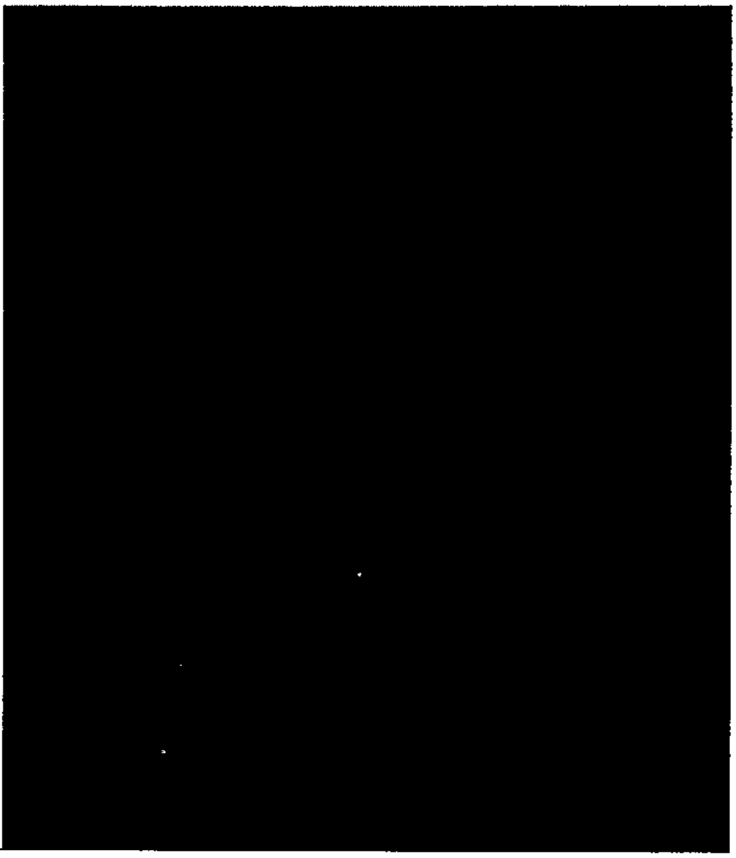
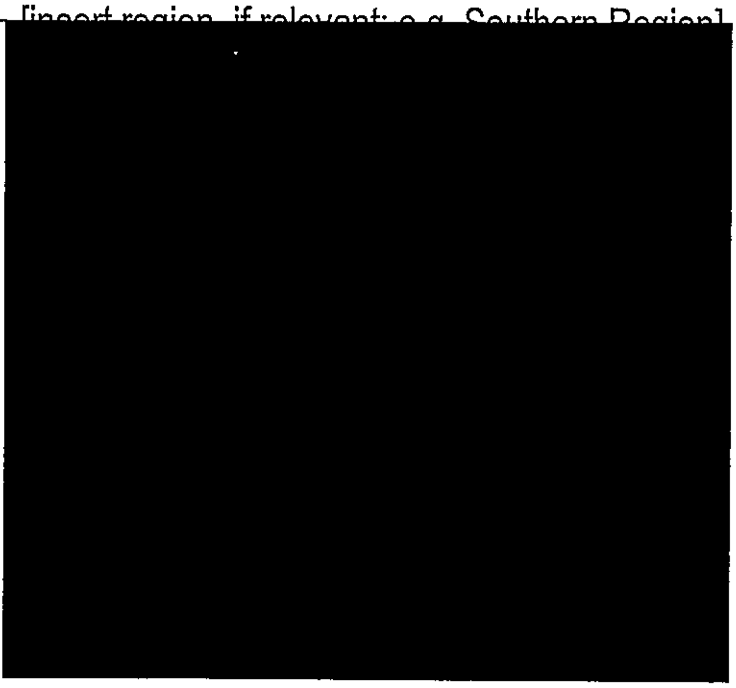


FOREST ECOSYSTEM CLASSIFICATION AND MAPPING FOR UPPER AND LOWER NORTH EAST CRA REGIONS

JANUARY 1999

[insert region, if relevant: e.g. Southern Region]



FOREST ECOSYSTEM CLASSIFICATION AND MAPPING FOR THE UPPER AND LOWER NORTH EAST CRA REGIONS

CRA UNIT, NORTHERN ZONE NATIONAL PARKS AND WILDLIFE SERVICE

A project undertaken for
the Joint Commonwealth NSW Regional Forest Agreement Steering Committee
as part of the
NSW Comprehensive Regional Assessments
project number NA35/EH

January 1999

For more information and for information on access to data contact the:

Resource and Conservation Division, Department of Urban Affairs and Planning

GPO Box 3927
SYDNEY NSW 2001

Phone: (02) 9228 3166
Fax: (02) 9228 4967

Forests Taskforce, Department of the Prime Minister and Cabinet

3-5 National Circuit
BARTON ACT 2600

Phone: 1800 650 983
Fax: (02) 6271 5511

ISBN 1 74029 039 9

This project has been jointly funded by the New South Wales and Commonwealth Governments and managed through the Resource and Conservation Division, Department of Urban Affairs and Planning, and the Forests Taskforce, Department of the Prime Minister and Cabinet

The project has been overseen and the methodology has been developed through the Environment and Heritage Technical Committee which includes representatives from the New South Wales and Commonwealth Governments and stakeholder groups.

Acknowledgement is due especially to all generous individuals and organisations who allowed us to use their hard earnt data and particularly: Steve Griffiths, John T Hunter, Coffs Harbour Council, SFNSW, and the Royal Botanic Gardens.

For their contribution to the implementation of this project, special acknowledgement is due to the following people:

Simon Ferrier for unsurpassed technical expertise and advice, methodological development and implementation and report writing

Veda Crossley for co-ordination of field survey work and data entry

Katrina McKay for report preparation, data manipulation, gap analysis and site selection and GIS support

Jessica Scott, Jo Spencer and Kelly Mantle for data manipulation, gap analysis and site selection

Sonia Diepeven and Jamie Love for data entry

Guy Hodgson for GIS support

Armidale GIS Unit for modelling, analysis and software development: Michael Drielsma, Jennie Pearce, Bill Forsyth, Giselle Whish and Glenn Manion

Lynne Dalton for administrative support

Doug Binns, Andrew Benwell, Phil Gilmour and Peter Richards for expert review and input into the forest ecosystem classification

Dr Jamie Kirkpatrick, Dr Mike Austin and Dr Mark Burgman for attendance at workshops and peer review

Van Klaphake, Darren Bailey, Annette McKinley, Barbara Stewart, John Nagle and Nick Cobcroft for botanical field survey and Sue Bower, Kevin Taylor, Jo Spencer, Margaret Wheeler, Chris Turbill, John and Corinne for field survey assistance

CONTENTS

PROJECT SUMMARY

This report describes a project undertaken as part of the comprehensive regional assessments of forests in New South Wales. The comprehensive regional assessments (CRAs) provide the scientific basis on which the State and Commonwealth Governments will sign regional forest agreements (RFAs) for major forest areas of New South Wales. These agreements will determine the future of these forests, providing a balance between conservation and ecologically sustainable use of forest resources.

