the analysis. Dissimilarity amongst survey sites was computed using the Kulczynski coefficient with a symmetric form of measure. Hierarchical agglomerative clustering was performed using a flexible unweighted pair group arithmetic averaging strategy with no adjacency constraint and a BETA value of -0.1. Homogeneity analysis (Bedward *et al.* 1992) was used to determine the point in the hierarchy at which a decline is observed in the rate of increase in within-group homogeneity yielded by further group subdivision. Visual inspection of the hierarchical dendrogram confirmed the integrity (Belbin 1991) of the groups defined at this point, thus amalgamation was not considered. Groups containing clusters of high integrity were identified for possible subdivision. In addition, the preliminary groupings were compared to previously derived floristic classification for the survey area to identify subgroups corresponding to recognised communities (Benson 1992; Keith 1994; French *et al.* in press). Subdivision was carried out by systematically increasing the number of groups across the whole dendrogram. This process was continued only while an increase in the number of groups led to the separation of subgroups of high integrity, and ceased when groups of high integrity became split. When the definition of groups was completed the analysis was repeated using the Bray–Curtis coefficient of dissimilarity (Clarke 1993) to examine the consistency of the grouping. Site classifications were crosschecked by examining the classification of the five nearest neighbours of each site. Sites with a different classification to all five nearest neighbours were re-evaluated.

After the sites sampled in the current survey (Cumberland Plain and Woronora Plateau) had been classified, clustering was repeated using all available survey data. The aim in this process was to allocate sites to previously defined clusters in preference to defining new assemblages. The assumption here was that in developing a regional scale classification of floristic assemblages from several smaller scale classifications, the most likely result was the amalgamation of units rather than further subdivision. Given the potential for variation in sampling technique amongst observers and conditions under which sampling was completed, it was considered that a conservative approach to analysis would be required. Nevertheless, the potential for the appearance of new clusters following the pooling of samples from previously poorly sampled strata was recognised. Clusters were not amalgamated unless there was a substantial overlap in floristic composition.

Two published vegetation classifications derived from clustering analyses were consulted to produce a list of communities described for the Woronora Plateau (Keith 1994; French *et al.* in press). These classes were merged with the classifications derived in the prior analysis, with areas of possible overlap noted. Sampled sites were initially clustered using the raw cover/abundance scores, then the data were converted to presence/absence to compensate for differences amongst observers in the estimation of cover/abundance and/or the cover/abundance scale used. Interpretation of the resulting dendrograms sought to resolve predefined clusters and amalgamate or split elements of existing classifications where necessary. Samples that would not integrate with other samples in the dendrogram were subjected to particular scrutiny, including examining the raw data for obvious anomalies. The full data set was reduced to various subgroups containing sites sampling relatively narrow biogeographical regions to assist in the interpretation of disparate strands of the dendrogram.

Description of vegetation communities

The floristic assemblages derived by cluster analysis were compared to community descriptions:

- contained in recent publications on the vegetation of the survey area (e.g. Keith 1994; French *et al.* in press; Benson & Howell 1990; Benson 1992; NPWS 1997a), and
- of Endangered Ecological Communities listed under the TSC ACT (Schedule 1).

Where an assemblage was judged to represent a previously described community, the assemblage was given the name in common use for that community. Where it was considered that an assemblage had not been previously described then a name was constructed using elements of the vegetation structure, topographical and geological preferences of the assemblage (e.g. Shale Hills Woodland).

Structural descriptions were compiled from the survey site data. The maximum height and projected foliage cover for each stratum were averaged across all sample sites representative of the community. The frequency with which each stratum was encountered in the community was also calculated. Summary statistics (mean, standard deviation and range) for the sample sites were calculated for elevation, slope, annual rainfall, ruggedness (900 m neighbourhood), solar radiation (January) and maximum temperature (January). The frequency with which sample sites were located on different substrates was also calculated.

Diagnostic species were defined using the concept of fidelity described by Keith and Bedward (1999). Species were allocated to one of four classes based on their frequency of occurrence in the target community compared to all other communities, and the median cover/abundance recorded in the target community compared to all other communities (Table 8). Positive diagnostic species are more likely to be found in the target community than in other communities. Negative diagnostic species are less common in the target community but occur frequently in other communities. Constant species occur frequently in several communities including the target community and uninformative species occur infrequently in all communities. Species classed as positive diagnostic, negative diagnostic or constant were tabulated for each community. Species classed as uninformative were not included unless they were tree species. Uninformative tree species were included to give an indication of the variation in canopy composition within communities.

