AN ASSESSMENT OF MAP VALIDATION TECHNIQUES: A CASE STUDY USING A VEGETATION MAP OF LITCHFIELD NATIONAL PARK, NORTHERN TERRITORY, AUSTRALIA.

A thesis in the Environmental Management Programme

Charles Darwin University

Faculty of Engineering, Health, Science and Environment

Master of Environment Management (MEM) Programme

By Hardev Singh

Student No: S267203

Supervisors: Dr Andrew Edwards from Darwin Centre for Bushfire Research (DCBR) CDU.

Dr Mandy Trueman from Department of Land Resource Management (DLRM) NT.

July 2016
STATEMENT OF AUTHORSHIP

I declare that this thesis is my own work and has not submitted in any form for any other degree or diploma at any university or other institutes of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text and list of references.

Signed: 

Date: 15th July, 2016
ACKNOWLEDGEMENT

I acknowledge, with gratitude, my debt of thanks to School of Environmental Science, Charles Darwin University for the enormous support to successfully complete my research project.

I am thankful to the Department of Land Resource Management (DLRM), NT and Darwin Centre for Bushfire Research (DCBR), CDU for their joint support to carry out this research by arranging both ground and aerial survey, also for providing data and other required instruments.

I am grateful to my supervisors, Dr Andrew Edwards and Dr Mandy Trueman for their continual encouragement, aid and foresight feedback on my work.

I am highly obliged to Andrew for his dedicated effort towards my academic research project and tireless feedback on my writings. This research project would have never been complete without his constant support, supervision and constructive suggestions.

My advisor, Dr Mandy Trueman, has been my guide since the beginning of this research. She has polished my skills for field survey planning, executing and successfully implementing the ideas into reality. Her scientific guidance and expertise have ignited a spirit of research in me.

My sincere gratitude to Mr. Nicholas Cuff (Senior Botanist, NT Herbarium) for his compliments, suggestions and feedback throughout my research project. His passion and dedication towards the research have renovated my vision to learn and share the knowledge acquired.

I am also thankful to Dr Penny Wurm, Mr. Julian Gorman and Mr. Keith Christian for their academic support and coordination of this research unit. I convey my gratitude to Ms. Dominique Lynch, (Darwin Centre for Bushfire research), for co-ordinating the field survey.

Heartfelt thanks to the CDU library staff and team of Academic Language and Learning Support Program (ALLSP).
DEDICATION

This thesis is dedicated to my family especially my loving Father and Mother whose constant prayers and words of encouragement gave me strength and hope during the entire period of my studies. I have special dedication to my nieces Geet and Kanika whose sweet voices every time I heard on phone made me feel good throughout the time of my studies. Lastly I would like to dedicate this thesis to the CDU staff and friends for their support and encouragement.
TABLE OF CONTENTS
LIST OF FIGURES............................................................................................................. vii
LIST OF TABLES ................................................................................................................... vii
ACRONYMS .......................................................................................................................... ix
ABSTRACT ............................................................................................................................. 1

CHAPTER 1 - INTRODUCTION .............................................................................................. 2
  1.1 VEGETATION MAPPING .................................................................................................. 2
    1.1.1 History of vegetation mapping ................................................................................. 3
    1.1.2 Vegetation classification ......................................................................................... 4
    1.1.3 Techniques and methods for vegetation mapping .................................................... 5
    1.1.4 Application of satellite imagery for vegetation mapping .......................................... 6
  1.2 MAP ACCURACY ASSESSMENT .................................................................................... 7
    1.2.1 Mapping validation techniques ............................................................................. 8
  1.3 CASE STUDY: LITCHFIELD NATIONAL PARK VEGETATION MAP ......................... 11
  1.4 AIMS AND OBJECTIVES ............................................................................................. 12
    Research aim .................................................................................................................. 12
    Research objectives ....................................................................................................... 12

CHAPTER 2 - METHODS ....................................................................................................... 13
  2.1 STUDY AREA DESCRIPTION ......................................................................................... 13
  2.2 VEGETATION CLASSES FOR FIELD SAMPLING AND ACCURACY ASSESSMENT .......... 14
  2.3 DATA COLLECTION ...................................................................................................... 16
    2.3.1 Vector data ............................................................................................................ 16
    2.3.2 Reference field data collection .............................................................................. 16
    2.3.3 Reference data calibration .................................................................................... 19
    2.3.4 Simplification of vegetation classes ...................................................................... 20
  2.5 ACCURACY ASSESSMENT ........................................................................................... 21
    2.5.1 The error matrix .................................................................................................... 21
    2.5.2 Overall accuracy ................................................................................................... 22
    2.5.3 Producer’s and User’s Accuracy ......................................................................... 22
  2.6 KAPPA ANALYSIS ...................................................................................................... 23
    2.6.1 KHAT statistic ...................................................................................................... 23
    2.6.2 Variance ................................................................................................................. 24
    2.6.3 Z statistic .............................................................................................................. 25
LIST OF FIGURES
Figure 1: Study area map ..................................................................................................................................................13
Figure 2: The Litchfield vegetation map, entitled "Provisional Vegetation Types over Litchfield National Park" (Trueman and Cuff, 2014) .........................................................................................................................14
Figure 3: Reference data collection techniques ..................................................................................................................16
Figure 4: Example of a geo-tagged photograph of the vegetation, taken from the helicopter .........................................................17
Figure 5: Location of reference data sites: Random Sampling Dataset ........................................................................................18
Figure 6: Location of reference data sites: Stratified Sampling Dataset .......................................................................................19
Figure 7: Accuracy of Litchfield vegetation map compared with reference data - random stratified sites .................................................................27
Figure 8: Accuracy of Litchfield vegetation map compared with reference data - systematic stratified sites .................................................................28
Figure 9: Accuracy of photos and observations compared with random stratified sites .................................................................31
Figure 10: Accuracy of photos and observations compared to stratified systematic calibrated data .............................................................34
Figure 11: Accuracy of NVIS-L3 map classification compared with random stratified reference data .......................................................37
Figure 12: Accuracy of NVIS-L2 map classification compared with random stratified reference data .......................................................39
Figure 13: Accuracy of Litchfield management units- map classification compared with random stratified reference data ........................................................................................................................................................................39

LIST OF TABLES
Table 1: Satellite sensors specification ..............................................................................................................................................6
Table 2: Vegetation classes for accuracy assessment, as defined by the Litchfield vegetation map (Trueman & Cuff, 2014) ..........................................................................................................................................................15
Table 3: Simplification of vegetation classes according to the National Vegetation Information System (NVIS) levels 2 and 3, and according to management units as defined in the Litchfield National Park Integrated Conservation Strategy ..................................................................................................................20
Table 4: Error matrix comparing Litchfield vegetation map with reference data - random stratified sites ..................................................................................................................29
Table 5: Error matrix comparing Litchfield vegetation map with reference data - systematic stratified sites ..................................................................................................................30
Table 6: Error matrix comparing site photos (post-classified) with reference data (random stratified sites) ..................................................................................................................32
Table 7: Error matrix comparing site observations (classified on-site) with reference data (random stratified sites) ..................................................................................................................33
Table 8: Error matrix comparing post classified photos accuracy for stratified systematic samples ..................................................................................................................35
Table 9: Error matrix comparing observational accuracy for stratified systematic samples ..................................................................................................................36
Table 10: Error matrix comparing map classified up to NVIS_L3 with random stratified calibrated reference data classified up to NVIS_L3 ..................................................................................................................38
Table 11: Error matrix comparing map classified up to NVIS_L2 with random stratified calibrated reference data classified up to NVIS_L2

Table 12: Error matrix comparing Litchfield map classified into management units with random stratified reference data classified according to Litchfield management units.

Table 13: Comparison of Kappa analysis for the Litchfield vegetation map with reference data generated using both random stratified and stratified systematic sampling.

Table 14: Comparison of Kappa analysis for the Litchfield vegetation map with random stratified reference data at broader classification scales; classified according to Litchfield management units and NVIS hierarchy level 2.
ACRONYMS
AVHRR-Advanced High Resolution Radiometer
CSIRO-Commonwealth Scientific and Industrial Research Organization
DLRM-Department of Land Resource Management
DCBR- Darwin Centre for Bushfire Research
ETM+-Enhance Thematic Mapper Plus
GIS- Geographical Information System
GAC-Global Area Coverage
LAC- Local Area Coverage
m- Metres
MODIS-Moderate Resolution Imaging Spectro-radiometer
NDVI-Normalised Difference Vegetation Index
NT-Northern Territory
NVIS-National Vegetation Information System
TM-Thematic Mapper
UNESCO-United Nations Educational Scientific and Cultural Organization
ABSTRACT
Litchfield National Park required a vegetation habitat map to undertake informed conservation management. The Northern Territory Government Department of Land Resource Management were commissioned to derive this map in a short timeframe that did not enable validation of the results. Good vegetation mapping procedure requires validation to inform the users of the map’s inherent errors. This can be done with field reference data by statistical analysis using a standard error matrix and Kappa analysis. The user’s and producer’s accuracy derived through the error matrix enables the users of the map to assess the individual vegetation class accuracy. However, the field sampling and data recording methods, although widely used, are not fully standardised.

Therefore, this research undertook an accuracy assessment of the Litchfield vegetation map to test the appropriateness of various field sampling techniques for reference data collection. This included the comparison of stratified versus stratified-random sampling and the comparison of vegetation classification on-site versus post-field photo interpretation. Two complete field datasets of 12 different vegetation classes were collected through air survey. All site classifications were calibrated based on the field photos, observations, imagery and landscape context. Subsequently, these datasets were used to create error matrices and conduct statistical analysis that assessed the accuracy of the map and the sampling and classification strategies.