Project objective/s

The scope of this work, as approved by the NSW CRA/RFA Environment and Heritage Technical Committee (EHTC), was to ''. Workshops involving EHTC and stakeholders were conducted at various key stages in the project and the methodology was subject to independent peer review.

Methods

Forest ecosystem classification in the north-east followed an approach recommended by a Forest Ecosystem Workshop convened by the Environment and Heritage Technical Committee in July 1997. Different approaches were approved by EHTC for three distinct biogeographic regions which are present within the north-east regions: the area south of the Hunter Valley, the area west of the New England highway, and the north-east area north of the Hunter Valley and east of the New England highway. An outline of the approach used in the southern area is provided in a separate report. For the north-east area, the approach entailed:

- the derivation of a forest ecosystem classification by splitting and amalgamation of existing SFNSW forest types based on analysis of variation between field survey plots in relation to environmental variables
- mapping of derived ecosystems within the existing mapped extent by use of decision rules relating variation to abiotic environmental variables
- predictive mapping of derived ecosystems across unmapped forest and cleared land based on modelling of the relationship between the mapped distribution of the ecosystem and abiotic variables

For the western area, the approach entailed:

- the derivation of a forest ecosystem classification by subjecting floristic data from field survey plots to numerical cluster analysis
- predictive mapping of derived ecosystems based on modelling of the relationship between the classified plots and abiotic environmental variables

A seamless vegetation coverage across the three distinct biogeographic areas was derived by expert integration of the disparate classifications.

One hundred and fifty-seven forest ecosystems were classified and mapped for the north-east area, including 141 dominated by eucalypts, and 16 dominated by non-eucalypt vegetation. Ninety-eight of the eucalypt dominated ecosystems were derived from splitting and amalgamation of forest types and descriptions of each of these ecosystems is provided in this report. The remaining 43 ecosystems comprised SFNSW forest types on which no splitting or amalgamation was conducted. Descriptions of these ecosystems is available in FCNSW (1989). A further 22 forest ecosystems were classified and mapped for the western area, including 21 dominated by eucalypts, and one shrubland ecosystem.

Key results and products

The resultant pre-1750 layer was refined in relation to historical data compiled from parish portion plans. The forest ecosystem map is available under licence from the NSW National Parks and Wildlife Service.

þ

2. METHODS FOR DERIVING FOREST ECOSYSTEMS

2.1 DATA AUDIT AND COLLATION

2.1.1 Fine scale vegetation mapping

Existing forest type mapping

Approximately 1,510,000 ha of the area east of the New England Highway and north of the Hunter River is covered by 1:25,000 forest type mapping based on the SFNSW Research Note 17 forest type classification. This coverage represents approximately 35% of the total area of extant forest in the region and 84% of the forest on public land (most of the mapping is within State Forests with some additional coverage of National Parks and other public land tenures). This forest type mapping was imported to ARCVIEW as a shape file coverage for stitching to other layers.

Other fine scale vegetation mapping

Some additional areas within the region have been mapped at a similar spatial and classification resolution to forest type mapping but using vegetation classes other than forest types. Where practicable such mapping was obtained in digital form and converted to forest type mapping by applying conversion tables (prepared by experts) assigning vegetation classes from other classifications to forest types. The mapped datasets which have been obtained, combined, and converted to RN17 forest types where required, are presented in Table 1. The total area of mapping generated by combining these other sources with existing forest type mapping is approximately 1,800,000 ha, covering 42% of all extant forest and 92% of forest on public land.

All available finescale vegetation mapping was conducted at a scale of 1:25 000 or finer. The vegetation mapping of Torrington State Recreation Area was conducted at a scale of 1:50 000. Areas with fine scale vegetation mapping have an estimated positional accuracy of map polygon boundaries of within 25m. All finescale vegetation mapping was imported into ARCVIEW as shape files and the vegetation type attribution of each mapping project was expertly converted to an analagous SFNSW RN17 forest type classification. The imported shapefiles were converted to ARCVIEW grids at a 50m resolution and merged into a single layer.

Table 1 Fine scale vegetation mapping which was collated for the CRA process

Data Set Source:	Upper		Lower	
	Eastern	Western	Eastern	Western
SFNSW Forest Typing and Royal Milli Typing	x		x	
Natural Resources Audit Council Multi-attribute Mapping	x			
Coffs Harbour Council Vegetation Mapping	x			
Henry James Tweed Vegetation Mapping	x		x	
Department of Land and Water Conservation Nambucca Vegetation Mapping			x	
National Parks and Wildlife Service Coastal Vegetation Mapping	x		x	
Vegetation mapping of Torrington State Recreation Area		x		

2.1.2 Floristic survey plots

A relatively large set of existing floristic field survey data was available for the region. Analyses of floristic variation within and between forest types is based on data from 0.1 ha plots within which the cover abundance (Braun-Blanquet index) of all vascular plant species has been recorded. Plot data of this type have been accumulated from a number of previous projects including the NPWS North East Forest Biodiversity Study, the NRAC Upper North East Regional Audit and the SFNSW EIS program. These data were further supplemented by extensive data collation in the CRA process. A list of all the floristic site data which was compiled and utilised in the CRA forest ecosystem derivation project is provided in Table 2.

The environmental and geographical spread of sites is generally good because much of the survey work has been specifically designed to stratify site locations in relation to major environmental and geographical gradients. Sites are also well distributed in relation to mapped forest types.

