Australian Government

Queensland Wetlands Program

Queensland Government

Wetland Mapping and Classification Methodology

Overall Framework

A Method to Provide Baseline Mapping and Classification for Wetlands in Queensland

VERSION 1.2

Attachment 6

Accuracy Assessment Methodology and Results

Table of Contents

List of Appendices

List of Figures

List of Plates

1. Introduction

Remotely sensed data classified to map or represent natural systems inevitably contains inaccuracies. Factors such as scale, image quality, image resolution and selection, and the temporal variability in natural systems, can all contribute to errors in both the spatial and attribute accuracy of information derived from remotely sensed data.

Accuracy assessment techniques can help to determine the quality of the information derived from the remote sensing data and can be undertaken through qualitative or quantitative processes. The purpose of the quantitative accuracy assessment techniques outlined below is to identify and measure map errors in satellite derived water bodies. Such a technique generally requires the collection of an independent data or "reference" set which is assumed to be the "truth" and which can be statistically compared to the mapping results (Congalton and Green 1999).

As part of the development of the Wetland Mapping and Classification Methodology for Queensland, a quantitative accuracy assessment of the standing water body mapping was carried out. This attachment details the methods used and the results obtained from this assessment.

2. Methods

2.1 Summary

The presence or absence of a wetland feature derived from spectrally classified satellite imagery was assessed against recent fine-scale aerial photography to derive a reference data set. Although the classification techniques used were primarily aimed at standing water body detection, wetlands were chosen as the primary assessment feature given that the intended use of the classification was for mapping of wetland features. Land cover classes assigned during the assessment were later used to quantify and assess the accuracy of the techniques in mapping water bodies.

The assessment was carried out for four study areas, selected to cover the range of environments encountered in Queensland (Figure 1): coastal central Queensland (Mackay scene), coastal wet tropics (Ingham scene), inland (Eulo scene) and South-East Queensland (Tweed Heads scene). Two image classification techniques were assessed: Normalised Difference Water Index (NDWI) (McFeeters 1996) (Attachment 5) and Standing Water Body Density Slicing (SWB) (Environmental Protection Agency 2005) (Attachment 4). For each technique a multi-temporal satellite image classification was assessed for the results from five different Landsat TM images. Additional classifications were carried out and assessed for the Mackay and Ingham scenes on a selected "wet scene" (i.e. a single scene selected from a known period of above-average rainfall). A further constraint was also employed and assessed that involved the retention of shadow-like features that intersected with identified wetland regional ecosystems (REs).

2.2 Sampling strategy

Reference samples were located by randomly selecting points within each of the four study areas. The number of sample points was determined after a preliminary assessment of the mapping was carried out in the Tweed Heads area. Sampling was stratified by inundation regime and mapped water body area as follows:

Inundation Regime (first number):

1: inundated on one satellite image scene

2: inundated on 2-4 satellite image scenes

3: inundated on 5 satellite image scenes

4: wet satellite image scene only

Mapped Water Body Area (second number):

0: < 0.25ha (not sampled)

1: 0.25-2ha

2: 2-100ha

3: 100-1000ha

4: > 1000ha

Strata were denoted by a combined Regime:Area number. For example, a stratum of 2:3 would be a water body that is inundated on 2-4 satellite image scenes and has a mapped area of between 100ha and 1000ha. Not all combinations of Regime:Area occurred in each study area.

The three assessment trials were sampled for each of the two classification techniques as follows:

1. Multi-temporal (MT) data (all four study areas):
    - 40 sites in each occurring Regime:Area strata
    - 500 sites in areas not mapped as water body
2. Wet scene data (WET) (Mackay and Ingham study areas only)

    For areas on the WET cover that are not mapped as water body on the MT cover:
    - 40 sites in each occurring Area strata
3. RE constrained data (Mackay and Ingham study areas only)

    For areas on the RE MT cover that are not mapped as water on the MT cover:
    - 40 sites in each occurring Area strata.

Further sampling constraints were also imposed to ensure adequate representation of the geographic area and to avoid re-sampling of small water bodies. This involved locating sites no nearer than 75 metres from ones already sampled, and location of sites in areas not mapped as a water body to be no nearer than 50 metres to a mapped water body polygon. Note that the use of these constraints meant that it was not possible to allocate the full quota of sites in some strata.

Each reference point was assessed as either “wetland present”, “wetland not present” or “uncertain” (due to indeterminate photo pattern or because the wetland location identified was not within the precision of the imagery used). Where a reference point was not located within a wetland but was within 100 metres of a wetland, the point was assessed as wetland not present and the distance from the wetland recorded to the nearest 10 metres. Land cover classes (Appendix 1) were assessed and assigned to each reference point to identify those land cover types that contribute to misclassifications and to distinguish between wetlands and water bodies for analysis. Note that mangrove areas and narrow riparian or littoral features were excluded from water body assessments due to the difficulty in determining standing water from aerial photography.