The overall accuracy of the Litchfield vegetation map was “moderate” (~60%). The vegetation types of lowland woodland, alluvial grassland, riparian and sandstone woodland were the most accurately mapped. The sampling methods comparison (Z statistic) showed that there was no significant difference in the error matrices generated from each sampling strategy, so it can be concluded that both methods were effective for collecting the reference data. When the Litchfield vegetation map was reclassified into 4 simple classes based on vegetation structure or management units, the overall accuracy of the mapping was much higher (72 – 80%), however still included misclassifications that would result in erroneous management. The field photo data recording type was more accurate (87%) compared to the on-site observation method (79%) for generating reference data for map validation.

Both reference data collection strategies, random stratified and systematic stratified, were equally suitable for providing data for the map accuracy assessment. This means that the easier, less costly approach can be used, in this case the easiest and cheapest strategy was random transect - stratified interval sampling.

Key Words: Accuracy assessment, Litchfield vegetation map, Random/stratified sampling, Error matrix and Kappa analysis.
CHAPTER 1- INTRODUCTION
Maps are an important source of information, illustrating the spatial distribution of different natural resources and manmade infrastructure (Dennis & Carte, 1998). Satellite remote-sensing platforms have enabled access to a synoptic view of the earth, providing a wide range of spectral data from various sensors and providing detailed information of the spatial extent and condition of different land cover types (Trotter, 1991) useful in regional scale natural resource conservation, planning and management. The emergence of readily accessible satellite data, combined with increases in desktop computing performance has enabled rapid processing of remotely sensed data for the production of maps illustrating the earth’s surface.

Broad scale mapping often uses satellite data as the basis of map creation (Congalton & Green, 2008). Derived spatial information (a map) is fundamentally a representation of the ‘real world’ at a scale convenient for analysis or display purposes. In terms of conservation management, it is important that these maps accurately and reliably reflect real-world features at a scale useful for management. While the methods used to derive these maps are often sophisticated, their accuracy can be limited by a number of factors including the availability of ancillary and ground data and by the spectral range of the available remotely sensed data (Herold & Clarke, 2003). It is essential that spatial information derived from remotely sensed data is verified with field data (Smits et al., 1999).

Satellite data are used to create maps of a whole range of features across the landscape. This project focuses on the assessment of map validation techniques, based on a vegetation map of Litchfield National Park prepared in 2013 by using Landsat satellite data. Therefore, this introduction will focus on two main topics within this field. Firstly, it will review the techniques used to map vegetation cover, which is relevant to the case study used in this project. Secondly, it will review techniques for map validation, including field data collection. Lastly, it will introduce the Litchfield vegetation map used as a case study and define the aims and objectives of this project.

1.1 VEGETATION MAPPING
Vegetation mapping has long been regarded as a key component of the information base required for effective natural resource planning and management throughout
the world including Australia (Neldner et al., 2014). Vegetation maps are used to locate and identify sites of conservation importance and to produce inventories for the spatial distribution of different plant communities (Nemec & Raudsepp, 2013). Modern vegetation mapping with high spatial and temporal resolution are widely used in the applied sciences including fire management, water quality monitoring and wildlife habitat characterization (Carpenter et al., 1999).

1.1.1 History of vegetation mapping
The first known vegetation maps were from the 15th century in the Cosmographia of Ptolemy, published in Italy in 1947. Known forests were indicated with patterns of trees of varying sizes (Fosberg, 1961). In the 16th century, forests were shown on maps with growing accuracy and frequency due to the application of forest maps for military purposes to indicate obstacles for communication, hunting reserves and timber resources, (Kuchler & Zonneveld, 2012). Meanwhile, the first genuine vegetation map was published by Sendtner (1854) with the geographical distribution of plant species such as ‘Spangnetum with Pinuspumilo” which illustrated the extent of vegetation types. At the end of the 19th century the first ecological approach for vegetation mapping was implemented for the new vegetation map of the world by Schimper (1898).

In the mid-20th century a new era of modern vegetation mapping evolved. The period after World War II provided important advancements in vegetation mapping with the innovation and availability of air survey data including black and- white, natural colour and false colour infrared photography. The Soviet Union, Europe and the United States all produced vegetation maps at that time (Kuchler, 1960). In 1961, Fosberg undertook European vegetation mapping based on scientific observation, aiming to provide a reference for the measurements of environmental change and ecological relationships. During this period UNESCO also published the 3 sheet map of the world with the major vegetation zones (White, 1981). In Australia, mapping was carried out by the Commonwealth Scientific and Industrial Research Organization (CSIRO) with mapping of Western Australia at 1:1,000,000 (Beard. 1974-1981 cited in Kuchler & Zonneveld, 2012).

Today, in addition to photogrammetric data, modern satellite technology makes it possible to map larger areas with increased spatial resolution, which today is even
capable of identifying individual plants (Goetz, Roack and Rowan 1983 cited in Kuchler, 1960). The number and range of stakeholders who create and use vegetation mapping are increasing with this advancement of technology. Kuchler and Zonneveld (2012) describe in their book “Vegetation mapping” that vegetation maps mainly consist of information about plant types and their areal extent within the mapped area, but only reflect the main interests of botanists and geographers. However, there are growing numbers of people who depend on accurate vegetation maps as a fundamental resource in their respective fields. This includes zoologists, geologists, pedologists, plant and animal ecologists and people involved and concerned with land use planning, agriculture and conservation, forestry and wildlife, with climatology and communication and education and military stakeholders.

1.1.2 Vegetation classification

On a vegetation map the landscape is generally divided into classes that each represents a distinct vegetation community (Kuchler, 1967). The way vegetation is classified and described depends on the nature and scope of research or management need. For example, the management aims and objectives, the required scale of the study or management, consideration of individual species or the overall habitat and the availability of different resources such as time, money, man power and ancillary data all contribute to decisions on how to map vegetation (Kent, 2011). To make a vegetation map, it is important to list out and describe the major vegetation communities that exist in the mapping area before proceeding to the final classification or vegetation survey (Xie & Yu, 2008).

There are two main types of vegetation classification and description: physiognomic or structural and; floristic.

Physiognomic or Structural: the description of vegetation is mainly based on the external morphology, stratification, size and life form of plants in the community (Raunkaier, 1934 cited in Kent, 2011). This classification is primarily used for mapping vegetation at a small scale over larger areas, such as community level formation of world vegetation (Kent, 2011). This kind of classification may include the description of individual strata and helps to determine the structure of a plant community and define the spatial distribution and pattern of growth forms within single or multiple plant communities (Gams, 1918 cited in Tansley, 1920).
Floristic: vegetation maps are often classified based on the dominant species and other species represented with the name of the dominant species such as oak forest and oak-white forest (Kuchler & Zonneveld, 2012). Usually a floristic description of vegetation has been applied at a large scale over small areas at the level of different plant communities to identify their presence or absence and to record abundance (Kent, 2011). For the mapping of vegetation based on floristics it is required to record the name of all species within the study area (Wilson et al., 1990). In Australia, national standards incorporate both a structural and floristic description of vegetation for the mapping of land cover (National Committee on Soil and Terrain, 2009).

1.1.3 Techniques and methods for vegetation mapping
There are a number of approaches that have been used to map vegetation. In general, field data are used in conjunction with some kind of aerial imagery or existing map of the landscape.

The survey approach to mapping vegetation is mainly based on ecological survey and has been widely applied internationally (Zonneveld, 1989). For example, the vegetation map of Australia at a scale of 1:6,000,000 was published by J.A. Carnahan (1973) for the Atlas of Australian resources to show the distribution of vegetation types. The vegetation was classified by undertaking a survey of various sites to differentiate the structural formation of vegetation types (Christian, 1982).

Another example is provided by Kamada and Okabe (1998) who described fine scale vegetation mapping techniques with the application of low-altitude aerial photography combined with ground vegetation survey. Photographs were taken from a vertical height of 100-300m using a black and white remote-controlled camera system mounted on a helium balloon. Some vegetation maps have been derived using quite different techniques. Source maps such as military maps, large area topographic maps and cadastral maps have been used by environmental researchers and forest scientists for the mapping and estimation of vegetation cover. For example, Carni et al (1998) described the use of cadastral and military maps for the mapping and comparison of past and present forest vegetation in NE Slovenia. Trueman et al (2013) used oral history and cadastral maps to map historical vegetation in the Galapagos Islands.
In the last two decades, satellite images and geographic information systems (GIS) have become the norm for vegetation mapping (Langley et al., 2001). Satellite imagery is good for mapping vegetation because it includes spectral reflectance in the visible and near-infrared regions of the electromagnetic spectrum, which are useful for distinguishing different types of vegetation (Xie et al., 2008). Comparative radiance in these defined spectral regions is incorporated into many vegetation indices, and is directly related to the fraction of photosynthetically active radiation (Asrar et al., 1984; Galio et al., 1985 cited in Xie et al., 2008). The level of radiance of photosynthetically or non-photosynthetically active vegetation to the sensor will provide differences based on the density, health and type of vegetation, this phenomenon is widely used by remote sensing scientists to classify vegetation (Berri et al., 2007).

1.1.4 Application of satellite imagery for vegetation mapping
For vegetation mapping, there are eight satellite sensors that have been commonly used (Table 1).