Table 2 Full floristic data collated for use in forest ecosystem derivation

Data Set Source:	Data Management	Numberof Plots	Type ofData	Plot Size	Reference
Flora Survey of Ben Halls Gap State Forest	Entered during NEFBS	21	Full floristics with abundances	20x20	Benson & Andrew 1990
Flora Survey of Broadwater National Park	Entered during NRAC	62	Full floristics with abundances and plot and structural data	20x20 and 10x10	Griffith 1985
Flora survey of Bundjalung National Park	Entered during NRAC	205	Full floristics with abundances and plot data	No fixed plot size	Griffith 1983
Flora Survey of the Coffs Harbour Local Government Area	Entered during CRA	108	Full floristics with cover abundance	20x50	Fisher & Gill 1996
Flora Survey of Demon Nature Reserve	Entered during CRA	39	Full floristic nested quadrats with frequency data converted to cover abundance	10 quadrats nested to 22.5mx22.5m	Hunter 1997

Eucalyptus dunnii survey	Entered during NEFBS	31	Full floristics with abundances	20x20	Benson & Hager 1993
John Hunter Granite Surveys	Entered during CRA	521	Full floristic nested quadrats with frequency data converted to cover abundance	10 quadrats nested to 22.5mx22.5m	Hunter 1997
Mount Neville Vegetation Survey	Entered during NEFBS	21	Full floristics with cover abundance	20X20	
Vegetation Survey of Myall Lakes National Park	Entered during CRA	34	Full floristics with cover abundance	20x20	NPWS 1997b
North East Forests Biodiversity Study Flora Sites	Entered during NEFBS	698	Full floristics with abundances and plot, physical and structural data	20x20 within 20x50	NPWS 1994
Natural Resources Audit Council Flora Survey	Entered during NRAC	418	Full floristics with abundances and plot, physical and structural data	20x20 within 20x50	NPWS 1995
Joint Old Growth Project Flora data	Entered during CRA	148	Full floristics with cover abundance; plot, physical and structural data	20x20 within 20x50	NPWS & SFNSW 1996
Hunter Valley Remnant Surveys	Entered during CRA	16	Full floristics with cover abundance	20x20	Peake & Roberts on 1998
Royal Botanic Garden Vegetation Data for the Guyra Mapsheet	Entered during NRAC	312	Full floristics with abundances; plot and structural data	20x20	Benson Ashby 1996
State Forest Environmental Impact Study and Monitoring Team flora data	Entered during NEFBS, NRAC and CRA	1,494	Full floristics with abundances; plot and structural data	20x20 within 20x50	Binns 1992; York, Binns & Shields 1991
Tomaree National Park Vegetation Survey	Entered during CRA	35	Full floristics with cover abundance	20x20	NPWS 1997a
Tweed Coast Vegetation Survey	Entered during NEFBS	53	Full floristics with abundances; plot and structural data	20x20 and 10x10	Pressey & Griffith 1992
Yuraygir National Park Flora Survey Sites	Entered during NEFBS and NRAC	186	Full floristics with abundances; plot and structural data	20x20 and 10x10	Griffith 1984
Eastlink Flora Survey	Entered during CRA	127	Full floristic nested quadrats with frequency data converted to cover abundance	10 quadrats nested to 32mx32m	Clarke *et al* 1995

Torrington State Recreation Area Vegetation Survey	Entered during CRA	201	Full floristics nested quadrats with frequency data converted to cover abundance	10 quadrats nested to 32mx32m	Clarke *et al* 1995
Total		4,730			

2.1.3 Abiotic environmental GIS layers

Complete regional GIS coverage were available for a large number of abiotic environmental variables, generally at a scale of 1:25,000 or 1:200,000. Abiotic environmental variables which were available for use in the analyses are presented in Table 3. For the eastern area, the majority of these variables were available at a 25m resolution. The coarsest variable was the geological layer which was infrequently used in the analysis. Latitude (northing) and longitude (easting) were also considered as additional explanatory variables in the analyses. For the purposes of analysis, all variables were re-sampled and used at 50m. The availability of environmental data for the western region was very restricted.

Table 3 Resolution of environmental variables available for the two bioregions

Environmental Variables	Resolution in Eastern Area (m)	Resolution in Western Area (m)
Mean annual rainfall	25	250
Mean Temperature	25	250
Minimum Temp of the Coldest Month	25	250
Ruggedness Indices	25	Not available
Slope	25	Not available
Solar Radiation Index	25	Not available
Topographic Indices	25	Not available
Topographic Wetness Index	25	Not available
Soil Moisture Index	200	Not available
Rainfall in the Driest Quarter	200	Not available
Soil Depth	200	Not available
Soil Fertility	200	200
Geological Classes	250	Not available
Topographic Position	25	Not available
Easting	25	25
Northing	25	25

2.1.4 Historical data

Historical data from parish portion plans was collated by SFNSW and is documented in a separate report. It involved collecting historical information on vegetation cover from a sample of original portion plans within parishes in the Upper North East and Lower North East regions. The utility of such data had been shown previously in work done by Ryan and Stubbs (1996) for an area comprising seven parishes in the County of Richmond in which it was concluded that "where the destruction of the vegetation has been complete, the historical record, and in particular the conditional purchase plans, is indispensable in reconstructing the pre-settlement pattern of vegetation". Information that was collected for each portion plan included the map sheet name, parish name, portion number, plan number, date of record, notation on vegetation type and topography, point data on corner trees used to fix survey points, comments or notes and grid references of corner survey points.

2.2 SITE SELECTION AND FIELD SURVEY

2.2.1 Site Selection

The environmental and geographical spread of the existing sites was generally good because much of the survey work had been specifically designed to stratify site locations in relation to major environmental and geographical gradients. Sites were also relatively well distributed in relation to mapped forest types. Survey work conducted by this CRA project was designed to fill environmental and/or geographical gaps in the coverage of sampling within each forest type, as well as major environmental and/or geographical gaps in sampling across parts of the region not covered by forest type mapping.

Automated gap analysis software designed by NPWS GIS Research and Development Unit was utilised to select sites (reference). Approximately 400 sites were selected by conducting gap analysis on a forest type by forest type basis. For each forest type the process would involve the use of the software to identify the most poorly sampled geographical/environmental envelope and to select a site in that envelope, and then to re-run the software to identify the next most poorly sampled envelope assuming that the previous selected site was to be surveyed. This process was conducted iteratively and at least 5 sites and up to 20 sites were placed in any given forest type depending on the level of sampling and extent of the type. Some very restricted types did not have any sites placed in them.