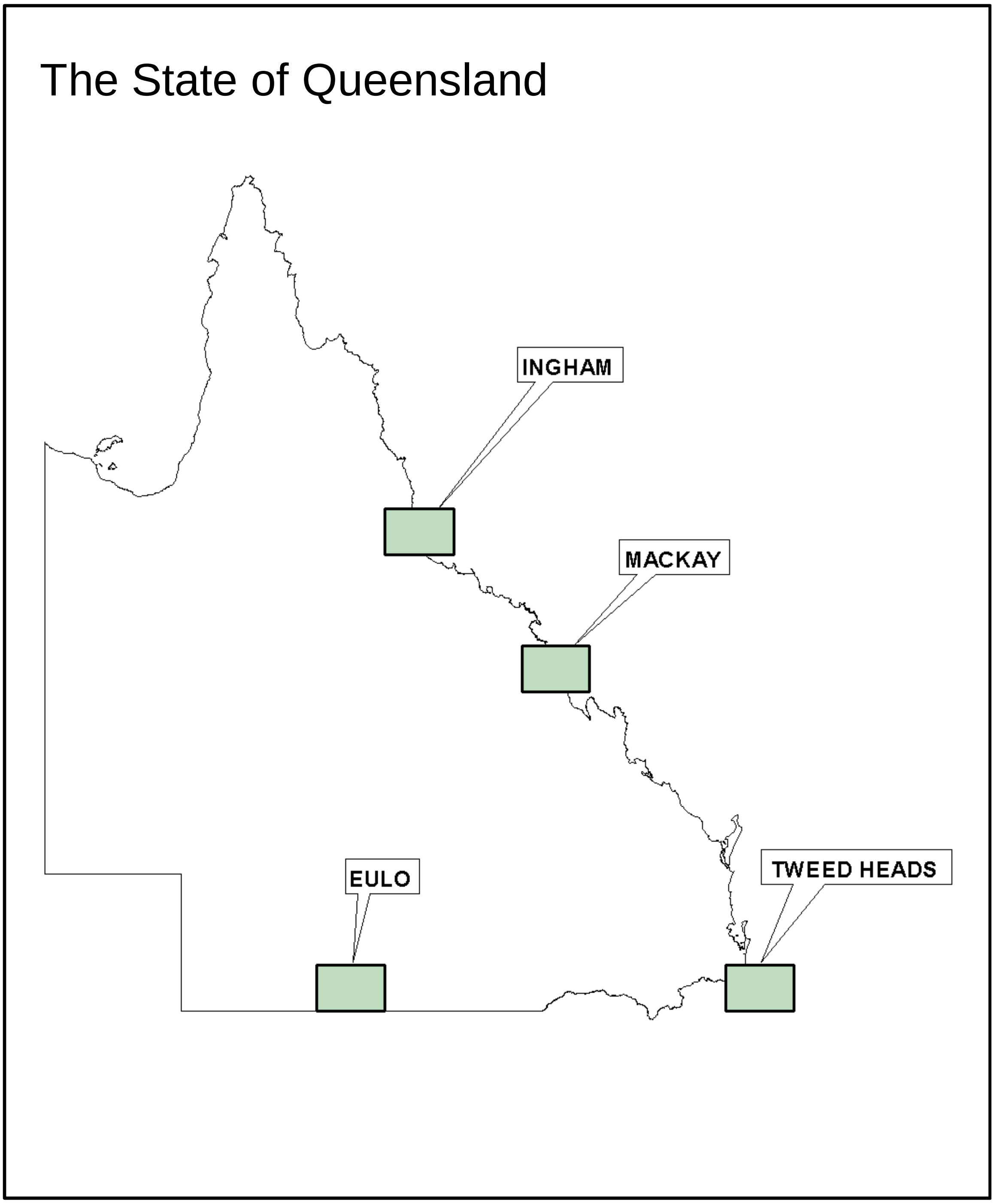

Figure 1. Location of study areas**.**

2.3 Data analysis

For the multi-temporal classifications undertaken by the two satellite image classification techniques, a two-way error matrix was produced for the wetland and water body assessments for each of the four study areas. Kappa statistics (Congalton and Green 1999) were calculated and assessed for each error matrix using purpose-designed software (Jenness and Wynne 2004). Kappa statistics measure the agreement between the predicted (classified) and observed (reference) categories of the data set after correcting for the agreement that occurs by chance. The resultant statistics were adjusted (weighted) for the proportion of area mapped as water body to correct for relative differences in the total area of water bodies mapped by each classification technique and therefore differences in the number of sites allocated to sampling strata. The key Kappa statistics (after Congalton and Green 1999) are:

a) *Producers accuracy*: the proportion of reference points in a mapped category that has been correctly mapped.

b) *Users accuracy*: the proportion of a mapped category that has been correctly mapped.

c) *Overall accuracy*: the percentage of points correctly classified by the mapping.

d) *Khat statistic*: the overall measure of agreement between the reference and mapped data, compared with chance agreement.