<table>
<thead>
<tr>
<th>Satellite sensors</th>
<th>Date of launch</th>
<th>Pixel size</th>
<th>Acquisition rate</th>
<th>Swath</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-1</td>
<td>1972</td>
<td>80 m</td>
<td>18 days</td>
<td>185km</td>
</tr>
<tr>
<td>Landsat 5 -(Thematic Mapper, TM)</td>
<td>1984</td>
<td>30-129m</td>
<td>16 days</td>
<td>185km</td>
</tr>
<tr>
<td>Landsat 7-(Enhance Thematic Mapper Plus – ETM+)</td>
<td>1999</td>
<td>30-60m</td>
<td>16 days</td>
<td>183km</td>
</tr>
<tr>
<td>MODIS (Moderate Resolution Imaging Spectro-radiometer )</td>
<td>1999</td>
<td>250m-1000m</td>
<td>24 hours/tasking plan</td>
<td>2330km</td>
</tr>
<tr>
<td>SPOT-1 to 7</td>
<td>2012</td>
<td>1.5m to 6.0m</td>
<td>26 days</td>
<td>60km</td>
</tr>
<tr>
<td>NOAA- (Advanced Very High Resolution Radiometer -A VHRR)</td>
<td>1978</td>
<td>1.1km</td>
<td>14 Each day</td>
<td>2399km</td>
</tr>
<tr>
<td>IKCONOS</td>
<td>1999</td>
<td>0.82m to 3.2m</td>
<td>3 days</td>
<td>11.3km</td>
</tr>
<tr>
<td>QuickBird</td>
<td>2001</td>
<td>0.65m</td>
<td>1-3.5 days</td>
<td>16.8km</td>
</tr>
</tbody>
</table>

Source: Satellite Imaging Corporation, 2016

LANDSAT data have been available to the scientific community since the first Landsat satellite was launched in 1972. The data are widely used for the mapping of land cover. A more refined LANDSAT data product became available with the addition of LANDSAT Thematic Mapper (TM) launched in 1982 and the Enhanced Thematic Mapper (ETM+) imaging sensors launched in 1999. NASA has archived millions of Landsat satellite images, providing a continuous record of observation of


Due to its coarse spatial resolution, Moderate Resolution Imaging Spectroradiometer (MODIS) data is not recommended for the mapping of local or regional scale vegetation cover (Xie & Yu, 2008). But an important example was presented by Knight et al (2006) to classify vegetation for the mapping of land cover of Albemarle-Pamlico estuarine by using MODIS-NDVI 250 m, 16 day composite data. Gowards et al (2003) suggested the use of IKONOS data for high accuracy vegetation map preparation. Coops et al (2006) described the application of QuickBird multispectral satellite data for the mapping and monitoring of small changes in vegetation cover, and highly recommended its application for the validation of vegetation mapping.

**1.2 MAP ACCURACY ASSESSMENT**

As discussed earlier, it is essential that maps derived from remotely sensed data are verified with field data (Smits et al., 1999). The verification process is called a map accuracy assessment and it evolved after the launch of Landsat-1 in 1972. Hord and Brooner (1976) and Genevan (1979) raised questions about map accuracy and proposed initial criteria and techniques for testing the accuracy of digitally derived maps (Congalton and Green, 2008). Further studies in the field of map accuracy assessment were carried out by Rosenfield et al., (1982), and Congalton et al., (1983). Modern methods of map accuracy assessment are now widely accepted by the research community in the field (Congalton and Green, 2008).

Map accuracy is typically reported as the proportion of the map dataset which agrees with the reality of what is on the ground (Smiths et al, 1999). Accuracy assessment
and map validation are an integral part of any mapping project to understand the extent to which maps of vegetation and other land cover classes reflect reality (Congalton, 1994 cited in Congalton and Green, 2008). Congalton (1999) outlined the following reasons for the need of a map accuracy assessment:

A. Curiosity and desire to know how good something is to achieve maximum reliability.
B. To achieve the ability of quantitative comparison of different mapping techniques.
C. The ability to use the spatial analysed data in the decision making process.

Further reasons for undertaking a map accuracy assessment are to provide an overall measure of map quality, to form a basis of evaluation for different classification methods or algorithms and to gain understating of the errors for better utilization of the maps for further management, applied research and planning projects (Foody, 2002).

1.2.1 Mapping validation techniques
Congalton (2001) identified major map validation techniques as outlined below.

Reference data
Sample design and reference data collection are an essential part of a map accuracy assessment (Congalton, 2001). ‘Reference data’ refers to the data representing reality as much as possible, (i.e. samples of reality). Reference data are the basis for calculating the accuracy of a map. Reference data are usually in the form of field observations at points in the landscape. The distribution of these points across the landscape is referred to as the sample design. Samples of reference data must be designed and collected carefully according to the aim of the analysis (Hay, 1979 cited in Foody, 2002). It is important to consider how information is distributed in a map, what can be the appropriate sampling unit, how the samples should be chosen and how many samples are required for assessment (Congalton and Green, 2008). Atkinson (1991) suggested that the sample design should be based on geo-statistics to produce the best accuracy assessment.

There are five common techniques widely used for collection of reference data: systematic sampling, simple random sampling, stratified random sampling, stratified
systematic unaligned sampling and cluster sampling (Congalton and Green 2008). The selection of an appropriate sampling method depends on the aim of the research and researcher’s interest (Foody, 2002). Simple random sampling is suitable for a large sample size to ensure that all land cover classes are represented adequately (Stehman and Czaplewski, 1998). The Systematic stratified sampling method is useful to select samples from certain classes within the study area (Congalton and Green, 2008).

**Visual inspection**
The first step of a map accuracy assessment is visual inspection of a classified map to generate the basic idea about how well it represents the reality of the ground, followed by re-classification and adjustment of the map accordingly. The major advantage of visual inspection of the map is to save time and effort in the overall accuracy assessment of the map. For example if a map does not seem accurate in the first visual assessment then it doesn’t make sense to assess it further. However, it is not sufficient to conclude the accuracy of a map with a visual inspection, many maps which look good, may have some serious errors and require further accuracy assessment (Congalton, 2001). For example, vegetation may be correctly labelled on a map, but for ecological management purposes it may require further accuracy assessments that perfectly show open woodland is separated from open forest.

**Non-site specific analysis**
Non site specific analysis of a particular map involves the comparison of the overall amount of accuracy for different areas without any regard to the locational component. For example, the total forest area in a particular state classified by two different analysts can be assessed based on the amount of total area present in a state forest statistic. Therefore, the classification which represents the same amount of area under forest can be assumed to be the more accurate map. Generally, these types of assessments are suitable for planning purposes and the guidance for management of natural forest resources, or prediction of agricultural crop production. The maps can be compared with the help of reference data, generally assumed to be correct (Congalton and Green 2008).

**Difference image creation**
This is the first error quantifying technique using GIS (Foody, 2002). The first step is to compare two maps of the same area and generate a difference map or image that
highlights where the two maps do not agree. This method requires that one of the two maps is accurate, so that the generated difference image shows the pattern of major errors in the second map (Congalton, 1999 cited in Foody, 2002). This method is generally suitable for management purposes and to detect the change in landscapes variables over time (Foody, 2002).

**Error budgeting**

Error budgeting is mainly used to determine the various components (users errors, used satellite data errors and image processing or classification errors) contributing most of the errors and which can be easily corrected by dividing into partitions and divisions (Congalton, 1999 cited in Foody, 2002). Error budgeting techniques have been widely implemented by different researchers in the last 20 years. The main methods for error budgeting are the listing of various components behind the error in a map individually and calculation of their contribution to the overall map error (Congalton, 2001). The main advantage of this method is it brings an opportunity to the analyst to know the main reasons behind the map errors and therefore helps to rectify these errors by targeting them individually to produce an accurate map (Congalton, 2001).

**Quantitative accuracy assessment**

Quantitative accuracy assessment includes a numerical measure of how much the map accurately represents the reality on the ground. This can be assessed through the process of creating an error matrix. It consists of a square array of numbers generally organised in rows and columns to express the number of sampled units belonging to each class compared to the intersecting mapped class at each of those sampled units (Congalton, 2001). In an error matrix, usually columns represent the reference data and rows indicate the map classes. An underlying assumption is that reference data are correct. Error matrices are widely used to calculate the error of inclusion such as commission error wrongly added area in a particular class from other classes and omission error; area excluded from the category in which it actually belongs (Congalton, 1999 cited in Foody, 2002). A recent example of calculating the omission and commission error using an error matrix was given by Edwards et al (2015) for validating fire severity maps in the tropical savannas of northern Australia. By comparing the 3148 waypoints from aerial transects with the
fire severity mapping and suggested the reliable classification for not severe fire and error of commission in the severe fire category.

The error matrix is very useful because it can be used as a platform for a series of analytical and descriptive statistics. For example, we can compute overall accuracy, producer’s accuracy (percentage of sites classified correctly through reference data) and user’s accuracy (proportion of sites in map accurately classified) by using the error matrix and by dividing various computed diagonals with reference data (Story and Congalton 1986 cited in Congalton, 2001). Overall map accuracy is measured using a range of analyses including Kappa, Margfit and conditional Kappa, all derived from an error matrix, as detailed further in the Methods section of this thesis.

Today, the error matrix is considered the standard accuracy assessment tool for map validation (Congalton 2001).

1.3 CASE STUDY: LITCHFIELD NATIONAL PARK VEGETATION MAP

Litchfield National Park represents a broad range of Top End flora and fauna and is widely regarded as a key reserve for the preservation of species and ecosystems unique to the tropical savannas of the Northern Territory. There are 974 native plant species and 338 native vertebrate species including 192 birds, 24 amphibians, 76 reptiles and 46 mammals (PWCNT, 2015). This high flora and fauna species diversity in the Park is considered to be due to the availability of various types of habitats including undisturbed remnant rainforest, sandstone plateaus, upland swamps and plains with black soil (DEWHA, 2008).