A further 200 sites were selected by conducting gap analysis across the entire region at once and iteratively implementing the process as described above. These sites were implemented to provide a sample of parts of the region not covered by forest type mapping. A further 48 sites were selected by conducting gap analysis on the western area only.

A further 51 sites were conducted in Barrington National Park and a total of 145 sites were conducted in total in Guy Fawkes, Chaelundi, Bellinger River and Nymboi-Binderay National Parks. These sites were implemented in conjunction with NPWS district offices who were undertaking Fire Management Planning and contracted the CRA Unit to conduct floristic surveys. Gap analysis was conducted on Barrington in isolation and the four Dorrigo National Parks as a whole and site selection implemented as described above.

2.2.2 Field Survey

The field survey methodology which was utilised followed the approach to plot-based sampling used in the North East Forests Biodiversity Study (NPWS 1994) and the Natural Resources Audit Council Vegetation Survey as outlined in NPWS (1995). The survey approach includes collection of floristic, structural, and physical data at a 20x20m plot nested within a 20x50m plot. Floristic data which is recorded includes all vascular plant species present at the plot and a visual estimation of the cover abundance of each species according to a modified Braun-Blanquet system of cover abundance classes (Mueller-Dombois in NRAC 48). Structural data includes the identification of the vegetation strata present in the plot and the predominant growth form, height range, percentage crown cover and three most dominant species within each stratum as well as diameter at breast height measurements and identification of the 12 upper stratum stems closest to the plot centre. Physical data which was recorded included altitude, slope, aspect, horizon elevation, soil depth and type, mapped geology and field geology and landform element. Other information which was collected included information on disturbance history and the overall condition of the site. For more detailed descriptions of the plot methodology and proformas refer to NPWS (1995).

A total of 820 full floristic plots were conducted during the CRA field surveys. This total was comprised of 576 in the north-east area, 48 plots in the western area, 145 plots in National Parks in the Dorrigo region and 51 plots in Barrington NP.

2.3 DATABASE DEVELOPMENT AND DATA CHECKING

2.3.1 Database development

An ACCESS 97 database was specifically designed for the storage, entry and manipulation of all CRA systematic flora data.

A data entry interface was developed to resemble the site proformas. These data entry forms were designed to constrain the type of data which could be entered in each field to only those options specified on the field proformas. For the majority of fields drop down boxes were used to constrain entry to a specified list of codes only. Other fields were constrained to the appropriate data type such as numeric or alphanumeric and restricted to a specified data range (such as >0.1 and <5). These constraints were designed to minimise data entry errors and in combination with systematic manual checking, proved to be effective in producing a quality, error free data entry process.

The database design was based on the Advanced Revelation (AREV) database which had been utilised by NPWS for the North East Forests Biodiversity Study (NPWS 1994) and the Natural Resources Audit Council Vegetation Survey as outlined in NPWS (1995). ACCESS 97 was preferred to AREV because it provided a friendlier user interface and more flexible approach to data manipulation, extraction and checking.

The database stored survey data in nine separate which were titled Sites, Vegetation Structure, Canopy Structure, Physical Attributes, Floristics, Disturbance Information, Environmental Variables, Survey Identification, and Recorders.

The database also included three libraries which related information held in codes in the tables to names and associated information. These libraries were a CAPS library (which related the Census of Australian Vascular Plant Codes to species names), a MAP library (which related 1:25,000 map codes and 1:100,000 map codes to map names) and an area library (which related area codes to National Park and State Forest names). These libraries were derived from the AREV database and the CAPS library was rigorously updated with reference to NPWS Head Office to account for recent changes in taxonomy and nomenclature.

2.3.2 Data checking

A series of manual and automatic checking procedures were undertaken on the complete collated database.

All imported and collated datasets were subject to the following checking procedures: taxonomic updates to ensure most recent nomenclature was used, automatic checks to ensure all records had species information; identification and rectification of all duplicate records as well as sites and modification of cover abundance information where required. The accuracy of grid references were checked by automatically comparing the 1:25,000 mapsheet number recorded in the field for a site, with that derived from the sites grid reference in the GIS. All anomalies were further checked and AMGs corrected where necessary.

Further checks were conducted to ensure that no species were duplicated in the floristics table, that all sites had complete information in each of the floristics table, vegetation structure table, canopy structure table, physical attributes table and modifications table. Checks were also conducted to ensure that all species in the floristics table had a cover abundance value and that all species codes in the floristics, vegetation structure and canopy tables were valid CAPS codes present in the CAPS library.

All site localities from the SFNSW data were manually checked against hard copy localities as presented in the EIS reports (where available) in addition to the automatic checking described above. Discrepancies were identified and corrected where possible or quarantined from analysis.

All State Forest grid references were checked in the GIS to ensure that they fell within State Forest boundaries. All sites which were outside State Forest boundaries were identified, checked and corrected or quarantined.

All NRAC sites were cross-referenced in a GIS to ensure that they fell within the designated reserve or parcel of land identified in the reserve area field. Anomalies were checked and corrected or quarantined from analysis.

2.3.3 Completed database

The completed database contained a total of 5,532 full floristic survey sites with abundance information which were utilised for the analysis described below. The database contained a further 835 rainforest sites conducted using a bounded irregular traverse which were not utilised in this project. The database also contained a further 2,270 floristic sites at which canopy only data or full floristics data without abundance information had been collected. These sites were derived from a number of sources with different plot specifications and were not utilised in the analysis for this project.

2.4 METHODOLOGICAL DEVELOPMENT

2.4.1 Analytical approach

The basic aim of this analysis was to use floristic survey data to guide decisions on splitting or amalgamating existing SFNSW forest types to yield a forest ecosystem classification suitable for use in CRA/RFA assessments. Forest types needed to be split and/or amalgamated in such a way that resulting forest ecosystems were relatively homogeneous in terms of floristic composition, and that this level of homogeneity was reasonably consistent across all derived forest ecosystems. A further requirement was that splits within a forest type were able to be mapped, and therefore were associated with variation in mapped abiotic environmental variables.