Table 8 Criteria used in the assignment of fidelity classes to diagnostic species in each assemblage

		Non-target communities		
		Freq. >0.5 *and* median C/A >2	**Freq. < 0.5 *or* median C/A >2**	**Freq. = 0**
Target assemblage	**Freq.[1]>0.5 *and* median C/A[2] >2**	Constant	Positive diagnostic	Positive diagnostic
	Freq. < 0.5 *or* median C/A ≤ 2	Negative diagnostic	Uninformative	Positive diagnostic
	Freq. = 0	Negative diagnostic	Negative diagnostic	–

1. Frequency (freq.) was calculated as the proportion of sample sites in which the species was recorded.

2. Median C/A is the median cover abundance score across samples.

Correlation with environmental variables (spatial data)

Correlations between floristic composition and environmental variables were explored using hybrid multidimensional scaling and principal axis correlation (Belbin 1991). The environmental variables, or spatial data, which were used in vegetation analysis are tabled in Appendix B. Ordination was performed on a dissimilarity matrix calculated using the symmetric form of the Kulczynski coefficient. Solutions were calculated in five and six dimensions from ten random starting configurations and a maximum of 50 iterations. The procedure was terminated if successive iterations produced an improvement in stress of less than 0.005. The choice of solution dimension was designed to minimise stress in the resulting solution and maximise the chances of revealing complex, fine-scale correlations in the data, while maintaining computing time at an acceptable level. Ratio regression was applied below ratio/ordinal cut values of 0.9, 0.2 and 0. By reducing the ratio/ordinal cut it was hoped that the solution stress could be further reduced (a non-metric solution would not be expected to deteriorate rank correlation in the fitted vectors).

Ordination and principal axis correlation were performed separately for sites located on the Cumberland Plain and Woronora Plateau. Separate analyses were used to allow for the possibility of different variables correlating highly with floristic patterns on the flat landscapes of the Cumberland Plain compared to the highly dissected Woronora Plateau. Some of the variables listed in Appendix B were not included in principal axis correlation because they were introduced only in the later stages of model development. The importance of the two ordinal variables (geology, and its derivative soil landscapes) could be assumed *a priori* (Benson 1992), thus no attempt was made to correlate them. A list of the vectors fitted to the ordination configuration with the lowest stress is contained in the results Section 3.2. The environmental data were derived from the digital data layers. Field data (e.g. slope and aspect) were not used because they frequently differed from the values derived from the digital data layers, particularly where survey sites sampled fine-scale topographical features. Topographical data layers were derived from a digital elevation model and represent a simplification (or smoothing) of the landscape. Since only the derived values could be used for modelling, correlations made between floristic patterns and derived values were most useful in modelling the distribution of vegetation communities. Variables were ranked in order of correlation and used preferentially in subsequent modelling of community distributions.

Spatial modelling of vegetation communities

Spatial interpolation of vegetation communities was carried out using a hybrid technique of decision-tree and expert systems developed by Keith and Bedward (1999). This technique utilises purpose-built software (ALBERO, Keith & Bedward 1999) to develop decision rules that quantify the environmental domain(s) occupied by each community. At each node in the decision tree ALBERO provides a list of variables which could be used to discriminate communities. The variables are selected on the basis of chi-squared statistics. The level of significance for the test ($p < 0.05$ in this case) and the variable upon which the split is performed are chosen by the user. This technique was used successfully in modelling 79 floristic assemblages in the South-East Forests Region of New South Wales (Keith & Bedward 1999).

Compared to statistical models, decision tree models are at a relative disadvantage in that fewer and fewer samples are utilised with each successive split in the tree (Keith & Bedward 1999). This means that areas with inadequate sampling are modelled less accurately, but in addition, the split point chosen for a variable may be influenced by the position in the tree at which the split is made. Several strategies were used to minimise the effects of this limitation. Variables that were highly correlated with floristic patterns through principal axis correlation were used preferentially in building the decision tree. Other variables were used sparingly, and only where there was an ecologically intuitive reason for doing so (i.e. splits were never determined on statistical significance alone). Variables that described small-scale gradients were fitted before variables describing gradients across the whole study area. This allowed the exploration of complex patterns using the maximum number of samples. Once this had been achieved the decision tree could be rebuilt commencing with large-scale variables, a conceptually easier task, and where appropriate, the splitting of small-scale variables further along the tree could then be forced to comply with decisions derived using larger sample sizes. Often when a split was made using a large-scale variable, further splits on the same variable were pursued in consecutive nodes. While due consideration was given to possible interactions between variables, in most cases it was deemed appropriate to standardise the value of a split where a particular variable was used to discriminate a pair of communities in different branches of the tree. In general, variables yielding splits that isolated small numbers of communities were pursued in preference to variables that split communities evenly. Terminal nodes were assigned to the community represented by the most samples.