Annual dry season fire is considered the major threat for the long term biodiversity of flora and fauna in the park (PWCNT, 2015). Specifically, 56 percent of the area of Litchfield Park has burnt annually in the last 8 years (Edwards et al., 2001). Therefore, the Parks and Wildlife Commission of the Northern Territory have developed a strategic management plan for the conservation and monitoring of local vegetation biodiversity in the Park (PWCNT, 2015).

As part of developing the plan, it was considered necessary to understand the distribution of vegetation types in the Park. Thus, the “Provisional Vegetation map over Litchfield National Park” was created specifically to provide information for the integrated conservation planning process and fire management (Trueman & Cuff,
The map was created in a short time-frame using historical ground data and readily available satellite imagery, but due to time restrictions, it was not validated using field data (Trueman & Cuff, 2014). As such, this map dataset provides an opportunity to test map validation techniques in relation to field data collection and analysis.

1.4 AIMS AND OBJECTIVES

Research aim
The main aim of this research project was to compare various sampling techniques for map validation applied to the Litchfield vegetation map. This included comparison of:

1. Field data collection sampling methodologies.
2. Field data recording type (on-site observation versus photographs).
3. Statistical validation methods for assessing the mapping against the field data (reference data).

Research objectives
The study focuses on four objectives:

- To use two field data sampling methods (stratified or stratified-random) and discuss the pros and cons of each.
- To assess the accuracy of post-classified field data (classified into vegetation type from photos) and the accuracy of directly observed field data (classified on-site at time of data collection) compared to a calibrated classification of field data.
- To assess the map accuracy at different classification scales.
- To apply Kappa analysis to test the appropriateness of sampling methodologies for map validation and comment on it for different applications.
CHAPTER 2- METHODS

2.1 STUDY AREA DESCRIPTION

The study area is located west of Batchelor and Adelaide River in the Northern Territory, Australia (Figure 1). The boundary of the study area is defined by the extent of the vegetation map with an area of 1,905km$^2$; it includes all of Litchfield National Park and a 2 km buffer (Trueman & Cuff, 2014).

Figure 1: Study area map
2.2 VEGETATION CLASSES FOR FIELD SAMPLING AND ACCURACY ASSESSMENT

The Litchfield vegetation map (Figure 2), was derived from Landsat 8 satellite imagery. Images were acquired in 2013 and mapping was completed in 2014 (Trueman & Cuff, 2014). Before field data collection, the classes were coded with acronyms to simplify recording during ground survey (Table 2). Further, vegetation classes were simplified (i.e. combined) to enable accuracy assessment of the map at broader classification scales. This is further explained in section 2.4.

Source: Department of Land Resource Management (DLRM) NT.

Figure 2: The Litchfield vegetation map, entitled “Provisional Vegetation Types over Litchfield National Park” (Trueman and Cuff, 2014)
Table 2: Vegetation classes for accuracy assessment, as defined by the Litchfield vegetation map (Trueman & Cuff, 2014).

<table>
<thead>
<tr>
<th>ORDER</th>
<th>VEG_TYPE</th>
<th>VEG_DESC</th>
<th>CHARACTERISTIC_SPECIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tabletop open forest</td>
<td>Open forest of mixed Eucalypt species on the Tabletop Range</td>
<td><em>Eucalyptus tetrodonta, E. miniata</em> and <em>Erythrophleum chlorostachys</em></td>
</tr>
<tr>
<td>2</td>
<td>Tabletop woodland</td>
<td>Woodland to open woodland of mixed Eucalypt species on the Tabletop Range</td>
<td><em>Eucalyptus tetrodonta, E. miniata</em> and <em>Erythrophleum chlorostachys</em></td>
</tr>
<tr>
<td>3</td>
<td>Upland swamp community</td>
<td>Shrubland of mixed species, typically <em>Lophostemon lactifluus</em>, in poorly drained areas of the Tabletop Range</td>
<td><em>Lophostemon lactifluus</em> and/or <em>Melaleuca viridiflora</em>/M. cajaputi.</td>
</tr>
<tr>
<td>4</td>
<td>Upland sedgeland</td>
<td>Closed sedgeland or grassland on the Tabletop Range</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Wet heath</td>
<td>Shrubland of mixed species, typically <em>Melaleuca nervosa, Pandanus spiralis, Grevillea pteridifolia</em>, on a sandstone substrate</td>
<td><em>Melaleuca nervosa, Pandanus spiralis, Grevillea pteridifolia, Sorghum plumosum, Osbeckia australiana.</em></td>
</tr>
<tr>
<td>6</td>
<td>Dry heath</td>
<td>Shrubland of mixed species, typically including <em>Calitrix exstipulata</em></td>
<td><em>Calitrix exstipulata, Grevillea spp.</em> (not <em>pteridifolia</em>), <em>Acacia spp.</em> (<em>e.g.</em> <em>tolmerensis</em>), <em>Jacksonia dilatata,</em> <em>Gardenia megasperma,</em> <em>Persoonia falcata,</em> <em>Owenia vernicosa,</em> <em>Verticordia spp.</em></td>
</tr>
<tr>
<td>7</td>
<td>Sandstone woodland</td>
<td>Woodland to open woodland of mixed Eucalypt species on a sandstone substrate</td>
<td><em>Eucalyptus tetrodonta, E. miniata, E. phoenicea,</em> <em>Terminalia latipes,</em> and <em>Erythrophleum chlorostachys</em></td>
</tr>
<tr>
<td>8</td>
<td>Riparian</td>
<td>Closed forest of mixed tree species, typically along steep drainage channels</td>
<td>This includes several types of forest including riparian forest, monsoon rainforest and spring communities</td>
</tr>
<tr>
<td>9</td>
<td>Lowland woodland</td>
<td>Woodland to open woodland of mixed tree species, including eucalypts</td>
<td><em>Eucalyptus miniata, E. tetrodonta, E. tectifica, Erythrophleum chlorostachys,</em> <em>Xanthostemon paradoxus,</em> <em>Terminalia ferdinandiana,</em> <em>T. grandiflora.</em> Also <em>Buchanania obovata,</em> <em>Livistona humilis,</em> <em>Grevillea pteridifolia,</em> <em>Pandanus spiralis</em> and <em>Acacia spp.</em></td>
</tr>
<tr>
<td>10</td>
<td>Melaleuca woodland</td>
<td>Woodland of mixed species on plains, typically including <em>Melaleuca viridiflora</em> or <em>M. leucadendra</em></td>
<td><em>Melaleuca viridiflora,</em> <em>M. leucadendra,</em> <em>Syzygium eucalyptoides,</em> <em>Pandanus spiralis,</em> and <em>Lophostemon lactifluus</em> may also be present.</td>
</tr>
<tr>
<td>11</td>
<td>Drainage woodland</td>
<td>Woodland of mixed species, typically <em>Lophostemon lactifluus</em>, in drainage areas</td>
<td><em>Lophostemon lactifluus</em> (dominant) and <em>Melaleuca cajaputi</em> and possibly <em>Corymbia bella,</em> <em>C. grandiflolia,</em> <em>Eucalyptus alba,</em> <em>E. bigalerita,</em> and the grasses/sedges <em>Eriachne triseta,</em> <em>Sorghum plumosum,</em> <em>Ectrosia sp.</em></td>
</tr>
<tr>
<td>12</td>
<td>Alluvial grassland</td>
<td>Closed grassland of mixed species on alluvial plains</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Rock</td>
<td>Sparsely vegetated</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Water</td>
<td>Inundated and not vegetated</td>
<td></td>
</tr>
</tbody>
</table>
2.3 DATA COLLECTION

2.3.1 Vector data
Prior to field data collection, ancillary vector data were compiled. Litchfield vegetation and road network datasets were collected from Department of Land Resource Management (DLRM), Northern Territory Government. For the preparation of the study area map, a vector shapefile was obtained from the Darwin Centre for Bushfire Research (DCBR), Charles Darwin University.

2.3.2 Reference field data collection
Field sampling of vegetation data was conducted from a helicopter using a combination of techniques from the 17th to 22nd March 2016 (Figure 3). Two complete datasets were collected through the air survey; the first using random location of sites, and the second using Stratified systematic location of sites; further outlined below. At each site the following data were collected:

- Geographical site location (latitude/longitude);
- Observed vegetation class (Table 2);
- A rating of the operator-confidence in the assigned vegetation class (high, medium or low) see Appendix 1,
- A geo-tagged photograph of the vegetation (Figure 4). These photographs were later viewed and assigned a vegetation class by an operator who was not the same person who collected the paired observation data.

Figure 3: Reference data collection techniques
Random (Interval) sampling: Site locations were determined by the location of the helicopter at 30-second intervals on any flight-path. Flight-paths were determined for other purposes, so the location of sites was considered random in relation to vegetation types. Observation data were collected using a cyber tracker sequence. Validation data were recorded at 322 sites (Figure 5). The geo-location of each site was determined from the photograph, or–where there was no valid geo-tagged photograph–from Cyber tracker.
**Stratified systematic sampling:** Site locations were pre-determined to cover all of the vegetation types in the Litchfield vegetation map as evenly as possible. This technique is highly recommended by Congalton and Green (2008) for the accuracy assessment of maps generated by remotely sensed images in the field of vegetation mapping. Observation data were recorded on paper. Validation data were collected at 105 sites (Figure 6). The geo-location of each site was determined from the photograph. These locations were within operational
distance (generally within 100m) from the pre-determined site locations, but not all coincided with the intended vegetation type.