Any analysis of floristic variation between survey sites must be based on some objective measure of difference in floristic composition or 'compositional dissimilarity' (Faith *et al.* 1987). The measure used in the current analyses is the Bray-Curtis index, demonstrated by Faith *et al.* to be a robust measure of compositional dissimilarity (and therefore ecological distance). The first step in all of the current analyses involves calculating the Bray-Curtis dissimilarity index for all possible pairs of sites being analysed. For example, if four sites are being analysed then the index is calculated for all 6 possible pairs of sites, namely Site 1 vs Site 2, Site 1 vs Site 3, Site 1 vs Site 4, Site 2 vs Site 3, Site 2 vs Site 4 and Site 3 vs Site 4. This procedure yields a 'sites by sites' dissimilarity matrix.

The approach adopted groups sites by considering both floristic and environmental data simultaneously. This new technique can therefore be thought of as a type of 'constrained' or 'canonical' numerical classification. In the same way that canonical ordination techniques (e.g. Canonical Correspondence Analysis, ter Braak 1986) fit an ordination to floristic data such that the ordination axes are functions of environmental variables, the classification technique employed here uses floristic data to divide sites into groups such that this grouping is also defined in terms of decision rules based on environmental variables.

Consider an example in which we wish to analyse the potential for a forest type to be split into two or more sub-types (or 'forest ecosystems'). Floristic data have been collected at a number of plots scattered environmentally and geographically throughout the mapped distribution of the forest type, and these data have been used to derive a dissimilarity matrix containing floristic dissimilarities (Bray-Curtis index) between all possible pairs of sites. The adopted strategy is to search through the set of all possible binary environmental splits for that split which maximises a measure of floristic difference between the two resulting groups of sites, relative to the floristic variation exhibited within these groups. Each binary environmental split is defined in terms of a cutpoint, such as 'soil moisture index = 0.75' (this cutpoint would divide survey sites into two groups, those with a soil moisture index less than 0.75 and those with an index greater than or equal to 0.75).

The floristic difference between the two groups formed by a split is measured using the statistic:

$$D = \bar{d}_B - \bar{d}_W$$

where $\bar{d}_B$, the average 'between group' dissimilarity, is defined as the average dissimilarity between pairs of sites on opposite sides of the environmental cutpoint (i.e. one site in group 1 and the other site in group 2) and $\bar{d}_W$, the average 'within group' dissimilarity, is defined as the average dissimilarity between pairs of sites on the same side of the cutpoint (i.e. either both sites in group 1 or both sites in group 2). The statistical significance of D is estimated using a Monte Carlo randomization procedure (Manly 1991) in which D is repeatedly calculated after randomly permuting the assignment of sites to groups. The value of D obtained using the real grouping of sites is then compared to the distribution of D obtained using random permutations. In the simplest form of this test the null and alternative hypotheses (H_0 and H_1) are:

$$H_0 : D = 0$$

$$H_1 : D > 0$$

This approach to measuring, and testing the significance of, differences between groups based on inter-site dissimilarities closely resembles the ANOSIM procedure described by Clarke and Green (1988) and Clarke (1993), which in turn appears to be an unattributed reinvention of an approach originally described by Mielke *et al.* (1976). The approach to significance testing employed by these authors has been generalised in the currrent analysis to allow testing of a wider range of hypotheses of the form:

$$H_0 : D \le D_t$$

$$H_1 : D > D_t$$

where D_t is a specified parametric value (or threshold) with which we wish to compare the observed value of D. We are therefore testing whether the observed value of D is greater than the specified threshold by estimating the probability (Type I error) that the true value of D is actually less than or equal to the threshold (Sokal and Rohlf 1981). By assigning D_t a value greater than zero splitting of types will occur only where the floristic difference between resulting groups is significantly greater than the specified threshold, not just significantly greater than zero. This helps to overcome the problem that a type containing a large number of survey sites may be split into groups exhibiting a floristic difference significantly greater than zero yet this difference is inconsequentially small.

The splitting procedure just described can be applied iteratively. Each group of sites formed by splitting can itself be subjected to further splitting. Iterative splitting generates a hierarchical classification of floristic groups in which each division in the hierarchy is defined in terms of an environmental decision rule. This strategy effectively combines elements of divisive polythetic classification (e.g. Lance and Williams 1975) with elements of decision tree modelling (e.g. Moore *et al.* 1991).

The approach to significance testing described above can also be used to investigate potential amalgamations of forest types or sub-types formed by initial splitting of different forest types. Floristic data from sites within the two groups being considered for amalgamation are used to calculate D, thereby providing a measure of floristic difference between the groups. Significance testing in this case requires a different configuration of null and alternative hypotheses than that used for splitting types. The appropriate hypotheses are:

$$H_0 : D \ge D_t$$

$$H_1 : D < D_t$$

In other words, we are testing whether the observed value of D is significantly less than a specified threshold by estimating the probability that the true value of D is actually greater than or equal to the threshold.

The new analytical approach represents a major improvement over the approach used to split and amalgamate north-east NSW forest types during the Interim Forest Assessment (NPWS 1996). Specific advantages of the new technique are:

- The technique provides an explicit and objective basis for making decisions on splitting or amalgamating forest types.
- The technique incorporates rigorous statistical significance testing.
- The technique considers floristic and environmental data simultaneously in a single integrated analysis. Emphasis is therefore placed on ensuring that splits within forest types are not only meaningful in terms of floristic variation but can also be defined in terms of environmental variation, and thereby mapped. The technique effectively filters out floristic variation which cannot be accounted for in terms of available environmental variables, and therefore cannot be mapped. (Some of this unexplained variation may relate to fine scale

3. METHODS FOR MAPPING FOREST ECOSYSTEMS

3.1 MAPPING AREAS WITH EXISTING FINE SCALE MAPPING

Mapping of derived forest ecosystems was most readily achieved in areas covered by existing SFNSW forest type (or equivalent) mapping. Splits in existing types were mapped for each type by using the GIS to apply the environmental decision rules defining these splits. This operation was automated in the software developed for the project (see 2.4.20). Since the derivation of ecosystems often involved iterative subdivisions within each original forest type, the final ecosystems were often derived by iterative applications of binary divisions of environmental variables. Amalgamations of types or sub-types were also mapped through application of a GIS merging operation. All final derived types were merged back into the original layer with the forest types that had not been subject to any splitting or amalgamation. This provided a complete coverage of derived forest ecosystems across all areas with existing fine scale mapping.