A purpose-built mapping program (ALBERO Mapper) was used to compile a 25 m-grid coverage from the decision rules and GIS data layers. The development of the coverage was iterative. For each rule set the modelled distribution of each community was assessed through comparisons with the distribution of the sample sites and descriptions of distributions contained in field notes and published reports. Where discrepancies were apparent the decision rules were deconstructed and reapplied as subrules to identify the section of the tree responsible. The decision rules were changed if inconsistencies in the construction of the decision tree were identified (e.g. in the way a community was modelled in different branches of the tree). Due to the highly fragmented nature of much of the survey area, parts of the environmental domain could not be sampled. Decision rules were modified to extend the range of some communities based on the interpretation of historical records (Benson 1992; Benson & Howell 1990). The final coverage was smoothed using a majority filter operating over a radius of 50 m.

2.3 Methods for assessing species diversity—selected fauna species

The purpose of the fauna survey component of the study was to assess the diversity of vertebrate fauna within the catchment. The methods chosen to complete this assessment included two broad approaches:

- Sampling of common/abundant vertebrate fauna through generic sampling methodology, and
- Supplementing generic vertebrate fauna surveys with surveys for rarer fauna species. These species would be chosen as surrogates of the complete faunal diversity in the catchment.

The NPWS Wildlife Atlas and information collated through the data review were assessed for records of regionally significant species anywhere within a 20 km radius of the study area to produce a list of selected species. Each of the twenty-four species used (see Table 9) were chosen because they were:

- known or likely to occur in the Georges River catchment
- and either
- threatened by land use changes within the catchment
- or
- rare, and therefore a good surrogate of species diversity.

The primary objective of the fauna study was to develop habitat distribution maps (models) for each of the selected fauna species and to link these to known fauna sightings. To meet this goal the following tasks were completed:

- a review of all known fauna reports, environmental assessments and affiliated survey work in the catchment (discussed in Section B1)
- a review and identification of selected fauna species known to occur in the catchment
- a systematic field survey for all vertebrate fauna groups
- additional targeted surveys for selected fauna species
- the development of computer-based models of species habitat distribution within the catchment for selected fauna species.

The key outputs from the above tasks include:

- the provision of spatial information incorporating a point locality database of data gathered during the field surveys (see Section 2.3.2) and data review (see Section 2.1); and
- predictive habitat models of fourteen of the selected species (see Section 2.3.4).

This section of the report provides a background to the methods used to assess and collate data on fauna, and to assist in the interpretation of this information. Together with vegetation community data (Section 2.2) and flora data (Section 2.4), these fauna data will be used to identify areas of significance for the management of biodiversity in the catchment. As such, this section of the report refers to the GIS (database) information that compiles this extensive data and to C-Plan, the software used to assess this information.

Table 9 Selected fauna species for the Georges River catchment

Group	No.[1]	Common name	Scientific name	Status[2] (TSC Act)
Amphibians	i	Giant Burrowing Frog	*Heleioporus australiacus*	V
	ii	Red-crowned Toadlet	*Pseudophryne australis*	V
	iii	Brown Toadlet	*Pseudophryne bibronii*	P
	iv	Green and Golden Bell Frog	*Litoria aurea*	E
	v	Green Tree Frog	*Litoria caerulea*	P
Reptiles	vi	Heath Monitor	*Varanus rosenbergi*	V
	vii	Broad-headed Snake	*Hoplocephalus bungaroides*	E
Birds	viii	Bush-stone Curlew	*Burhinus grallarius*	E
	ix	Glossy Black-Cockatoo	*Calyptorhynchus lathami*	V
	x	Swift Parrot	*Lathamus discolor*	E
	xi	Turquoise Parrot	*Neophema pulchella*	V
	xii	Barn Owl	*Tyto alba*	P
	xiii	Powerful Owl	*Ninox strenua*	V
	xiv	Rock Warbler	*Origma solitaria*	P
	xv	Brown Treecreeper	*Climacteris picumnus*	P
	xvi	Regent Honeyeater	*Xanthomyza phrygia*	E
Mammals	xvii	Koala	*Phascolarctos cinereus*	V
	xviii	Squirrel Glider	*Petaurus norfolcensis*	V
	xix	Eastern Horseshoe-bat	*Rhinolophus megaphyllus*	P
	xx	Eastern Freetail-bat	*Mormopterus norfolkensis*	V
	xxi	Undescribed Freetail-bat	*Mormopterus species 1*	P
	xxii	Common Bentwing-bat	*Miniopterus schreibersii*	V
	xxiii	Eastern False Pipistrelle	*Falsistrellus tasmaniensis*	V
	xxiv	Large-footed Myotis	*Myotis adversus*	V

1. These numbers will be used in reference to the associated species throughout this report.
2. Status at 25 August 2000. E—Endangered (Part 1, Schedule 1, TSC Act); V—Vulnerable (Schedule 2, TSC Act); P—Protected (NP&W Act 1974).