Figure 6: Location of reference data sites: Stratified Sampling Dataset

2.3.3 Reference data calibration

All the data collected during the field survey were arranged in digital files by using a range of software including; Microsoft Excel, ArcGIS 10.2.2, and DNRGPS. These arranged files were handed to the two different experts from the NTG working in the relevant field for the post field classification. These experts assigned a calibrated vegetation class to each site in each dataset by assessing the visual appearance of vegetation in the photograph, the observation recorded at the site (including the associated confidence level recorded during the survey), the visual appearance of the satellite imagery at the site, and the surrounding landscape context. These
calibrated datasets (both Random and Stratified) were then used as the reference data for the accuracy assessment of the Litchfield vegetation map. They were also used for the accuracy assessment of each of the two data recording types: photographs and observation.

2.3.4 Simplification of vegetation classes

In order to test the appropriate vegetation map classification scale, the Litchfield vegetation map classes and field reference data were combined to approximately match the NVIS classification hierarchy levels 2 & 3 (Table 3) (Brocklehurst et al., 2007). Furthermore, for ecological conservation and biodiversity management, all classes were combined into management units according to the values defined in the Integrated Conservation Strategy for Litchfield National Park (PWCNT, 2015).

Table 3: Simplification of vegetation classes according to the National Vegetation Information System (NVIS) levels 2 and 3, and according to management units as defined in the Litchfield National Park Integrated Conservation Strategy

<table>
<thead>
<tr>
<th>Map class code from Table 2 (equivalent to NVIS L4/5)</th>
<th>NVIS L3 code*</th>
<th>NVIS L2 code**</th>
<th>Litchfield Management unit code***</th>
<th>Management unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>a</td>
<td>a</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DW</td>
<td>b</td>
<td>b</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>EW</td>
<td>d</td>
<td>b</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>EWT</td>
<td>d</td>
<td>b</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>GL</td>
<td>e</td>
<td>c</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>MW</td>
<td>c</td>
<td>b</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>OF</td>
<td>f</td>
<td>d</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>RIP</td>
<td>g</td>
<td>d</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>SL</td>
<td>e</td>
<td>c</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>b</td>
<td>a</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>SW</td>
<td>d</td>
<td>b</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>WH</td>
<td>a</td>
<td>a</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

* a is ‘Grevillea Shrubland’, b is ‘Lophostemon woodland’, c is ‘melaleuca woodland’, d is ‘eucalyptus woodland’, e is ‘sorghum grassland/Cyprus sedgeland’, f is ‘eucalyptus open forest’, and g is ‘carallia open forest’.

** a is ‘Shrubland’, b is ‘woodland’, C is ‘closed grassland/closed sedgeland’, and d is ‘open forest.’
2 is ‘sandstone Plateaus’, 3 is ‘Monsoon Rainforest and Swamps’, 4 is ‘Melaleuca Woodlands’ and 5 is ‘Lowland and Alluvial Plains’.

2.5 ACCURACY ASSESSMENT

The two calibrated datasets were each treated as reference data for assessing the accuracy of the Litchfield vegetation map using statistical error analysis as outlined below. This process was repeated using the simplified vegetation classifications, using only the random sampling reference dataset. Additionally, within each dataset, the photograph data and the observation data were each compared with the calibrated data as a measure of the accuracy of the type of data recording (not including the kappa analysis).

2.5.1 The error matrix

An error matrix was generated by intersecting the reference points with the mapping in GIS to obtain the mapped vegetation class at each reference data site, the data were exported to Microsoft Excel. A pivot table was generated to display the reference data in columns compared with the base data (map data, photos or observations) in the rows, with the cells containing the total number of sites in each vegetation class. This is a standard method for assessing map accuracy because it shows how many sites in each vegetation class have been mapped correctly and incorrectly. It also allows for the calculation of overall accuracy and user’s and producer’s accuracy as explained below (Congalton and Green 1993).

The column and row totals were calculated according to the following formulas:

\[ n_{i+} = \sum_{j=1}^{k} n_{ij} \]

Where, \( n_{i+} \) represents the grand total of the reference data,

\( n_{ij} \) is the number of samples in the \( i^{th} \) row and \( j^{th} \) column in the matrix which was categorized into \( k^2 \) cells.

and,

\[ n_{+j} = \sum_{i=1}^{k} n_{ij} \]
Where, $n_{+j}$ represents the grand total of base data (under accuracy assessment)

\[ n_{ij} \] is the number of samples classified in the $i^{th}$ row and $j^{th}$ column in the matrix which was categorized into $k^2$ cells.

### 2.5.2 Overall accuracy

Overall accuracy is the sum of the diagonal values representing correctly classified sites based on the comparison of the base map and the reference data, divided by the total number of sites included in the entire error matrix (Congalton and Green, 2008). The following equation was used for the calculation of overall accuracy in the error matrix analysis:

\[
\text{Overall accuracy} \text{ (\%)} = 100 \times \frac{\sum_{i=1}^{k} n_{ii}}{n}
\]

Where, $n_{ii}$ is the number of samples classified in the $i^{th}$ row and $i^{th}$ column in the matrix categorized into $k^2$ cells. These are the values in the diagonals of the matrix.

$n$ is the total number of samples in the matrix.

### 2.5.3 Producer’s and User’s Accuracy

User’s and producer’s accuracy represent the accuracy of individual categories, which allow for a more detailed assessment of the inaccuracies of individual classes.

Producer’s accuracy reports the percentage of reference sites in a particular class that have been correctly classified. User’s accuracy estimates the proportion of areas in the map that actually belong to a specific class in the reference data.

The following equations were used to calculate the producer’s and user’s accuracy in the error matrix analysis:

\[
\text{Producer’s accuracy} \text{ (\%)} = 100 \times \frac{n_{jj}}{n_{+j}}
\]

\[
\text{User’s accuracy} \text{ (\%)} = 100 \times \frac{n_{ii}}{n_{i+}}
\]
Where, \( n_{ii} \) is the number of samples classified in the \( i^{th} \) row and \( i^{th} \) column in the matrix categorized into \( k^2 \) cells.

\[
n_{i+} \text{ is the grand total of the number of samples in the rows}
\]

and \( n_{j+} \) is the grand total of the number of samples in the columns.

Users and producers accuracy are also referred to as errors of commission and omission, respectively (Congalton and Green, 2008).

2.6 KAPPA ANALYSIS

Kappa analysis is used to statistically assess the differences between error matrices and is a standard technique for map accuracy assessment (Congalton and Green, 1999). In studies related to the map accuracy assessment, Kappa analysis has been used in a number of studies to determine the difference of one error matrix to the other, for example, comparison of two different mapping methods, comparison of classification accuracy of two different analysts and change detection (Bishop et al., 1975 cited in Congalton and Green, 2008). It is also commonly used to test the overall accuracy of a map, by comparing a mapping classification to a random classification (Congalton and Green, 1999).

The following Kappa analysis statistics were computed in this study:

2.6.1 KHAT statistic

The KHAT statistic is the Kappa measure which reflects the agreement of data under assessment with the reference data (Cohen, 1960). Results range from +1 to -1. Congalton (1991) described the grouping of KHAT values: if results are greater than 0.80, it represents strong agreement, between 0.40 and 0.80 represents a moderate level of accuracy agreement and results below 0.40 shows poor agreement.

It is computed according to the following formula:

\[
\hat{K} = \frac{n \sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} n_{i+} n_{+j}}{n^2 - \sum_{i=1}^{k} n_{i+} n_{+j}}
\]

In this project the KHAT value was used as a measure of how well the classified map data agreed to the reference data in the error matrix. This was applied to the error matrices derived from both the random (Interval) sampling and stratified systematic
sampled reference data. It was also applied to the map and reference data that had been simplified into the broader classification scales of NVIS level 2 and Litchfield management units.

2.6.2 Variance

The variance around the KHAT values is another important measure in Kappa analysis that expresses the sample variance of Kappa (Congalton and Green, 1999). The KHAT statistic is normally distributed. The calculation of the variance allows for testing the significance of the KHAT statistic for single or multiple error matrices. Variance values of agreement between the reference data and remotely sensed classification should be greater than 0 to represent the satisfactory level of classification accuracy in comparison to a random classification (Congalton and Green, 1999).

The sample variance of Kappa was computed using the Delta method:

$$\text{var}(\xi) = \frac{1}{n} \left\{ \theta_1 (1 - \theta_1) + \frac{2(1 - \theta_1)(2\theta_1\theta_2 - \theta_3)}{(1 - \theta_2)^3} + \frac{(1 - \theta_1)^2(\theta_4 - 4\theta_2^2)}{(1 - \theta_2)^4} \right\}$$

Where,

$$\theta_1 = \frac{1}{n} \left\{ \sum_{i=1}^{k} n_{ij} \right\}$$

$$\theta_2 = \frac{1}{n^2} \left\{ \sum_{i=1}^{k} n_{i+} n_{+i} \right\}$$

$$\theta_3 = \frac{1}{n^2} \left\{ \sum_{i=1}^{k} n_{ii} (n_{i+} + n_{+i}) \right\}$$

$$\theta_4 = \frac{1}{n^3} \left\{ \sum_{i=1}^{k} \sum_{j=1}^{k} n_{ij} (n_{j+} + n_{+i})^2 \right\}$$

$n_{ij}, n_{ii}, n_{j+}, n_{i+}, n_{+i}$, and $n_{+i}$ were determined previously.

This was computed to accompany all KHAT values.
2.6.3 Z statistic

The Z statistic is used to test the statistical significance of a single error matrix (difference from random) or to test if two error matrices are significantly different (Congalton and Green, 1999).

The Z statistic for testing the significance of the single error matrices was obtained using the equation below. This was computed to accompany all KHAT values.

\[
Z = \frac{\hat{K}}{\sqrt{\text{var}(\hat{K})}}
\]

where, \(\text{var}(\hat{K})\) is the variance of the error matrix and \(\hat{K}\) is the KHAT value.