In addition, the Kappa statistics were used to derive the following statistics and comparisons:

a) *Estimate of total area of wetlands not mapped*: This statistic estimates image classification omission, which is the area of additional wetland area that could have been mapped when compared to reference land cover class maps and images. The estimate is based on the users accuracy of the category "not classified as wetland" and the area of all non-wetland classes mapped by each image classification method.

b) *Proportion correct of additional area on WET image but not on MT*: This is the ratio of the number of correct reference sites in the additional areas on the wet scene to the total number of reference sites allocated to the wet scene data.

c) *Proportion of area not mapped by other method as wetland (MT) that is correct*: This is calculated from the users accuracy of additional areas in each MT classification method relative to the other MT method (applied to whole wetland cover classes only).

d) *Pairwise comparison* of the Khat statistics (K_i) for the two satellite image classification techniques: $H_0 : (K_1 - K_2) = 0$ i.e. if $Z \geq Z_{\alpha/2,\, p = 0.05}$ then H_0 is rejected and the analyses are considered to be significantly different at the 95 percent test level, where $Z_{\alpha/2,\, p = 0.05} = 1.96$.

3. Results

3.1 Wetland accuracy assessment

Appendix 2 lists the producers accuracy, users accuracy and other summary statistics for wetlands for each study area assessed. The overall accuracy is consistently high across all four study areas ranging from 93 to 98 percent. The producers accuracy for non-wetland areas (i.e. how accurately non-wetland areas were identified) and the users accuracy for wetland and non-wetland areas (i.e. accuracy of the mapped areas) was also consistently high ranging from 77 to 99 percent. However, the producers accuracy for areas mapped as wetland (i.e. how accurately a wetland area was actually classified as wetland) was low, ranging from 14 to 54 percent. Thus, although there was a high level of certainty that areas mapped as wetland were actually wetland, the overall Khat statistic was relatively low, ranging from 24 to 64 percent. This indicates that relative to the actual area of wetlands present in a study area, the area of wetlands that was not mapped by either classification technique was large.

A comparison of the SWB (density slicing) and NWDI classification techniques across the four study areas show: the SWB method generally mapped more wetland areas (apart from Eulo); the SWB method has a greater overall accuracy resulting in a higher Khat value (apart from Tweed Heads); each technique mapped wetlands that were not mapped by the other and these areas show a high level of accuracy (generally > 90 percent). Pairwise comparison of the results of the two classification techniques (Appendix 2) for each study area shows that with the exception of the Ingham study area where the SWB method mapped more wetland area, the accuracy assessment results do not differ significantly at the 95 percent test level (Jenness and Wynne 2004). Differences in the Ingham study area are probably attributable to a larger area of wetlands mapped by the SWB method, resulting in higher users accuracy and therefore a higher overall Khat value.

In the two coastal areas (i.e. Mackay and Ingham) assessed for the additional wetland areas mapped from the wet image (compared with the multi-temporal image), both methods were highly accurate (95-100 percent). These additional areas were relatively small compared with the total area of wetlands mapped. However, visual assessment of the Eulo wet image suggests that, in inland areas, additional areas of wetland mapped from the wet image may be significantly larger. The RE constraint also resulted in further small additional areas being mapped which were also highly accurate (>90 percent; not reported here).

Figures 2 to 5 provide examples of the proportion of reference samples that were correctly mapped as wetlands for each Regime:Area sampling strata in the Ingham and Tweed Heads study areas. These figures generally reflect the high overall accuracy values calculated for both classification techniques in all four study areas. Stratification regimes that had relatively lower accuracy were generally those that had less frequent inundation regimes and smaller areas of wetland mapped. These lower accuracy figures are likely to be due to the temporal variability in the inundation regime on the imagery while the aerial photography presents static landscape conditions. In addition, shallower, saltier and more turbid wetland features tend to reflect more in Band 5 and are therefore often not included in the standing water body mapping (A. Knight, pers. comm.). For the assessment of reference sites, the smaller the area of a mapped wetland, the greater the likelihood that a reference sample may be interpreted as not located directly within the wetland. This is due to the spatial error associated with the imagery used and the difficulty occasionally experienced in identifying wetland extents from aerial photography.

Land cover classes that contributed most to misclassifications of wetland areas were plantation forests (especially pine forests) and areas of shadow in native vegetation in steep terrain. This is due to the similar spectral characteristics of these land cover types and standing water. However, as evidenced by the generally high overall accuracy of both classification techniques in all study areas, these errors are minimal and could be manually identified and corrected during later stages of the wetland map development. Land cover classes with low rates of misclassification included palustrine/lacustrine wetlands.