Notes on selected species: Since the field studies were completed a number of additional species from the region have been listed under the TSC Act, including the Eastern Pygmy Possum (*Cercartetus nanus*), Grey-headed Flying fox (*Pteropus poliocephalus*), and Heath Frog (*Litoria littlejohni*). These species did not recieve targeted survey effort and no attempt was made to model potential habitat for them.

Migratory waders and seabirds were excluded from the selected species list although several threatened species are known from the study area, including the migratory Broad-billed Sandpiper *(Limicola falcinellus)*, Great Knot *(Calidris tenuirostris)* and

Terek Sandpiper *(Xenus cinereus).* A significant breeding population of Little Terns *(Sterna albifrons)* inhabits areas of Towra Point. These species are subject to considerable ongoing research and management action and are not addressed in this report. The threatened Pied Oystercatcher *(Haematopus longirostris)* and Sooty Oystercatcher *(H. fuliginosus)* are also found in the catchment. Implications for management of these species will in part be addressed in third part of this report, *Aquatic biodiversity of the Georges River catchment*.

2.3.1 Site selection

The field survey program was planned to systematically sample the range of habitats present in the catchment for all vertebrate fauna. The survey design would:

- provide new records for selected species from which to undertake habitat modelling tasks; and
- form a sound basis for the formulation of future biodiversity management objectives and strategies.

Sites were selected based on a stratification of the study area derived through a combination of environmental attributes including geology, rainfall, vegetation and size of vegetation remnants. This ensured an adequate environmental spread of survey sites. Existing sites (obtained through the data review process) were considered in relation to this stratification and further sites were undertaken to fill the environmental gaps. At the site planning stage, vegetation communities had not been mapped and hence were not available to aid in situating the fauna sites.

2.3.2 Survey methodology

Consistent methodology in field survey enables the researcher to draw some statistical conclusions about the presence and absence of species in different habitats and environments. This data provides the information that underpins the extrapolation of the types of environmental conditions that are likely to support a species. The methods used were also chosen to minimise wasted sampling effort and maximise sampling results, given the time and cost limitations of the study.

Fauna surveys were undertaken by qualified contracted consultants in accordance with methods developed by the NSW Government for the Comprehensive Regional Assessment process. The specific methods used for each fauna group are detailed in Table 10 and described below.

Table 10 Survey sites (number) for each survey method

Survey technique		Number of sites
Herpetofauna	Diurnal census	**42**
	Nocturnal frog census	**25**
Birds	Diurnal census	**167**
Arboreal mammals	Transect spotlighting	**13**
Bats	Harp trapping	**46**
	Echolocation recording	**42**
All groups	Opportunistic off-site census	**104**

Herpetofauna census

- **Diurnal herpetofauna census—**involved a one person-hour search of an area of 0.5 ha (50x100 m). Surveys were conducted during daylight hours when temperatures ensured highest reptile activity. Note that this may not have always been at the hottest time of the day on exposed sites. Sites were sampled by searching fallen logs, litter, fallen bark, rock outcrops and other suitable microhabitat.

Contracted consultant: Ecotone Pty Ltd

- **Nocturnal frog census**—surveys were carried out along the edges of a 200 m section of stream. The equivalent of one person-hour was spent at each site. Any frogs either seen or heard during this period were recorded. Surveys were only undertaken within two days of rain and when ambient temperatures exceeded 20°C.

Contracted consultants: Ms Kylie Madden and Mr Steven House

Additional frog sites: Biosphere Pty Ltd

Bird census

- **Diurnal bird census**—Twenty person-minutes were spent recording any birds seen or heard within a 2 ha site, during periods of peak bird activity, i.e. at either dusk or dawn. Censuses were not undertaken on rainy or windy days or during times of extreme heat. In order to maintain consistency in the survey approach, birds identified close to the site or flying over the site were recorded as off-site. Those recorded before or after the survey period were recorded as opportunistic on-site.

Contracted consultants: Mr Richard Turner and Mr Graham Turner

Arboreal mammal census

- **Spotlighting**—was used as the primary method for location of arboreal mammals, aided by regular stops to listen for evidence of these species. Two people walked a predetermined 2 km transect over a period of one hour.

Contracted consultant: Ecotone Pty Ltd

Bat census

Two techniques were employed to detect bats:

- **Harp trapping**—traps were set in areas deemed to be suitable bat flyways. Each trap was left for two nights at a site. Bats were removed and identified at dawn each morning, then released if suitable hollows could be located nearby; otherwise they were released at dusk that day.
- **Echolocation recording—**detects high-flying bats that are typically not captured using the harp trap technique. In this technique an Anabat recorder is used to detect the echolocation signals of bats flying through the field-range of the detector. Each recorder was fitted with a delay switch and left at a site for one night with a maximum recording period of 30 minutes. The call records made in this way were later fed into software equipped with template calls for each of the species. The computer was able to identify species from signature sequences in the call.