The Z statistic for testing the difference between two error matrices was obtained using the equation below. This was computed to compare the level of suitability of the two different field data collection methods, random (Interval) sampling and stratified systematic sampling.

\[
Z = \frac{\hat{K}_1 - \hat{K}_2}{\sqrt{\text{var}(\hat{K}_1) + \text{var}(\hat{K}_2)}}
\]

Using a p value of 0.05 for significance, the corresponding critical value of the Z statistic is 1.96. Therefore, if the Z statistic is greater than 1.96, the result is significant. In the case of a single error matrix, it is significantly different to random. In the case of comparing two error matrices; they would be significantly different to one another, suggesting that the different data sampling methods lead to different validation results.
CHAPTER 3- RESULTS

This chapter illustrates the findings of the Litchfield vegetation map validation and the comparison of methods used to generate the reference data. The findings are illustrated as follows:

- Map validation:
  - Random stratified sampling.
  - Stratified system sampling.
- Accuracies achieved based on the data recording methods (photos post) and (direct observation) classification.
- Accuracies for management, NVIS-L3, L2 and Litchfield management units.
- Accuracies achieved through kappa analysis.

3.1 RANDOM STRATIFIED SAMPLING

The overall accuracy of the Litchfield vegetation map, assessed against reference data obtained through random stratified was 57%. Analysis showed that the user’s accuracy ranged between 8% and 100% in different vegetation classes and producer’s accuracies were calculated to be in the range between 3% and 100% (Table 4, Figure 7). The highest user’s and producer’s accuracies (100%) were observed for vegetation type upland swamp community, which had the lowest number of reference sites. Results indicated that 55% of dry heath (DH) sites, 74% of drainage woodland (DW) sites and 7% of lowland woodland (EW) sites were wrongly attributed to sandstone woodland (SW), lowland woodland (EW) and melaleuca woodland (MW) respectively. Likewise, 50% and 14% of tabletop woodland (EWT) sites were wrongly mapped as tabletop open forest (OF) and sandstone woodland (SW) respectively.

Reference data illustrated that 83% of sites mapped as dry heath (DH), 67% of sites mapped as drainage woodland and 4% of sites mapped as lowland woodland (EW) were wrongly committed to these categories from sandstone woodland (SW), alluvial grassland (GL) and melaleuca woodland (MW) respectively (Table 4).
Figure 7: Accuracy of Litchfield vegetation map compared with reference data - random stratified sites

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, EWT is ‘tabletop woodland’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’

Numbers above Producer’s accuracy bars are the number of sites of each vegetation type in the reference data. Numbers above User’s accuracy bars are the number of sites mapped as each vegetation type.

### 3.2 STRATIFIED SYSTEMATIC SAMPLING

Results of the comparison of reference data obtained through stratified systematic sampling with the Litchfield vegetation map showed a 58% overall map accuracy. Analysis showed that the user’s accuracy ranged between 9% and 100% in different vegetation classes whereas producer’s accuracy reported between 25% and 100% (Table 5).

Only 17% of sites mapped as dry heath (DH) matched the reference data; 67% of the sites were actually found to be sandstone woodland (SW). Similarly, only 25% of tabletop woodland (EWT) sites were accurately classified on the map and 63% of sites were omitted and placed in tabletop open forest (OF) (Table 5). In the case of alluvial grassland (GL), 44% of sites were wrongly classified in the mapping, out of which 33% of sites were committed into drainage woodland (DW). 17% of sandstone woodland sites were wrongly mapped as dry heath (DH). The accuracy was lowest
for drainage woodland (DW); 63% of the reference data sites in this vegetation type were mapped as lowland woodland, and all of the reference data sites that coincided with areas mapped as this vegetation type were actually of alluvial grassland (GL).

**Figure 8**: Accuracy of Litchfield vegetation map compared with reference data - systematic stratified sites.

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, EWT is ‘tabletop woodland’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’

Numbers above Producer’s accuracy bars are the number of sites of each vegetation type in the reference data. Numbers above User’s accuracy bars are the number of sites mapped as each vegetation type.
Table 4: Error matrix comparing Litchfield vegetation map with reference data - random stratified sites

<table>
<thead>
<tr>
<th>Reference Data</th>
<th>Litchfield Map</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>EWT</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>RIP</th>
<th>SL</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User's Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>DW</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>EW</td>
<td>1</td>
<td>25</td>
<td>85</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>72</td>
</tr>
<tr>
<td>EWT</td>
<td>1</td>
<td>21</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
<td>91</td>
</tr>
<tr>
<td>GL</td>
<td>2</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>MW</td>
<td>2</td>
<td>7</td>
<td></td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>OF</td>
<td>1</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32</td>
<td>27</td>
</tr>
<tr>
<td>RIP</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>SL</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>SS</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>SW</td>
<td>5</td>
<td>1</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td>61</td>
</tr>
<tr>
<td>WH</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Grand Total</td>
<td>9</td>
<td>34</td>
<td>97</td>
<td>64</td>
<td>13</td>
<td>16</td>
<td></td>
<td></td>
<td>3</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td>322</td>
<td></td>
</tr>
<tr>
<td><strong>Producer's Accuracy (%)</strong></td>
<td><strong>33</strong></td>
<td><strong>3</strong></td>
<td><strong>88</strong></td>
<td><strong>33</strong></td>
<td><strong>75</strong></td>
<td><strong>17</strong></td>
<td><strong>92</strong></td>
<td><strong>63</strong></td>
<td><strong>0</strong></td>
<td><strong>100</strong></td>
<td><strong>65</strong></td>
<td><strong>33</strong></td>
<td><strong>Overall Accuracy (184/322*100)=57%</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, EWT is ‘tabletop woodland’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’
Table 5: Error matrix comparing Litchfield vegetation map with reference data - systematic stratified sites

<table>
<thead>
<tr>
<th>Reference Data</th>
<th>Litchfield Map</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>EWT</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>RIP</th>
<th>SL</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DW</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EW</td>
<td></td>
<td></td>
<td>20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EWT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MW</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OF</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SW</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>4</td>
<td>8</td>
<td>24</td>
<td>16</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>23</td>
<td>2</td>
<td>103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy (%)</td>
<td>25</td>
<td>0</td>
<td>83</td>
<td>25</td>
<td>56</td>
<td>0</td>
<td>100</td>
<td>67</td>
<td>100</td>
<td>50</td>
<td>74</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, EWT is ‘tabletop woodland’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’
3.3 PHOTOS AND OBSERVATIONS CLASSIFICATION ACCURACY

3.3.1 Random stratified samples photos and observations accuracy

The classification method using photos (post-field classification) was found to be overall 87% accurate compared with the calibrated reference data collected by the random stratified sampling strategy (Table 6). This was higher than the accuracy of the observation method (on-site classification), at 79% for the same sampling strategy (Table 7). Producer’s accuracies for photos recording type ranged from 30% to 100% whereas producer’s accuracies for observations recording type were ranging between 33% and 100% (Figure 9).

The largest error in the classification of photos was that 67% of the open forest (OF) reference data sites had been wrongly classified as eucalypt woodland (EW). The highest error in the on-site observations was that 31% of the riparian (RIP) reference data sites were classified as melaleuca woodland (MW).

Figure 9: Accuracy of photos and observations compared with random stratified sites.

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’

Numbers above Producer’s accuracy bars are the number of sites of each vegetation type in the reference data.
GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’.

Table 6: Error matrix comparing site photos (post-classified) with reference data (random stratified sites)

<table>
<thead>
<tr>
<th>Photos Classified classes</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>RIP</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>80</td>
</tr>
<tr>
<td>DW</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>18</td>
<td>94</td>
</tr>
<tr>
<td>EW</td>
<td>1</td>
<td>3</td>
<td>119</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>135</td>
<td>88</td>
</tr>
<tr>
<td>GL</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>MW</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>OF</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>RIP</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>85</td>
</tr>
<tr>
<td>SS</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>SW</td>
<td>1</td>
<td>9</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>46</td>
<td>80</td>
</tr>
<tr>
<td>WH</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Grand Total</td>
<td>6</td>
<td>20</td>
<td>130</td>
<td>4</td>
<td>4</td>
<td>10</td>
<td>11</td>
<td>3</td>
<td>46</td>
<td>4</td>
<td>238</td>
<td>80</td>
</tr>
</tbody>
</table>

Producer’s Accuracy (%) | 67 | 85 | 92 | 100| 75 | 30 | 100 | 100| 89 | 75 | Overall Accuracy (208/238)*100= 87%
Table 7: Error matrix comparing site observations (classified on-site) with reference data (random stratified sites)

<table>
<thead>
<tr>
<th>Reference data</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>RIP</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User's Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>DW</td>
<td></td>
<td>31</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>EW</td>
<td></td>
<td></td>
<td>106</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>120</td>
</tr>
<tr>
<td>GL</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>MW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>OF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>RIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>SS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>SW</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>52</td>
<td></td>
<td>69</td>
</tr>
<tr>
<td>WH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Grand Total</td>
<td>10</td>
<td>34</td>
<td>147</td>
<td>7</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td>4</td>
<td>61</td>
<td>6</td>
<td></td>
<td>301</td>
</tr>
<tr>
<td>Producer's Accuracy (%)</td>
<td>90</td>
<td>91</td>
<td>72</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>62</td>
<td>75</td>
<td>85</td>
<td>33</td>
<td></td>
<td>Overall Accuracy 237/301*100= 79%</td>
</tr>
</tbody>
</table>

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’.
3.3.2 Stratified systematic samples photos and observations accuracy

The classification method using photos (post-field classification) was found to be overall 90% accurate compared with the calibrated reference data collected by the stratified systematic sampling strategy (Table 8). This was higher than the accuracy of the observation method (on-site classification), at 75% for the same sampling strategy (Table 9). Producer’s accuracies for photos recording type ranged from 33% to 100% whereas producer’s accuracies for observations recording type were ranging between 50% and 100% (Figure 9).