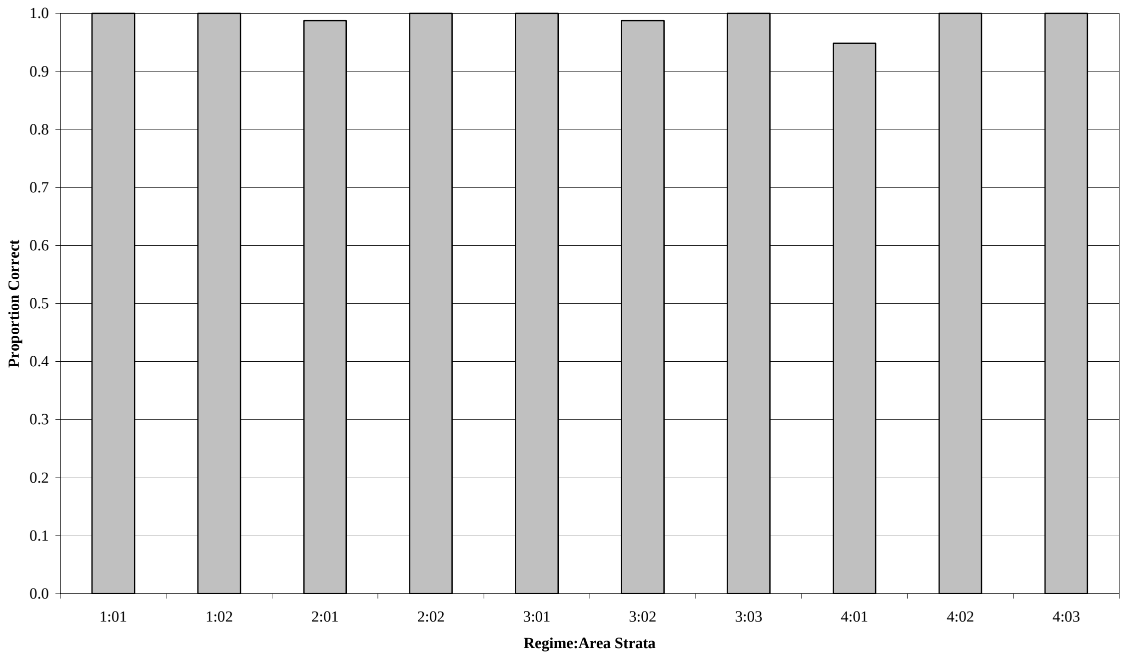

Figure 2. Proportion of correct reference sites by Regime:Area strata for the NDWI classification technique in the Ingham study area.

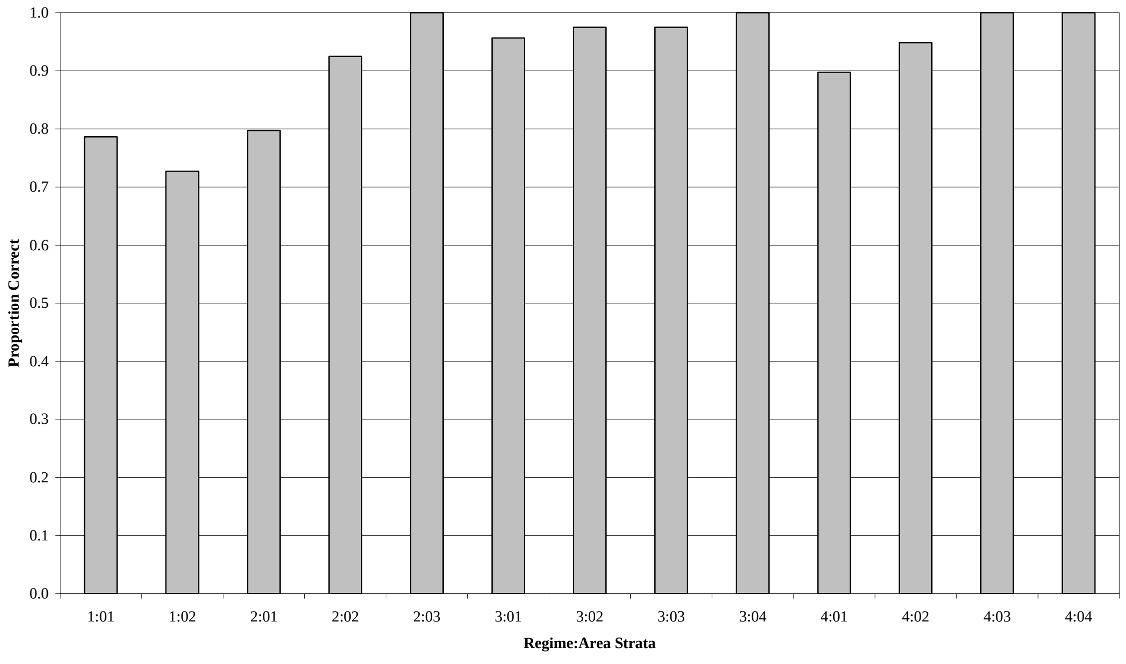

Figure 3. Proportion of correct reference sites by Regime:Area strata for the SWB classification technique in the Ingham study area**.**