3.2 DERIVING PRE-1750 DISTRIBUTION MODELS

3.2.1 Modelling of mapped forest ecosystems in relation to abiotic environmental variables

The distribution of each derived forest ecosystem was modelled in relation to abiotic environmental variables using data extracted from areas covered by existing SFNSW forest type mapping. For each forest ecosystem a random sample of 1ha grid cells was drawn from all cells mapped as containing that ecosystem. A second sample of cells was drawn from all cells mapped as not containing the ecosystem. Samples were selected in a manner which minimizes problems of spatial autocorrelation and model overfitting. A logistic regression model relating the probability of presence of each forest ecosystem to abiotic environmental and geographical variables (see2.1.3 for description of variables) was then fitted using generalised additive modelling (Yee and Mitchell 1991), a technique already applied extensively by NPWS in forest assessment work in NSW. The modelling was conducted via a modelling module (produced by Watson, 1996) which fitted regression models under S-PLUS statistical software (StatSci, 1995).

The predictive accuracy of ecosystem-environment models can be inferred from work conducted previously on species-environment models and information derived during the modelling process. Confidence limits were estimated for each of the probability surfaces interpolated from species-environment models. These indicate the prediction error expected throughout the study area. Another useful measure estimated for all fitted models is the percentage of deviance explained by the model. These measures should be interpreted with some caution as they measure prediction error by simple resubstitution which tend to underestimate the true prediction error of the model. Three additional measures are also produced during the modelling process which describe the performance of the model in terms of model discrimination, calibration and refinement.

3.2.2 Modelling of western ecosystems in relation to abiotic environmental variables

The distribution of each derived forest ecosystem in the western subregion was modelled in relation to abiotic environmental variables using data extracted from the sites within each ecosystem instead of the mapped extent of the ecosystems as was used in the eastern subregion. There were only four broad environmental predictors available for use in the western subregion compared with the 16 predictors available for the eastern subregion. The resolution of the western variables was also coarse in comparison with the eastern variables (see 2.1.3). Apart from these differences, the modelling was conducted using the same approach as that described above.

3.2.3 Extrapolation of models

These fitted models were then used to extrapolate the distribution of forest ecosystems across unmapped areas of forest and cleared land. Extrapolation within the modelling module was conducted using ARCVIEW Spatial Analyst (ESRI, 1996). The modelling resulted in a probability surface or extrapolated distribution for each forest ecosystem at 100m resolution. The modelling for the western ecosystems was extended to the eastern edge of the New England Tablelands instead of cutting off at the New England Highway (see2.7).

There were a number of forest ecosystems for which models could not be fitted. For a number of restricted types it was not possible to predict pre-1750 distributions robustly because of the very small size of the sample provided by the current distribution. For a number of other ecosystems, it was not possible to find sufficiently strong environmental relationships on which to develop a model. The ecosystems which were not modelled for these two reasons are:

- Type 10: Black Sallee
- Type 12: Blue Mountain Ash
- Type 13: Blue-leaved Stringybark
- Type 39: Dry Heathy New England Stringybarks
- Type 46: Eastern Red Gums
- Type 133: Snow Gum – Black Sallee
- Type 158: Wet Spotted Gum – Tallowwood

In addition to these, fitted models were not able to be derived for non-eucalypt dominated ecosystems. The CRA process was conducted under very tight timeframes and the emphasis was on forests and eucalypt-dominated vegetation. Therefore, the classification used for non-eucalypt ecosystems was very broad and applies virtually at a formation level. Since these ecosystems are so broad they generally contain major floristic variation within them and this variation is paralleled by significant environmental heterogeneity. It is generally not possible to derive robust models for types with such wide environmental characteristics. The ecosystems which were not modelled for this reason are:

- Type 16: Bull Oak
- Type 18: Casuarina Woodland
- Type 22: Coast Cypress Pine
- Type 64: Heath
- Type 66: Herbfield and Fjaeldmark
- Type 77: Mangrove
- Type 96: Natural Grassland
- Type 112: Paperbark
- Type 121: Rock

- Type 125: Saltbush
- Type 141: Swamp
- Type 151: Wattle
- Type 168: Rainforest
- Type 169: Scrub

3.3 MAPPING IN UNMAPPED AREAS AND CLEARED LAND

3.3.1 Derivation of a pre-1750 forest ecosystem layer

A single layer depicting the pre-1750 distribution of each forest ecosystem in the eastern region was derived from the overlay of all the forest ecosystem probability surfaces by randomly proportinally assigning each gridcell to a forest ecosystem according to the relative probabilities of each ecosystem at that gridcell. It is important to keep in mind that for areas without fine scale vegetation mapping the modelled distributions were used to *predict the proportion* of a modelled ecosystem only. No attempt was made to pinpoint the exact geographical location of these proportional allocations. Therefore, the nature of the random, proportional assignment process which was utilised to derive the most accurate areal figures, means that the exact spatial representation of the data is not designed to be accurate. Whilst areal calculations derived from such an approach are valid and reliable, any printed map is only one of many equally valid representations.

Integration of eastern and western map layers

The same approach was used to derive a single layer depicting the pre-1750 distribution of each forest ecosystem in the western area. The classifications for the two areas or bioregions had already been converted to an single classification (see2.8). However, since the western analysis had encompassed the entire Tablelands area instead of the area west of the New England Highway only, there was a zone of overlap between the two mapping schemes from the New England Highway east to the eastern edge of the New England Highway. Therefore, a merge operation was conducted in ARCVIEW GIS to combine the two layers and in the overlap zone the western ecosystems overrode the mapped distribution of eastern ecosystems only for those ecosystems for which the western ecosystem models were deemed (via expert opinion) more robust then the eastern ecosystem models.

Integration of historical data

An extensive analysis of data from historical portion plans was used to inform the pre-1750 distribution of eucalypt forest vegetation and thus to constrain the modelling of eucalypt-dominated forest ecosystems. For a description of the data collection methodology see SFNSW (in prep).

Each data point from a sample of historical portion plans was assigned to open eucalypt forest or non eucalypt vegetation based predominantly on interpretation of the corner tree type recorded by the surveyors and secondarily on the description provided by the surveyors. From this information, the proportion of open eucalypt forest to non eucalypt vegetation was calculated for each vegetation unit derived during the Interim Assessment Process (NPWS 1996). Vegetation units for which no historical portion plan data was collected, were assigned the proportions of their nearest neighbour in the dendrogram for which data was available. This then provided a full coverage of the likely proportion of pre-1750 eucalypt forest on a vegetation unit by vegetation unit basis. Gridcells were then randomly proportionally allocated to eucalypt forest or not according to the vegetation unit proportion. The non-eucalypt gridcells were then cut out from the pre-1750 ecosystem layer within cleared land and did not contribute to the derivation of pre-1750 area values for eucalypt ecosystems.