Contracted consultant: Ecotone Pty Ltd

2.3.3 Data storage—fauna

All data gathered as part of the fauna survey program was entered into the NPWS Wildlife Atlas Biodiversity Survey Subsystem (BSS). This database allows for the entry of presence/absence in relation to each site. Records gathered opportunistically as a result of data review were entered directly into the Wildlife Atlas.

An extensive amount of preprocessing of data is required before the initiation of a modelling run. To facilitate this all selected species data is extracted from the Wildlife Atlas and BSS and imported into an MS Access database. Here the data is manipulated and files exported in the format required for fauna modelling.

2.3.4 Fauna habitat modelling

The principle objectives of the fauna surveys were to:

- gain information on existing fauna values within the study area, and
- spatially extend (or interpolate) known occurrences of selected species into habitat maps for the entire study area based on modelled relationships between fauna and remotely mapped attributes (climate, topography, geology etc.).

A computer-based modelling package was used to statistically examine relationships between the location of species and environmental attributes such as climate; terrain; vegetation type and geology. These relationships can then be used to identify identical features in the landscape using a GIS, thereby predicting potential habitat.

The models were produced for each species to enable definition and evaluation of explicit conservation requirements of each species. Not all species supported sufficient records within the catchment from which to develop statistical models. In such cases an Expert Habitat Model was developed by an expert familiar with the habitat requirements of the relevant species.

Modelling of fauna habitats relied on the development of spatial data layers that describe variations in the natural environments present within the catchment. These include data layers that describe variations in soils, climate, terrain, and vegetation communities. Appendix B lists the spatial data layers developed during this study, and cross-references those used in relation to fauna and to flora and vegetation modelling exercises respectively.

The aim of the modelling process was to spatially interpolate known occurrences of fauna species from field surveys throughout the study area by finding statistical relationships between the biota and the background environmental variables.

Preprocessing of data identified the underlying environmental values for each record of selected fauna species. Statistical analyses of the presence and absence of the species for each of the environmental variables was completed. A module running under S-PLUS statistical software produced by Watson (1996) and combined with ArcView was used.

The predictive species modelling package provides the user with a choice between the two most commonly used logistic regression procedures, generalised linear modelling (referred to as GLMs) and generalised additive modelling (referred to as GAMs) (Watson 1996).

- **Generalised linear modelling** (GLM) is essentially an extension of ordinary linear regression that approximates a linear (straight line) function to relate a response (dependent) variable to one or more predictor (independent) variables. Two of the basic assumptions of linear regression are that the relationship between response and predictor variables can be approximated by a straight line, and that the variance associated with the response is homogenous throughout the full range of the response variables. GLM allows a class of models that provide non-linearity and heterogenous variance in response functions, thus avoiding the assumptions made by linear regression (NPWS 1994a).

- **Generalised additive modelling** (GAM) is essentially an extension of GLM, the major difference being GAM uses a non-parametric, smooth function which relates the response to the predictor. The functions are smooth curves estimated from the data using techniques originally developed for smoothing scatter-plots (data distribution maps). The software uses cubic splines to approximate smooth functions to derive the GAM. As a result, the principal difference between the two modelling techniques is that GAM allows the survey data to determine the shape of the response curves, whereas GLM is constrained by parametric forms, that is, cubic and quadratic polynomial response curves (NPWS 1994a; Watson 1996).

Presence-only data was modelled in relation to 1 000 pseudo absences which were randomly chosen from all land within the study area. Pseudo-absence points were given a weighting of n/1 000 (where n = the number of presence records for each species). This manner of weighting facilitated approximation of degrees of freedom, deviance and significant levels appropriate to presence-only modelling. These weightings also enabled predictions to be expressed in terms of a relative index of likelihood of occurrence ranging from 0 to 1 (later converted to a percentage).

The statistical analyses provides a combination of variables which best describes the pattern of data for each species. This combination of variables or equation is then run through ArcView to map the intersection of these variables in the landscape. A map describing the likelihood of occurrence across the entire catchment is produced.

GLMs and GAMs were run for each species for which there were at least ten known locations within or adjacent to the catchment. Modelling runs were undertaken several times varying the available environmental layers each time. Ecological experts evaluated the models of each selected species, selecting the model which best described the species habitat according to their knowledge of species ecology within the catchment and the Sydney region. Where necessary models were further refined in order to avoid spurious statistical inferences.