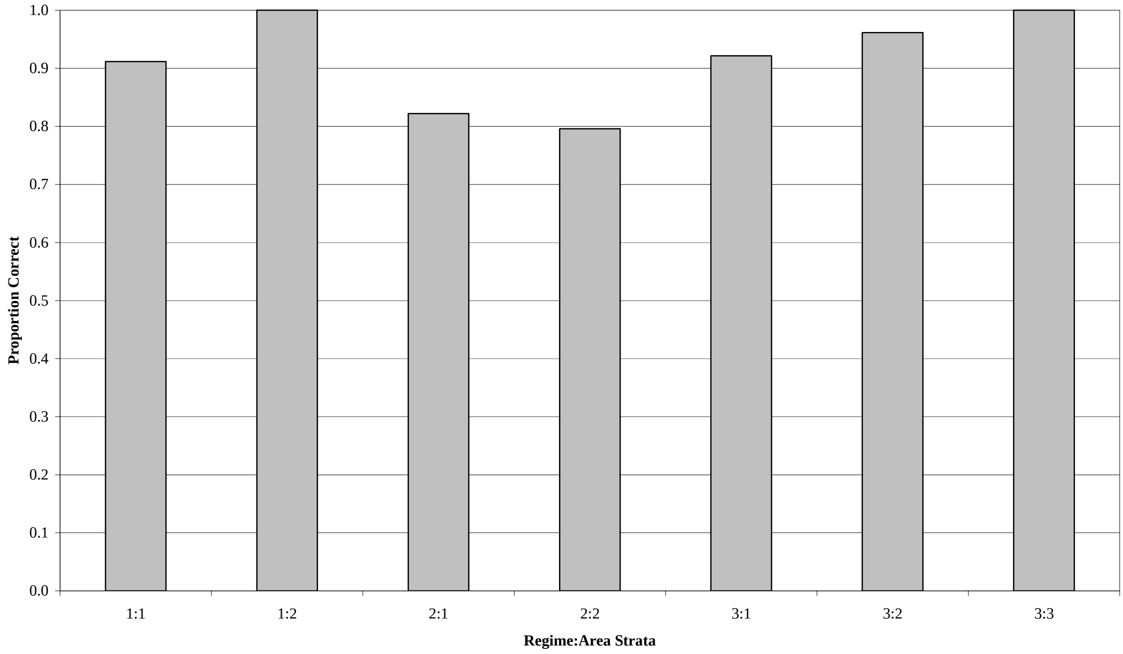

Figure 4. Proportion of correct reference sites by Regime:Area strata for the NDWI classification technique in the Tweed Heads study area.

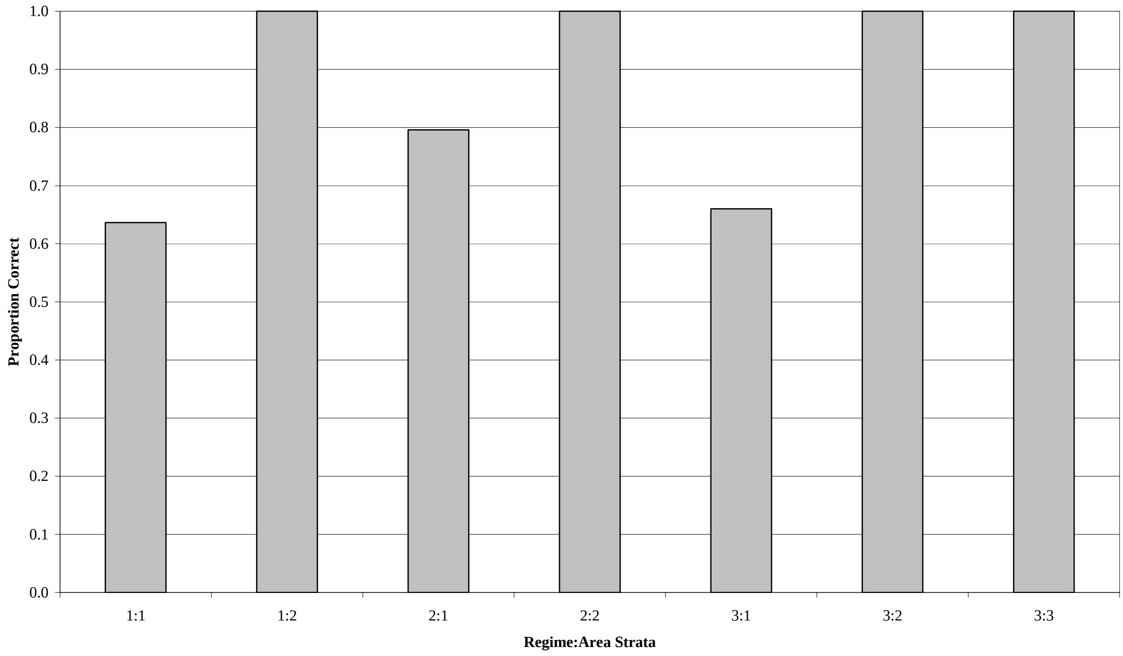

Figure 5. Proportion of correct reference sites by Regime:Area strata for the SWB classification technique in the Tweed Heads study area.