Integration of fine scale map information for non-eucalypt ecosystems

Existing fine scale forest ecosystem mapping where it was available was then merged over the top of the complete pre-1750 modelled coverage. Mapping for non-eucalypt dominated ecosystems for areas outside existing fine scale mapping was then incorporated into the coverage from the CRA Aerial Photograph Interpretation Project (CRAFTI). This was available for the Upper North East Region only. Rainforest mapping across all tenures from the CRAFTI project was merged into the derived layer and other non-eucalypt ecosystems were merged into the layer only where they occurred on vegetation outside existing fine scale mapping. Therefore, the CRAFTI rainforest mapping was used to over-ride existing fine scale mapping attribution where it overlapped with such mapping, whilst other non-eucalypt dominated vegetation mapping did not and was itself over-ridden by the existing fine scale map information.

In the Lower North East region, where the CRAFTI data was unavailable, rainforest mapping from the Broad Old Growth Mapping Project (RACAC 1996) was merged into the derived layer. This mapping was available for public land only. The only additional non-eucalypt dominated vegetation mapping which was available was the Eastern Bushlands Database which was captured from Landsat TM at a scale of 1:100,000. This layer contained two map units which included significant areas of non-forest ecosystems: Coastal Sclerophyll Complex and Plateau/Rocky Complex. These map units included non-forest vegetation within mosaics of forested vegetation. A profile of each of the two map units was derived by analysis of the proportions of forest to non-forest within each of the two map units in areas covered by existing fine-scale mapping. The profiles so derived were then applied to areas outside existing fine-scale mapping by random proportional allocation to non-forest or eucalypt forest within each map unit according to the derived proportions. The non-eucalypt areas so identified were then cut out from the ecosystem layer and did not contribute to the derivation of pre-1750 area values for eucalypt ecosystems.

3.3.2 Derivation of an extant forest ecosystem layer

For the Upper North East Region, the final extant forest ecosystem layer was derived by masking the pre-1750 ecosystem layer with the extant forest layer from the CRAFTI project. This layer was derived at a scale of 1:25,000 and mapped patches of extant forest down to 10ha in size. For the Lower North East region, the final extant forest ecosystem layer was derived by masking the pre-1750 ecosystem layer with the extant forest layer from the Eastern Bushlands Database (NPWS 1994b). This layer was derived at a scale of 1:100,000.

4. METHODS FOR MAPPING FOREST ECOSYSTEMS

4.1 MAPPING AREAS WITH EXISTING FINE SCALE MAPPING

Mapping of derived forest ecosystems was most readily achieved in areas covered by existing SFNSW forest type (or equivalent) mapping. Splits in existing types were mapped for each type by using the GIS to apply the environmental decision rules defining these splits. This operation was automated in the software developed for the project (see 0). Since the derivation of ecosystems often involved iterative subdivisions within each original forest type, the final ecosystems were often derived by iterative applications of binary divisions of environmental variables. Amalgamations of types or sub-types were also mapped through application of a GIS merging operation. All final derived types were merged back into the original layer with the forest types that had not been subject to any splitting or amalgamation. This provided a complete coverage of derived forest ecosystems across all areas with existing fine scale mapping.

4.2 DERIVING PRE-1750 DISTRIBUTION MODELS

4.2.1 Modelling of mapped forest ecosystems in relation to abiotic environmental variables

The distribution of each derived forest ecosystem was modelled in relation to abiotic environmental variables using data extracted from areas covered by existing SFNSW forest type mapping. For each forest ecosystem a random sample of 1ha grid cells was drawn from all cells mapped as containing that ecosystem. A second sample of cells was drawn from all cells mapped as not containing the ecosystem. Samples were selected in a manner which minimizes problems of spatial autocorrelation and model overfitting. A logistic regression model relating the probability of presence of each forest ecosystem to abiotic environmental and geographical variables (see 0 for description of variables) was then fitted using generalised additive modelling (Yee and Mitchell 1991), a technique already applied extensively by NPWS in forest assessment work in NSW. The modelling was conducted via a modelling module (produced by Watson, 1996) which fitted regression models under S-PLUS statistical software (StatSci, 1995).

The predictive accuracy of ecosystem-environment models can be inferred from work conducted previously on species-environment models and information derived during the modelling process. Confidence limits were estimated for each of the probability surfaces interpolated from species-environment models. These indicate the prediction error expected throughout the study area. Another useful measure estimated for all fitted models is the percentage of deviance explained by the model. These measures should be interpreted with some caution as they measure prediction error by simple resubstitution which tend to underestimate the true prediction error of the model. Three additional measures are also produced during the modelling process which describe the performance of the model in terms of model discrimination, calibration and refinement.

4.2.2 Modelling of western ecosystems in relation to abiotic environmental variables

The distribution of each derived forest ecosystem in the western subregion was modelled in relation to abiotic environmental variables using data extracted from the sites within each ecosystem instead of the mapped extent of the ecosystems as was used in the eastern subregion. There were only four broad environmental predictors available for use in the western subregion compared with the 16 predictors available for the eastern subregion. The resolution of the western variables was also coarse in comparison with the eastern variables (see 0). Apart from these differences, the modelling was conducted using the same approach as that described above.

4.2.3 Extrapolation of models

These fitted models were then used to extrapolate the distribution of forest ecosystems across unmapped areas of forest and cleared land. Extrapolation within the modelling module was conducted using ARCVIEW Spatial Analyst (ESRI, 1996). The modelling resulted in a probability surface or extrapolated distribution for each forest ecosystem at 100m resolution. The modelling for the western ecosystems was extended to the eastern edge of the New England Tablelands instead of cutting off at the New England Highway (see 0).