In some instances insufficient data were available to develop robust statistical models. In such cases, attempts were made to build models of species habitat based entirely on expert knowledge. This entailed experts describing the habitat requirements of the species, for example a certain forest type or geomorphological unit, then relating this to the available layers within the GIS. At this stage Boolean overlay techniques were utilised to spatially define the expert knowledge. These maps were checked and refined again by the relevant expert, as appropriate. This was an iterative process and continued until the expert felt that an adequate model had been produced.

Section C3 describes the models selected for each species.

2.3.5 Fauna assemblage investigation

Faunal assemblages were identified for each of the major vegetation communities mapped during this study.

A separate fauna profile has been produced for each vegetation community (Appendix F). Profiles summarise the characteristic species found for each fauna group. This has been based on the results of multivariate statistical analysis of systematic data collected during this study. The profiles also present information on the species richness and diversity of the faunal assemblages.

Several vegetation communities have very restricted distributions within the study area, resulting in small samples and preventing the application of statistical analyses for the assemblages of some communities. In these instances we have listed those species that have been recorded. This also applies to terrestrial and arboreal mammals, nocturnal birds and endangered species for all profiles.

Statistical analysis

i) Species richness

Species richness is the number of species occurring at a site. The species richness of diurnal birds, reptiles, bats and frogs was averaged for each of the major vegetation communities. Vegetation communities were ranked as having species richness of High, Medium or Low for each of these fauna groups. These ranks were based on relative scores between each vegetation community within the catchment.

ii) Species diversity

Diversity was calculated using the Shannon–Wiener Diversity Index (H'), where P_i = proportion of the total count arising from the i^{th} species.

$$H' = -\Sigma P_i \log(P_i)$$

Species diversity gives an indication of the number of species at a site, and of their relative abundance. The more equal their abundance at a site, the more diverse that site is. Species diversity was averaged across sites for each vegetation community, and ranked as High, Medium or Low for each fauna group. Again, these ranks were determined based on relative scores of each vegetation community within the catchment

iii) Similarity matrices

Bray–Curtis similarity matrices were created for the diurnal birds, reptiles and frogs. This method creates pair-wise comparisons for every site using the Bray–Curtis coefficient. Sites that are perfectly similar will have a coefficient of 1, and perfectly dissimilar will have a coefficient of 0. For the bats we used Euclidean distance to determine the similarity/dissimilarity coefficient. This is because Euclidean distance allows the inclusion of sites that have no captures recorded. These matrices are shown for assemblages in the results section of this report.

Similarity matrices are useful for identifying groups of sites that contain similar faunal assemblages. Matrices were the basis of several other techniques performed, including the ANOSIM and cluster analysis.

iv) ANOSIM (analysis of similarities)

Analysis of similarities (ANOSIM) was used to test for differences between the fauna assemblages of the major vegetation communities of the GRC. It was applied to diurnal bird, reptile, bat and frog assemblage data. This statistical test is similar to analysis of variance (ANOVA), however it allows for the use of multivariate rather than univariate data. ANOSIM ascertains if there is a greater degree of similarity between replicates (sites) within a vegetation community than there is between sites from different vegetation communities. The output is an R statistic and a significance value for each vegetation community pairing. This information is presented in the assemblages section of Part 4 in the form of matrices.

The results of the ANOSIM are the basis of much of the information provided in the vegetation community fauna profiles (Appendix F). The profiles list the relationships between each vegetation group for diurnal birds, reptiles, bats and frogs. These relationships are summarised as either being similar or dissimilar based on the statistical significance score generated by the ANOSIM.

v) Cluster analysis

Cluster analysis is a graphical presentation of a similarity matrix. Sites are linked to one another based on their similarity coefficient to form a dendrogram. Using this we can investigate the faunal assemblages of sites that group together using a statistical procedure known as SIMPER.

vi) SIMPER

SIMPER is a technique that isolates fauna species that characterise site-groupings identified in the cluster analysis. It can only be used to describe groups defined *a posteriori*, that is, based on the output of the cluster analysis. Where sites grouped strongly by vegetation, this technique provides an excellent means of defining the faunal assemblages of the mapped vegetation communities. However, sometimes vegetation may be less important for defining fauna assemblages than other ecological or stochastic factors. Where fauna assemblages did not group strongly by vegetation community, a listing of the most commonly recorded species has been included in this report (see Appendix D) as a guide to what species might be encountered.

2.4 Methods for assessing species diversity—selected flora species

To supplement the general vegetation surveys conducted as part of the assessment of vegetation communities (see Section 2.2), selected species of flora were used as surrogates of floral diversity in the catchment. Each of the seven species finally used (see Table 11) was chosen because it was:

- known or likely to occur in the Georges River catchment
- and either
- threatened by land use changes within the catchment
- or
- rare, and therefore a good surrogate of species diversity.