3.2 Water body accuracy assessment

Summary statistics calculated for land cover classes identified as a water body are shown in Appendix 3. The overall accuracy of the water body mapping is consistently very high across all four study areas ranging from 96-99 percent. This is marginally higher overall than the same statistic for wetlands. The producers accuracy and users accuracy for non-water body areas were also marginally higher than those for wetlands ranging from 96 to 99 percent (mostly >98 percent). There was also an improvement in the producers accuracy figures for water body areas compared with those for wetlands with values ranging from 31 to 88 percent. This reflects the two classification techniques primary focus on standing water body classification. However, even with respect to standing water bodies, it is still evident that significant areas have not been mapped by either classification technique. Furthermore, the users accuracy for water body areas was generally significantly lower than that for wetlands in the coastal study areas of Tweed, Ingham and Mackay, dropping to as low as 64 percent (NDWI) and 67 percent (SWB) in the Tweed Heads study area. This could be attributed to the exclusion of mangrove and estuarine areas from the water body analysis. These areas were highly accurate in the wetland assessment and their exclusion reduced the overall sample of "accurate" reference points relative to the total area of water bodies assessed.

Misclassifications in water body areas were primarily due to reference samples that were located in areas that had been missed by the classification techniques. These mainly included riparian water bodies and palustrine water bodies. As was the case with the wetland areas, the number of misclassified reference sites was minimal when compared to the total number of reference sites assessed. Non-water land cover classes that had been misclassified as water bodies were urban/infrastructure and narrow riparian systems. The error for riparian systems is likely to be the result of relatively large pixel size and minimum mapping unit.

4. Discussion

4.1 Accuracy assessment

The results presented here provide evidence that both techniques (DS and NDWI) accurately mapped areas of wetland and standing water bodies. It could be expected that the overall accuracy figures presented here would be improved by factoring for scale and resolution differences when assessing reference sites. This could be achieved by assessing reference sites within 40 metres of a wetland or water body as being correctly mapped. For water bodies, the users accuracy could be improved by re-adjustment of mapped proportions to only the area mapped within the tested land cover classes.

The results provide evidence that each technique misses areas of wetland and water bodies that the other maps accurately. Despite this, pairwise comparison shows that the classification techniques generally do not differ significantly in terms of overall accuracy. The results also show that there are slight increases in the area accurately mapped by the inclusion of the wet scene and the use of the RE constraint. It therefore follows that a combination of the NDWI and SWB techniques for the multi-temporal and wet imagery gives a more accurate wetland map than either of the two methods independently.

The low producers accuracy figures achieved in this assessment suggest that both methods miss substantial areas of wetlands and water bodies. This is to be expected given the difficulty that exists in separating and classifying the spectral reflectances of natural systems in remotely sensed imagery that has been captured at arbitrary time intervals. The methods employed by the wetlands mapping methodology will help to overcome these shortfalls and any inherent errors through manual interpretation and the inclusion of other reliable known sources of wetland data (e.g. RE mapping and Geodata layers) in the final wetland map. However, further accuracy assessment of the final product should be undertaken to verify the overall accuracy of map products and to enable both the producers and users of the map to have an understanding of the reliability and limitations of the mapped information.

4.2 Wetlands versus water bodies

The primary focus of the accuracy assessment was to assess reference sites for wetland or "not wetland". In many cases wetlands may not be classified as a water body because they do not have a substantial amount of open water visible to satellites. Therefore the low producers accuracy for water bodies and wetlands is primarily due to limitations in remotely assessing wetlands due to the extent and conditions of standing water at the time of image capture rather than errors in the satellite image classification techniques used to delineate water bodies. Examples of areas assessed as wetlands but not mapped as water are:

- riparian areas where the amount of standing water is too narrow to be reliably represented by the satellite image pixels (Plate 1);
- where there is dense vegetation such as mangroves, Melaleuca or fringing riparian forests that obscures water (Plates 1, 2 and 3); or
- areas where water was not present at the time of the satellite such as salt flats (Plates 2 and 3).

4.3 Validation

Validation of data and mapped products should be ongoing throughout the implementation of the methodology and undertaken prior to the production of final products. Validation is generally inherent in the methodology and may take the form of decision or assessment rules. Appropriate validation steps for the Wetland Mapping and Classification Methodology include:

- steps/rules in the methodology that check for internal inconsistencies in databases and maps;
- recording of sites (already on Queensland Herbarium site data base) and field sampling (existing plus extra) for individual wetland polygons or regions that are uncertain;
- recording the source of all derived products and mapped polygons;
- checking of final products by an independent survey and mapping team to ensure the methodology is followed;
- distribution of preliminary maps to regional staff and/or a technical panel for review; and
- assignment of qualitative reliability codes for areas as per Queensland Herbarium's vegetation mapping methodology (Neldner et al. 2004).

Plate 1. Riverine Wetlands from the Tweed area that are not mapped as a water body. Water bodies (blue) were mostly obscured by vegetation or were too narrow for satellite pixels. Accuracy assessment sample point shown by red dot.