The largest error in the classification of photos was that 80% of the wet heath (WH) reference data sites had been wrongly classified as upland swamp community (SS). The highest error in the on-site observations was that 50% of the dry heath (DH) and 50% of the wet heath (WH) reference data sites were classified as sandstone woodland (SW) and drainage woodland (DW).

![Stratified systematic samples photos and observations classification accuracy](image)

Figure 10: Accuracy of photos and observations compared to stratified systematic calibrated data.

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’

Numbers above Producer’s accuracy bars are the number of sites of each vegetation type in the reference data.
Table 8: Error matrix comparing post classified photos accuracy for stratified systematic samples

<table>
<thead>
<tr>
<th>Reference data</th>
<th>Photos Classified Classes</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>SL</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User's Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DH</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>DW</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>EW</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>GL</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>MW</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>OF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>RIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>11</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>SL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>SW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>WH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grand Total</td>
<td>5</td>
<td>8</td>
<td>40</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>0</td>
<td>3</td>
<td>23</td>
<td>5</td>
<td>104</td>
</tr>
<tr>
<td>Producer's Accuracy (%)</td>
<td>80</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>92</td>
<td>0</td>
<td>33</td>
<td>83</td>
<td>0</td>
<td></td>
<td>Overall Accuracy (94/104)*100=90%</td>
</tr>
</tbody>
</table>

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’.
Table 9: Error matrix comparing observational accuracy for stratified systematic samples

<table>
<thead>
<tr>
<th>Observation classes</th>
<th>DH</th>
<th>DW</th>
<th>EW</th>
<th>GL</th>
<th>MW</th>
<th>OF</th>
<th>RIP</th>
<th>SL</th>
<th>SS</th>
<th>SW</th>
<th>WH</th>
<th>Grand Total</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>11</td>
<td>3</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>DW</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>34</td>
<td>3</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>EW</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>20</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>GL</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>MW</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>OF</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>RIP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>SL</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>SS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>SW</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>WH</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Grand Total</td>
<td>4</td>
<td>8</td>
<td>38</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>20</td>
<td>2</td>
<td>99</td>
<td>99</td>
</tr>
</tbody>
</table>

GL is ‘alluvial grassland’, DW is ‘drainage woodland’, DH is ‘dry heath’, EW is ‘lowland woodland and tabletop woodland’, MW is ‘melaleuca woodland’, RIP is ‘riparian’, SW is ‘sandstone woodland’, OF is ‘tabletop open forest’, SL is ‘upland sedgeland’, SS is ‘upland swamp community’ and WH is ‘wet heath’.
3.4 CLASSIFICATION SCALE ACCURACY

3.4.1 NVIS-L3
Table 10 showed the error matrix of map reclassified up to NVIS-L3 and compared with the reference data from the random stratified data set. It illustrated the overall 60% map accuracy at this classification scale. The plot clearly depicted the individual category accuracies and showed that the user’s accuracies recorded in between the range of 8% to 67% and producer’s accuracies ranged from 6% to 75% (Figure 11).

69% sites of vegetation class ‘b’ (lophostemon woodland) and 83% sites of class ‘c’ (melaleuca woodland) were wrongly omitted into the vegetation type ‘d’ (eucalyptus woodland). Whereas, from class ‘d’ (eucalyptus woodland) 14% and 11% sites were inaccurately omitted into the vegetation class ‘f’ (eucalyptus open forest) and ‘a’ (grevillea shrubland). Results also showed that 12%, 2% and 1% sites were committed to this class which originally belonged to ‘b’ (lophostemon woodland), ‘c’ (melaleuca woodland) and ‘g’ (carallia open forest) (Table 10).

![NVIS-L3 classification accuracy](image)

Figure 11: Accuracy of NVIS-L3 map classification compared with random stratified reference data

a is ‘grevillea shrubland’, b is ‘lophostemon woodland’, c is ‘melaleuca woodland’, d is ‘eucalyptus woodland’, e is ‘sorghum grassland/cyperus sedgeland’, f is ‘eucalyptus open forest’ and g is ‘carallia open forest’.
Table 10: Error matrix comparing map classified up to NVIS_L3 with random stratified calibrated reference data classified up to NVIS_L3

<table>
<thead>
<tr>
<th>Reference data</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>Grand Total</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Litchfield Map</td>
<td>a</td>
<td>4</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>13</td>
</tr>
<tr>
<td>b</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>9</td>
<td>25</td>
<td>5</td>
<td>160</td>
<td>1</td>
<td>2</td>
<td></td>
<td>202</td>
<td>79</td>
</tr>
<tr>
<td>e</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>55</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>32</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
<td>27</td>
</tr>
<tr>
<td>g</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>67</td>
</tr>
<tr>
<td>Grand Total</td>
<td>14</td>
<td>36</td>
<td>6</td>
<td>229</td>
<td>8</td>
<td>13</td>
<td>16</td>
<td>322</td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy (%)</td>
<td>29</td>
<td>6</td>
<td>17</td>
<td>70</td>
<td>75</td>
<td>92</td>
<td>63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall Accuracy (195/322)*100=60%

a is ‘grevillea shrubland’, b is ‘lophestemon woodland’, c is ‘melaleuca woodland’, d is ‘eucalyptus woodland’, e is ‘sorghum grassland/ cyperus sedgeland’, f is ‘eucalyptus open forest’ and g is ‘carallia open forest’.

3.4.2 NVIS-L2

When all data were reclassified according to a NVIS-L2 classification, the comparison of the Litchfield vegetation map with reference data gave 72% overall map accuracy (Table 11). User’s and producer’s accuracies ranged in between 15% to 92% and 29% to 76% respectively (Figure 12). 53% and 17% of sites were inaccurately omitted from the vegetation class of ‘a’ (shrubland) and committed to ‘b’ (woodland) and ‘c’ (grassland/sedgeland) respectively. Out of vegetation class ‘b’ (woodland), 14% and 10% of sites were misclassified into vegetation class ‘d’ (open forest) and ‘a’ (shrubland).
Figure 12: Accuracy of NVIS-L2 map classification compared with random stratified reference data. a is ‘shrubland’, b is ‘woodland’, c is ‘closed grassland/ closed sedgeland’ and d is ‘open forest’.

Table 11: Error matrix comparing map classified up to NVIS_L2 with random stratified calibrated reference data classified up to NVIS_L2

<table>
<thead>
<tr>
<th>Reference data</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>Grand Total</th>
<th>User's Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Litchfield Map</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>5</td>
<td>27</td>
<td>1</td>
<td>33</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>9</td>
<td>201</td>
<td>2</td>
<td>6</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>38</td>
<td>22</td>
<td>60</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>17</td>
<td>268</td>
<td>8</td>
<td>29</td>
<td>322</td>
<td></td>
</tr>
</tbody>
</table>

Producer’s Accuracy (%) = (234/322) * 100 = 72%

a is ‘shrubland’, b is ‘woodland’, c is ‘closed grassland/ closed sedgeland’ and d is ‘open forest’

3.4.3 Litchfield management units

When the map and reference data were reclassified to match the Litchfield management units, the overall classification accuracy of the map was recorded as 81%. The highest agreement of accuracy was observed under the unit ‘2’ (sandstone plateaus) which showed 96% user’s and producer’s accuracy where the lowest user’s and producer’s accuracies were depicted for ‘4’ (melaleuca woodlands) being 25% and 10% (Figure 13).

The results showed that the management unit ‘3’ (monsoon rainforest and swamps) were misclassified into other units such as 12% in ‘2’ (sandstone plateaus), 12% in
‘4’ (melaleuca woodlands) and 8% in ‘5’ (lowland and alluvial plains). The highest percentage of error was observed in the classification of unit ‘4’ (melaleuca woodlands) as 75% samples were wrongly classified in unit ‘5’ (lowland and alluvial plains) (Table 12).

![Figure 13: Accuracy of Litchfield management units- map classification compared with random stratified reference data.](image)

2 is ‘sandstone plateaus’, 3 is ‘monsoon rainforest and swamps’, 4 is ‘melaleuca woodlands’ and 5 is ‘lowland and alluvial plains’.

Table 12: Error matrix comparing Litchfield map classified into management units with random stratified reference data classified according to Litchfield management units

<table>
<thead>
<tr>
<th>Reference data</th>
<th>Lithified map management Units</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Grand Total</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>146</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>152</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>17</td>
<td>5</td>
<td>1</td>
<td>28</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
<td>9</td>
<td>16</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>30</td>
<td>93</td>
<td>126</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td>152</td>
<td>25</td>
<td>40</td>
<td>105</td>
<td>322</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s Accuracy (%)</td>
<td>96</td>
<td>68</td>
<td>10</td>
<td>89</td>
<td>Overall Accuracy ((260/322)*100=81%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2 is ‘sandstone plateaus’, 3 is ‘monsoon rainforest and swamps’, 4 is ‘melaleuca woodlands’ and 5 is ‘lowland and alluvial plains’.

### 3.5 RESULTS OF KAPPA ANALYSIS

The Kappa analysis indicates that the map has moderate agreement with the reference data using both sampling techniques (Table 13) because the KHAT is
between 0.40 and 0.80 and the Z statistic is higher than 1.96 (Congalton and Green 2002). The pairwise comparison Z statistic of 0.96 (Table 13) showed that the error matrices (Tables 4 and 5) are not significantly different to one another.