The primary objective of the flora surveys was to develop habitat models to more accurately map the known the sites of these species and predict their potential distributions within the catchment. The development of such maps removes the bias often associated with assessments that rely solely on known individual species records.

To meet this goal the study encompassed the following tasks:

- a review of all known flora reports, environmental assessments and affiliated survey work in the catchment (discussed in Section 2.1)
- a review and identification of selected flora species known or likely to occur in the catchment
- systematic field survey for these selected flora species
- the development of computer-based models of species habitat distribution within the catchment for selected flora species.

The key outputs from the above tasks include the provision of spatial information incorporating a point locality database of data gathered during the field surveys and data review projects as well as predictive models of six of the seven selected species. Together with vegetation community data (Section 2.2) and fauna data (Section 2.3), these flora data will be used to identify areas of significance for the management of biodiversity in the catchment.

This section of the report provides a background to the methods used to assess and collate data on flora, and to assist in the interpretation of this information. As such, it refers to the GIS database information that compiles this extensive data.

The field survey program was planned to systematically sample the range of habitats present in the catchment for all flora species. The survey design would:

- provide new records for selected species from which to undertake habitat modelling tasks, and
- form a sound basis for the formulation of future biodiversity management objectives and strategies.

Consistent methodology in a field survey enables the researcher to draw some statistical conclusions about the presence and absence of species in different habitats and environments. This data provides the information that underpins the extrapolation of the types of environmental conditions that are likely to support a species.

Table 11 Selected flora species for the Georges River catchment

Number[1]	Scientific name	Family	Status[2] (TSC Act)
I	*Pimelea spicata* R. Br.	Thymelaeaceae	E
II	*Persoonia nutans* R. Br.	Proteaceae	E
III	*Pultenaea parviflora* Sieber ex DC.	Fabaceae	E
IV	*Pterostylis saxicola* D.L. Jones & M.A. Clem.	Orchidaceae	E
V	*Pterostylis gibbosa* R. Br.	Orchidaceae	E
VI	*Cynanchum elegans* (Benth.)	Asclepiadaceae	E
VII	*Gyrostemon thesioides* (Hook.f.) *A.S. George*	Gyrostemonaceae	E

1. These numbers will be used in reference to the associated species throughout this report.
2. Status at 25 August 2000. E—Endangered (Part 1, Schedule 1, TSC Act .

Maps were constructed by examining statistical relationships between known locations and climatic, topographic and substrate features.

Section 4 of this report includes profiles on each of the species studied, including characteristics and distribution of the species studied, explains the modelling procedure and provides maps of potential species habitats. An explanation of how maps should be interpreted is also provided.

2.4.1 Site selection

Sites where selected species were either known or likely to occur were targeted for detailed survey work. Habitat preferences for each species at existing sites were noted while new survey effort searched proximate areas of similar habitat.

2.4.2 Field surveys

Additional field surveys were conducted in order to target habitat for these species in August 1999. Searches of the study area were undertaken over fifteen days. The survey sites are listed in Table 12.

Table 12 Location of selected species searches (August 1999)

Sites
1. Cabramatta Creek Industrial Area
2. Carysfield Park, Bass Hill
3. East of Harris Creek, Holsworthy
4. Georges River Nature Reserve (Council), end of Georges River Rd
5. Georges River Nature Reserve (Council), Picnic Park, Ingleburn
6. Louisa Reserve, Bass Hill
7. Macquarie Fields/Bunbury Curran Creek
8. Maxwells Creek/Liverpool Showground, off Jedda Rd
9. North of Artillery Rd, Holsworthy Military Area
10. North of Florence Ave., Minto Heights
11. Simos Beach Nature Reserve (Council), Macquarie Fields
12. Voyager Point, off Heathcote Rd
13. Voyager Point
14. Williams Creek and National Park Rd, Holsworthy
15. Yeramba Lagoon, Picnic Point

2.4.3 Species modelling package

The aim of the modelling process was to examine statistical relationships between species locations and layers of geographical data which describe the broad physical and climatic features of the landscape. Once identified these relationships can be extrapolated across the landscape to indicate the relative likelihood of species occurrence using GIS. The species modelling program (Watson1996) was used to carry out these tasks. This program forms a module of S-Plus software and can be integrated with ArcView mapping software using the Species Predict extension.