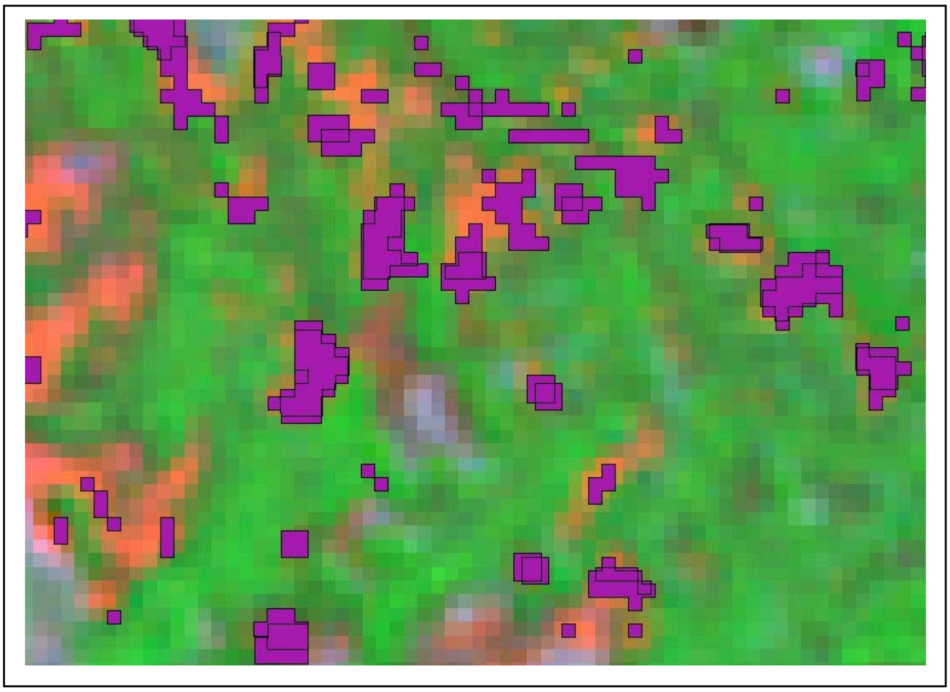

Plate 2. Estuarine wetland from the Mackay area that was not mapped as a water body (red arrow). Water bodies (purple) were obscured by mangroves (green areas) or salt flats (orange areas) that were not inundated by the tide at the time of the satellite image capture.

Plate 3. Melaleuca swamps from the Maryborough area not mapped as a water body. The water in the Melaleuca swamps (red arrow pointing to dark indigo coloured areas) was mostly obscured by vegetation, or the swamps were not inundated at the time of the satellite image capture. Existence of the swamp was verified by regional ecosystem mapping and, in this case, a site inspection.

5. References

Congalton, R.C. and Green, K. (1999) *Assessing the accuracy of remotely sensed data: Principles and practices.* Lewis Publishers/CRC Press: USA.

Neldner, V.J., Wilson, B.A., Thompson, E.J. and Dillewaard, H.A. (2004) *Methodology for survey and mapping of Regional Ecosystems and vegetation communities in Queensland*. Version 3.0. Queensland Herbarium, Environmental Protection Agency, Brisbane.

Jenness, J. and Wynne, J.J. (2004) *Cohen's Kappa and classification table derived metrics: An Arcview 3.x Extension for accuracy assessment of spatially-explicit models*. www.jennessent.com/arcview/kappa_stats

McFeeters, S.K. (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* 17(7):1425-1432.

Appendix 1. Land cover classes used in accuracy assessment

Land Cover Class Code	Description	Wetland Assessment	Water Body Assessment
M	mangroves, salt flats etc – tidal/estuarine wetlands	wetland	uncertain (excluded from analysis)
P	palustrine or lacustrine wetland	wetland	water body
RW	riparian water body	wetland	water body
R	riparian system (less than 40m wide)	wetland	not a water body
SM	small water bodies <0.25ha in area	wetland	water body
FL	floodplain	not a wetland	not a water body
C	canal/river estuary	wetland	water body
S	sea	wetland	water body
D	dam/reservoir	wetland	water body
N	not wetland or water body	not a wetland	not a water body
RF	rainforest on drainage line/gully heads	uncertain (excluded from analysis)	not a water body
1	natural vegetation	not a wetland	not a water body
2	pasture	not a wetland	not a water body
3	forest plantation	not a wetland	not a water body
4	crops	not a wetland	not a water body
5	urban/infrastructure	not a wetland	not a water body

Appendix 2. Summary of Kappa and zonal statistics for the accuracy assessment of wetland features within each study area.