Table 13: Comparison of Kappa analysis for the Litchfield vegetation map with reference data generated using both random stratified and stratified systematic sampling

<table>
<thead>
<tr>
<th>Error Matrix</th>
<th>KHAT</th>
<th>Variance</th>
<th>Z statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random stratified</td>
<td>0.57</td>
<td>0.0009</td>
<td>19.35</td>
</tr>
<tr>
<td>Stratified systematic</td>
<td>0.51</td>
<td>0.003</td>
<td>9.78</td>
</tr>
<tr>
<td>Z statistic (Pairwise comparison)</td>
<td></td>
<td></td>
<td>0.96</td>
</tr>
</tbody>
</table>

Statistical Kappa analysis of the dataset simplified to the Litchfield management units represented a moderate agreement of accuracy as KHAT values laid in between 0.40 and 0.80 whereas simplification of Litchfield management map to NVIS level 2 reported poor agreement of accuracy being below 0.40. The Z statistic results (Table 14) showed that these error matrices were significantly different to one another, as the pairwise comparison result was greater than 1.96.

Table 14: Comparison of Kappa analysis for the Litchfield vegetation map with random stratified reference data at broader classification scales; classified according to Litchfield management units and NVIS hierarchy level 2

<table>
<thead>
<tr>
<th>Error Matrix</th>
<th>KHAT</th>
<th>Variance</th>
<th>Z statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map Management units compared with Calibrated Management units</td>
<td>0.7</td>
<td>0.0008</td>
<td>7.58</td>
</tr>
<tr>
<td>Map NVIS L2 compared with Calibrated class NVIS L2</td>
<td>0.34</td>
<td>0.001</td>
<td>9.31</td>
</tr>
<tr>
<td>Z statistic (Pairwise comparison)</td>
<td></td>
<td></td>
<td>8.48</td>
</tr>
</tbody>
</table>
CHAPTER 4- DISCUSSION

4.1 MAP ACCURACY

Generally, the potential accuracy of a vegetation map depends on the remotely sensed data used for classification, application of an appropriate classification scheme and field reference data sampling methods (Card and Strong, 1989). The standard map accuracy assessment approach involves the comparison of a map with reference data using an error matrix (Lewis, 2011). This gives measures of the map accuracy in the form of overall accuracy and user’s and producer’s accuracies, and enables the assessment of statistical significance using Kappa analysis (Congalton, 1991 as cited in Lewis, 2011).

In this research the Litchfield vegetation map was assessed as having accuracy just less than 60%, which is moderate by the measure of Congalton and Green (2008). By other standards, this level of accuracy is considered unacceptable. For example, both the Survey and Mapping Guidelines of Regional Ecosystems and Vegetation Community produced by Queensland Herbarium & Environment Protection Agency (Neldner, et al., 2005) and the American National Vegetation Classification System for vegetation community mapping (Rapp et al., 2005) suggest the acceptable measure of vegetation map overall accuracy should be 80%. However, accuracy ranging from 50% to 70% for regional mapping programs is commonly accepted through the American National Vegetation Classification System (Rapp et al., 2005).

Several of the vegetation types had high user’s and producer’s accuracies—lowland woodland (EW), alluvial grassland (GL), riparian (RIP) and sandstone woodland (SW). This indicates a high level of separability between these vegetation classes. Probably, the main explanation for this is the variation in vegetation thickness and associated spectral reflectance (Penuelas et al., 1993). Apart from this these classes commonly occur in a unique environmental gradient in regions with high water availability and land suitability such as plains and hill slopes (Lewis, 2011).

Dry heath vegetation showed the lowest user’s and producer’s accuracy using both reference datasets. The main reason for this is that dry heath vegetation occurs as small patches which are hard to identify in Landsat imagery of 30m resolution. These small vegetation patches are generally surrounded by sandstone woodland, which
has a similar suite of species within diffuse boundaries. Therefore, it increases the chances of misclassification (Nordberg and Evertson, 2003). For both the field datasets the majority of errors were observed in the misclassification of drainage woodland (DW) to lowland woodland (EW), tabletop woodland (EWT) to tabletop open forest (OF) and sandstone woodland (SW) to drainage woodland. The vegetation types that were extensive and generally dominant in the study area were covered by plentiful reference data sites for the accuracy assessment using both sampling methods.

4.1.1 Random stratified and stratified systematic sampling

Analyses concluded that both the sampling methods proved equally effective in assessing the map accuracy. They produced nearly the same overall map accuracy and the result of the Z statistic in the Kappa analysis showed that there is no significant difference between the error matrices. Therefore, we recommend using the cheaper and quicker field sampling method, the random stratified sampling method, because it did not require pre-planning and is often chosen so that all classes have an equal opportunity of being selected for the sampling (Congalton 1991). However, vegetation types that occupy only a small area have a low chance of being sampled with this method. In this study, upland sedgeland (SL) was not sampled as it occupied the smallest area of all the vegetation types mapped.

4.2 PHOTOGRAPHS VERSUS ON-SITE OBSERVATION ACCURACY

Post field classification of site photographs had high overall accuracy (87% and 90% for random and stratified systematic sampling respectively) compared with the reference data that had been further calibrated using additional information. The accuracy of site classifications that were attributed through direct observation compared with the reference data was still reasonably high (79% and 75% respectively) but was lower than the post-field classified photographs.

4.3 DIFFERENT CLASSIFICATION SCALE ACCURACY

This research concluded that both NVIS classifications (level 2 and 3) had higher overall accuracy than the map classification which was equivalent to NVIS level 4 and 5. The highest overall accuracy (72%) was at the broadest scale-NVIS Level 2, in which the vegetation had been reduced to four simplified classes. However, at this
level there was still a high degree of confusion between individual class accuracy in particular woodland mapped as open forest or shrubland; and shrubland was mapped as woodland.

When the data were simplified to the broader classification of the Litchfield Management Units, the overall accuracy of the map compared with the reference data was considered very high (81%). However, at this scale, most of the Melaleuca Woodlands were not captured in the mapping and were instead classified as lowlands and alluvial plains. The conservation plan recommends different management strategies for these two management units, so this error in the mapping will lead to inappropriate management strategies being applied to Melaleuca woodlands. For example; it is the priority of the Litchfield National Park integrated conservation strategy (PWCNT, 2015) to maintain the biodiversity of habitats by controlling the major threats such as; fire, weeds and feral animals. One conservation target for melaleuca woodlands is that less than 65% of this habitat and parts of it should be burnt every year so that 33% of this habitat should remain unburnt over 3 years. Also, rangers have to observe fair or better conditions of weed absence and feral animal disturbance. Both of these goals are highly dependent on accurate information of spatial distribution of the community instead of the total area covered by this habitat.

Therefore, for ecological conservation, planning and management purposes for Litchfield National Park, it is recommended to use the vegetation map classification scheme at the level of management units. However it is very important to take into account the inaccuracies in the mapping. Serious management errors could result from relying on the map, so it is recommended for the mapping to be improved using the field data obtained in this study.

4.4 UTILITY OF KAPPA ANALYSIS FOR MAP ACCURACY ASSESSMENT

Kappa analysis is mainly carried out to assess the statistical agreements or disagreements of accuracy between the map and reference data. The results of the Kappa analysis for the random stratified and stratified systematic data resulted in KHAT values of 0.57 and 0.51, respectively. However both it and the overall accuracy do not capture the intricacies of the errors. The management units mapping is a great example as it had high overall accuracy and good kappa results,
but actually contained gross errors that would result in poor management. So, while kappa analysis and overall accuracy are useful measures, they cannot be used alone without further inspecting the information contained in the error matrices. On the other hand it was useful to be able to use the kappa analysis (pairwise Z statistic) to compare the two sampling methodologies. So maybe Kappa analysis isn’t really necessary for assessing the accuracy of the vegetation map (as this can be done using overall accuracy). But it has been useful for comparing different methods of reference data collection.

CHAPTER 5- CONCLUSION AND FUTURE WORK

This research work demonstrates a map accuracy assessment using field reference data collected using a variety of methods. It compared two reference data collection strategies – random stratified and systematic stratified – and showed that both were equally suitable for providing data for the map accuracy assessment. This means that the easier, less costly approach can be used; in this case the easier strategy was the random stratified. This study also compared two methods of data recording – photographs that were classified later versus on-site observation and classification. Results showed that the photographs provided a higher degree of classification accuracy than the on-site observations. Therefore, photographs would be the suggested method for future studies.

The main aim of the Provisional Vegetation Type map of Litchfield National Park was to provide accurate mapping of different vegetation types within the Park for integrated strategic planning with regards to ecological conservation and fire management. This study showed that the mapping is not fully effective for this purpose because some of the vegetation types have not been mapped correctly. In particular, the Melaleuca Woodlands were not effectively captured in the mapping and could be mismanaged as a result. Therefore, it is suggested that the field data obtained for this study be used to refine the Litchfield vegetation map and provide an updated product.
REFERENCES


**Appendices**

Appendix 1: Field survey sheet

<table>
<thead>
<tr>
<th>Date:</th>
<th>Time:</th>
<th>Place:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who:</td>
<td>Sampling interval:</td>
<td></td>
</tr>
</tbody>
</table>

**Confidence**

<table>
<thead>
<tr>
<th>Class</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riparian Forest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open Forest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euc. Woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandstone Woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melaleuca woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drainage Woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swamp Woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wet Shrubland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry Shrubland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedge/Grassland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainforest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Callitris</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Woodland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>