The framework for the usage of this package in the context of this report is outlined in the Eden Fauna Modelling Project (NPWS 1997b). The program interpolates species locations with environmental variables and assigns each variable a species significance factor. This significance factor relates to the relative importance of the variables for overall species distribution. Environmental variables with high significance factors are assumed to be the most important for probable species occurrence. The values of the significant variables at each site are then analysed to produce a spatial probability or likelihood of occurrence scale across the landscape (NPWS 1997b). In this process the data range of the environmental variables corresponding to known species locations is obtained and mapped across the landscape. These data values correspond to the optimum range of potential habitat for each particular species. Thus, the relative probability of occurrence is highest within the optimum range for the most significant environmental variables. The resolution and accuracy of the environmental variables is therefore an important issue in the modelling process and always remains inherent in the output produced. Appendix B lists the environmental variables (spatial data) used in the modelling process.

The output from this modelling process is twofold. Firstly, there is a map of the area spatially displaying the relative probability of species occurrence. Secondly, there is additional mathematical information that describes the performance of the model. This information is in the form of a graph for each environmental variable used in the model as well as in the form of textual data for the deviance, model type and overall fit of the model. The graph shows the values for the individual variable in relation to the probability of occurrence and the location of sites. The textual information provides mathematical data regarding the overall fit of the model regarding how accurately the probability of occurrence can be modelled by the variables chosen.

Deviance is a measure of how accurately the modelling procedure has represented the data. The Null Deviance is an explanation of the inherent noise of the data prior to analysis. The Residual Deviance is a measure of the deviance or closeness of fit after the model has been run. Thus, the Deviance Explained is the representation of the overall fit of the model to the data. Deviance Explained is calculated as a percentage of the Null Deviance and Residual Deviance, where a perfectly fitted model would have 100% of its deviance explained. Most models are typically much less than this and are a result of the amount of variation within the data before modelling (NPWS1997b).

The modelling process outlined above is however limited by the number of quality site locations and the ability for the computer to accurately measure a statistical relationship between the sites and environmental variables. To overcome a lack of site locations for a number of species, a more ad hoc approach was taken. This is termed an expert model, where spatial information is manually analysed through an understanding of the habitat requirements and growth characteristics of the individual species. This is carried out through the collation of all available information, an examination of the written site descriptions and data values for each variable at every site, and investigation of habitat areas at known site locations. From this background, the major factors influencing the distribution of the species across the landscape are hypothesised and the spatial data is subsequently analysed until a sound match of the data to recorded species locations is found.

2.5 Data analysis

The data resulting from the API, vegetation and fauna surveys were collated in ArcView for analysis and storage. This database provides a means for viewing geographical information in map formats. Assessments of the complex data sets were then carried out using C-Plan, software developed by the NSW Government as a planning tool for regional conservation assessment.

Used in the Comprehensive Regional Assessments of the late 1990s in northern and southern New South Wales, C-Plan was designed to assist in the identification and prioritisation of lands for conservation. More specifically, it has been set up to identify lands that would add the greatest natural conservation value to the state reserve system. In doing this, the software makes a range of assumptions that are important to the way the software is used to assist the development of a biodiversity management framework for the GRC.

C-Plan makes the base assumption that lands already in the NSW reserve system, including lands that are administered under the *National Parks & Wildlife Act 1974* and lands administered as drinking water catchment areas, already provide an appropriate management framework for the conservation of biodiversity values. In considering the implications of this assumption, it is important to be cognisant of the fact that many biodiversity values are currently in decline, therefore it is logical to assume that the existing reserve system is unable to adequately manage the full range of biodiversity values. There is no assumption made that addition to the reserve system is the best way to manage those values. To the contrary, it is acknowledged that a range of management mechanisms must be adopted to manage biodiversity values across the landscape.

C-Plan adopts a range of criteria developed by the user to assess the biodiversity data. These criteria must be ordered and organised according to the priorities of the user, i.e. the stakeholders involved in developing a management framework. It can rank sites based on species richness as the dominant criterion. Under this criterion, species richness is used as a surrogate of biodiversity, i.e. it is assumed that the number of different species is a measure of biodiversity. This approach also accounts for ecosystem diversity to some degree, however there are no distinctions made in regard to genetic variability as this level of detail about species in the catchment has not been assessed.

Finally, C-Plan can incorporate weightings so that threatened species have greater priority in the management framework than more common species. In ranking the patches (planning units) C-Plan can assign greater value to patches that support habitat for threatened species, and even higher priority to sites known to support threatened species. Therefore in considering two similar sites, each of which has the same species richness (number of species), only one of which supports threatened species, the site supporting the threatened species can be assigned a higher rank. For example, with site 1 and site 2 below, site 2 can be ranked higher:

- site 1 has 86 recorded species, none of which are considered to be under significant threat; and
- site 2 has 86 recorded species and supports habitat for at least one threatened species.

It is important to remember that C-Plan is a tool to assist with the development of a management framework. If it is considered that the solutions proposed by C-Plan are inappropriate or do not sufficiently address the management goal then it is important to re-assess the criteria being used by C-Plan in making that assessment.