INGHAM Multi-temporal (weighted*)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as wetland”	0.999	0.932	0.996	0.948
Statistics for “classified as wetland”	0.178	0.994	0.349	0.875
Overall accuracy	0.933		0.946	
Khat value	0.28		0.48	
Khat significance	0.00001		0.00000	
Estimate of total area of wetlands not mapped (ha)	53,349		40,212	
Proportion correct of additional area on WET image but not on MT	0.987		0.952	
Proportion of area not mapped by other method as wetland (MT) that is correct	0.984		0.968	
Pairwise comparison {H_0 : $(K_1 – K_2) = 0$}				
($Z_{\alpha/2,\ p = 0.05} = 1.96$)	Z = 2.145679			
EULO Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as wetland”	0.999	0.946	0.999	0.958
Statistics for “classified as wetland”	0.257	0.988	0.291	0.975
Overall accuracy	0.946		0.958	
Khat value	0.39		0.43	
Khat significance	0.00000		0.00000	
Estimate of total area of wetlands not mapped (ha)	85,472		66,562	
Proportion of area not mapped by other method as wetland (MT) that is correct	0.962		0.976	
Pairwise comparison {H_0 : $(K_1 – K_2) = 0$}				
($Z_{\alpha/2,\ p = 0.05} = 1.96$)	Z = 0.384424			
TWEED Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as wetland”	0.996	0.985	0.995	0.981

Statistics for "classified as wetland"	0.541	0.809	0.487	0.778
Overall accuracy	0.982		0.977	
Khat value	0.64		0.59	
Khat significance	0.00000		0.00000	
Estimate of total area of wetlands not mapped (ha)	2,012		2,540	
Proportion of area not mapped by other method as wetland (MT) that is correct	0.640		0.649	
Pairwise comparison $\{H_0 : (K_1 - K_2) = 0\}$				
$(Z_{\alpha/2,\ p = 0.05} = 1.96)$	Z = 0.424682			
MACKAY Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for "not classified as wetland"	0.999	0.948	0.999	0.956
Statistics for "classified as wetland"	0.146	0.978	0.172	0.893
Overall accuracy	0.948		0.955	
Khat value	0.24		0.28	
Khat significance	0.00156		0.00110	
Estimate of total area of wetlands not mapped (ha)	44,921		37,967	
Proportion correct of additional area on WET image but not on MT	1.000		0.951	
Proportion of area not mapped by other method as wetland (MT) that is correct	0.893		0.847	
Pairwise comparison $\{H_0 : (K_1 - K_2) = 0\}$				
$(Z_{\alpha/2,\ p = 0.05} = 1.96)$	Z = 0.282134			

* The weights applied are proportions that weight Kappa statistic values so that the sampling rates for each method may be compared. The sampling rates for each image classification method are standardised to the total area mapped by each method. Weighting was not applied to the two wet scenes (i.e. Ingham and Mackay), and the statistics for the wet scene data, which is based on limited sampling, is a direct ratio of the number of correct reference sites to the total number of reference sites allocated to the wet scene data.

Appendix 3. Summary of Kappa and zonal statistics for the accuracy assessment of water body features within each study area.

INGHAM Multi-temporal (weighted*)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as water body”	0.999	0.983	0.990	0.976
Statistics for “classified as water body”	0.641	0.936	0.497	0.696
Overall accuracy	0.982		0.969	
Khat value	0.752		0.564	
Khat Significance	0.00000		0.00000	
Estimate of total area of wetlands not mapped (ha)	13,337		18,559	
Pairwise comparison $\{H_0 : (K_1 - K_2) = 0\}$				
$(Z_{\alpha/2,\ p = 0.05} = 1.96)$	Z = 1.765			
EULO Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as water body”	0.999	0.962	0.999	0.970
Statistics for “classified as water body”	0.313	0.982	0.364	0.972
Overall accuracy	0.962		0.970	
Khat value	0.461		0.512	
Khat Significance	0.00000001		0.00000	
Estimate of total area of wetlands not mapped (ha)	60,147		47,544	
Pairwise comparison $\{H_0 : (K_1 - K_2) = 0\}$				
$(Z_{\alpha/2,\ p = 0.05} = 1.96)$	Z = 0.477			
TWEED Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for “not classified as water body”	0.992	0.995	0.992	0.998
Statistics for “classified as water body”	0.726	0.642	0.881	0.675
Overall accuracy	0.987		0.991	
Khat value	0.674		0.760	
Khat Significance	0.00000		0.00000	

Estimate of total area of wetlands not mapped (ha)	671		267	
Pairwise comparison {H_0 : $(K_1 - K_2) = 0$}				
($Z_{\alpha/2, p = 0.05} = 1.96$)	Z = 0.681			
MACKAY Multi-temporal (weighted)	**NDWI**		**SWB Density Slicing**	
	Producer	User	Producer	User
Statistics for "not classified as water body"	0.999	0.980	0.997	0.988
Statistics for "classified as water body"	0.318	0.924	0.378	0.718
Overall accuracy	0.979		0.985	
Khat value	0.465		0.488	
Khat Significance	0.00000406		0.00001298	
Estimate of total area of wetlands not mapped (ha)	17,277		10,355	
Pairwise comparison {H_0 : $(K_1 - K_2) = 0$}				
($Z_{\alpha/2, p = 0.05} = 1.96$)	Z = 0.152			

* The weights applied are proportions that weight Kappa statistic values so that the sampling rates for each method may be compared. The sampling rates for each image classification method are standardised to the total area mapped by each method. No accuracy assessment is provided for water body features in the wet scenes.