There were a number of forest ecosystems for which models could not be fitted. For a number of restricted types it was not possible to predict pre-1750 distributions robustly because of the very small size of the sample provided by the current distribution. For a number of other ecosystems, it was not possible to find sufficiently strong environmental relationships on which to develop a model. The ecosystems which were not modelled for these two reasons are:

- Type 10: Black Sallee
- Type 12: Blue Mountain Ash
- Type 13: Blue-leaved Stringybark
- Type 39: Dry Heathy New England Stringybarks
- Type 46: Eastern Red Gums
- Type 133: Snow Gum – Black Sallee
- Type 158: Wet Spotted Gum – Tallowwood

In addition to these, fitted models were not able to be derived for non-eucalypt dominated ecosystems. The CRA process was conducted under very tight timeframes and the emphasis was on forests and eucalypt-dominated vegetation. Therefore, the classification used for non-eucalypt ecosystems was very broad and applies virtually at a formation level. Since these ecosystems are so broad they generally contain major floristic variation within them and this variation is paralleled by significant environmental heterogeneity. It is generally not possible to derive robust models for types with such wide environmental characteristics. The ecosystems which were not modelled for this reason are:

- Type 16: Bull Oak
- Type 18: Casuarina Woodland
- Type 22: Coast Cypress Pine
- Type 64: Heath
- Type 66: Herbfield and Fjaeldmark
- Type 77: Mangrove
- Type 96: Natural Grassland
- Type 112: Paperbark
- Type 121: Rock
- Type 125: Saltbush
- Type 141: Swamp

- Type 151: Wattle
- Type 168: Rainforest
- Type 169: Scrub

4.3 MAPPING IN UNMAPPED AREAS AND CLEARED LAND

4.3.1 Derivation of a pre-1750 forest ecosystem layer

A single layer depicting the pre-1750 distribution of each forest ecosystem in the eastern region was derived from the overlay of all the forest ecosystem probability surfaces by randomly proportinally assigning each gridcell to a forest ecosystem according to the relative probabilities of each ecosystem at that gridcell. It is important to keep in mind that for areas without fine scale vegetation mapping the modelled distributions were used to *predict the proportion* of a modelled ecosystem only. No attempt was made to pinpoint the exact geographical location of these proportional allocations. Therefore, the nature of the random, proportional assignment process which was utilised to derive the most accurate areal figures, means that the exact spatial representation of the data is not designed to be accurate. Whilst areal calculations derived from such an approach are valid and reliable, any printed map is only one of many equally valid representations.

Integration of eastern and western map layers

The same approach was used to derive a single layer depicting the pre-1750 distribution of each forest ecosystem in the western area. The classifications for the two areas or bioregions had already been converted to an single classification (see 0). However, since the western analysis had encompassed the entire Tablelands area instead of the area west of the New England Highway only, there was a zone of overlap between the two mapping schemes from the New England Highway east to the eastern edge of the New England Highway. Therefore, a merge operation was conducted in ARCVIEW GIS to combine the two layers and in the overlap zone the western ecosystems overrode the mapped distribution of eastern ecosystems only for those ecosystems for which the western ecosystem models were deemed (via expert opinion) more robust then the eastern ecosystem models.

Integration of historical data

An extensive analysis of data from historical portion plans was used to inform the pre-1750 distribution of eucalypt forest vegetation and thus to constrain the modelling of eucalypt-dominated forest ecosystems. For a description of the data collection methodology see SFNSW (in prep).

Each data point from a sample of historical portion plans was assigned to open eucalypt forest or non eucalypt vegetation based predominantly on interpretation of the corner tree type recorded by the surveyors and secondarily on the description provided by the surveyors. From this information, the proportion of open eucalypt forest to non eucalypt vegetation was calculated for each vegetation unit derived during the Interim Assessment Process (NPWS 1996). Vegetation units for which no historical portion plan data was collected, were assigned the proportions of their nearest neighbour in the dendrogram for which data was available. This then provided a full coverage of the likely proportion of pre-1750 eucalypt forest on a vegetation unit by vegetation unit basis. Gridcells were then randomly proportionally allocated to eucalypt forest or not according to the vegetation unit proportion. The non-eucalypt gridcells were then cut out from the pre-1750 ecosystem layer within cleared land and did not contribute to the derivation of pre-1750 area values for eucalypt ecosystems.

Integration of fine scale map information for non-eucalypt ecosystems

Existing fine scale forest ecosystem mapping where it was available was then merged over the top of the complete pre-1750 modelled coverage. Mapping for non-eucalypt dominated ecosystems for areas outside existing fine scale mapping was then incorporated into the coverage from the CRA Aerial Photograph Interpretation Project (CRAFTI). This was available for the Upper North East Region only. Rainforest mapping across all tenures from the CRAFTI project was merged into the derived layer and other non-eucalypt ecosystems were merged into the layer only where they occurred on vegetation outside existing fine scale mapping. Therefore, the CRAFTI rainforest mapping was used to over-ride existing fine scale mapping attribution where it overlapped with such mapping, whilst other non-eucalypt dominated vegetation mapping did not and was itself over-ridden by the existing fine scale map information.

In the Lower North East region, where the CRAFTI data was unavailable, rainforest mapping from the Broad Old Growth Mapping Project (RACAC 1996) was merged into the derived layer. This mapping was available for public land only. The only additional non-eucalypt dominated vegetation mapping which was available was the Eastern Bushlands Database which was captured from Landsat TM at a scale of 1:100,000. This layer contained two map units which included significant areas of non-forest ecosystems: Coastal Sclerophyll Complex and Plateau/Rocky Complex. These map units included non-forest vegetation within mosaics of forested vegetation. A profile of each of the two map units was derived by analysis of the proportions of forest to non-forest within each of the two map units in areas covered by existing fine-scale mapping. The profiles so derived were then applied to areas outside existing fine-scale mapping by random proportional allocation to non-forest or eucalypt forest within each map unit according to the derived proportions. The non-eucalypt areas so identified were then cut out from the ecosystem layer and did not contribute to the derivation of pre-1750 area values for eucalypt ecosystems.

4.3.2 Derivation of an extant forest ecosystem layer

For the Upper North East Region, the final extant forest ecosystem layer was derived by masking the pre-1750 ecosystem layer with the extant forest layer from the CRAFTI project. This layer was derived at a scale of 1:25,000 and mapped patches of extant forest down to 10ha in size. For the Lower North East region, the final extant forest ecosystem layer was derived by masking the pre-1750 ecosystem layer with the extant forest layer from the Eastern Bushlands Database (NPWS 1994b). This layer was derived at a scale of 1:100,000.