Calibration and validation of state wide land cover mapping

Elizabeth Morse-McNabb[1], Kathryn Sheffield[2], Rob Clark[1], Susan Robson[3], Hayden Lewis[4]

[1] Future Farming Systems Research Division, Department of Primary Industries Bendigo
Elizabeth.Morse-McNabb@dpi.vic.gov.au

[2] Future Farming Systems Research Division, Department of Primary Industries, Parkville
[3] Future Farming Systems Research Division, Department of Primary Industries, Horsham
[4] Future Farming Systems Research Division, Department of Primary Industries, Tatura

Abstract

The Victorian Land Use Information System (VLUIS) produces an annual land information product for the state of Victoria. It comprises three important characteristics of land information: tenure (ownership), use (property type), and cover (physical surface). The land cover component is created using a remote sensing approach. Using MODIS EVI 16-day image composites (MOD13Q1) collected over 12 months, phenology metrics such as start, end and length of the growing season are derived using the TIMESAT program. These metrics are used as input into the C5.0 program to derive a rule set to produce an annual state wide land cover classification product.

A significant portion of the resources and time required to generate a land cover product is spent collecting ground data to develop the rule sets for calibration and to validate the created product. Ground data collection is undertaken once a year from October through to December. Data is collected using a stratified random sampling approach. Initially the state is stratified by Primary Production Landscapes (PPLs) which divide the state into six regions based on terrain, geomorphology, soil types, climatic conditions, and land use and management information. Using the 1:25,000 scale map grid of the state, a random selection of map grids from each PPL is taken. Within these grids, parcels that were accessible within 20 m of formed roads are selected and, from within this selection, only parcels greater than 25 ha, after imposing a 150 m internal buffer, are chosen to allow for at least one MODIS pixel (250 m x 250 m) in each parcel.

This basic sampling framework has been used to collect ground data since 2009. To produce a random sample at minimal cost some map grids are resampled each year, while others are randomly added or taken out of the target sample. Some rare land cover classes, such as woody and non-woody horticulture, are not adequately sampled using this approach. A secondary sampling phase is used to target underrepresented land cover classes to ensure an adequate sample size for calibration and validation is achieved and from these supplementary data sets, a random selection is taken for validation. This paper outlines the approach taken for collecting calibration and validation ground data used to produce the VLUIS state wide land cover classification product.

Keywords: Calibration, Validation, Land cover, Sampling design

Biography of Authors:
Dr Elizabeth Morse-McNabb is a Senior Research Scientist in Remote Sensing with the Department of Primary Industries, located in Bendigo, Victoria. Dr Kathryn Sheffield is a Research Scientist in Remote Sensing with the Department of Primary Industries based in Parkville, Victoria. Rob Clark, Susan Robson and Hayden Lewis are Senior Technical Officers located in Bendigo, Horsham and Tatura, respectively.

Introduction
The production of land cover maps from satellite imagery is now commonplace; due to a vast array of satellite image products and their ease of access, low cost and larger area coverage. Accessing and processing imagery is a straightforward process with many software programs available with inbuilt classification mechanisms (e.g. ENVI™, Erdas Imagine™ and IDRISI Selva™). The collection of classification 'calibration' or 'image training' data is critical to the quality of land cover mapping and can be done in many ways, from ground based sampling (in-field mapping), to using free on-line sources of high resolution satellite imagery (Clark, Aide et al. 2010). Often substantial resources are used to collect calibration data and create a classified image. However; the result will only be a 'colourful picture' if proper validation of the resulting map is not done. Calibration and validation are both critical aspects involved in the creation of land cover maps.

Calibration Data

Supervised image classification requires calibration data to 'train' the chosen classification model. To ensure that the calibration data set describes all classes equally well, an approximately equal number from each category must be measured. It is also possible to create a weighted sampled representative of the area each class occupies; however, accuracy assessment of uneven sample numbers can be difficult (Strahler, Boschetti et al. 2006). The classification detail dictates the amount of calibration data required for a representative sample; the more detailed a classification scheme (i.e. more classes), the more calibration data required (Congalton 1991). Furthermore, to ensure that the calibration data can adequately describe the classification, the categories must be exhaustive and mutually exclusive i.e. they must cover the entire area of interest and each area to be classified should satisfy the criteria for only one category. This can be difficult to attain, especially in ecological studies (Congalton 1991). For example, after detailed field work to separate tree species it may not be possible to resolve the species in the imagery either due to the pixel size or spectral resolution of the imagery. This is particularly true for coarse resolution data (such as MODIS (Justice, Vermote et al. 1998), AVHRR (NOAA Satellite and Information Service 2012) and SPOT Vegetation (CNES 2012) where any single class label will be to some extent erroneous (Strahler, Boschetti et al. 2006). To be able to map a single land cover type to the chosen minimum mapping unit, it may be necessary to have a hierarchical classification, use a 'fuzzy' classification approach, or quantify the proportion of each category in the minimum mapping unit (Stehman and Czaplewski 1998). Congalton (1991) suggests that it is advantageous to use a hierarchical classification, as the need to generalise may arise and this can be achieved by collapsing categories. Calibration data can be targeted, both for categories and regions, and can be based in entirely homogeneous regions; however, data collected in this way must not be used for validation (Stehman and Czaplewski 1998). Foody et. al (2006) found that reductions in training size of ~90% were possible when focused on a specific class rather than all possible classes in the area of interest.

Validation Data

Validation of remotely sensed classification products uses a suite of techniques to assess map quality including: overall accuracy, errors of omission and commission by class, regional errors and class membership probability (Fassnacht, Cohen et al. 2006; Strahler, Boschetti et al. 2006).

Ground 'truth' data, to validate a map derived from remotely sensed data, needs to be obtained using a statistically valid sampling design. Problems associated with obtaining a statistically valid independent accuracy assessment include spatial autocorrelation, gaining a representative sample, sampling costs and sample characteristics (Congalton 1991; Muchoney and Strahler 2002). Pixels drawn from the same site are not independent since they are spatially auto-correlated (Muchoney and Strahler 2002).

To avoid bias, pixels used in validation should be independent of the classification training set and be collected using a sampling design independent of data collected for calibration purposes (Strahler, Boschetti et al. 2006). However; this 'ideal' approach is often difficult and time consuming and therefore there are three common ways to improve efficiency and still collect a statistically valid sample. First, ground data is often randomly split into two data groups, using an 80/20 (Muchoney and Strahler 2002) or 70/30 (Clark, Aide et al. 2010) split for calibration/validation data. Second, random stratified sampling, often in conjunction with some form of clustering, is used to provide a balance between a statistically rigorous sampling design and the practicalities and cost of ground data collection (Congalton 1991; Brown de Colstoun, Story et al. 2003; Nusser and Klaas 2003; Strahler, Boschetti et al. 2006; Clark, Aide et al. 2010). Third, using high resolution satellite imagery or aerial photography as a source of validation data is also a common approach and is particularly useful in inaccessible or heterogeneous areas (Wulder, White et al. 2007; Xie, Sha et al. 2008; Clark, Aide et al. 2010).

Stratification can be done on a basis of regions (to support region-specific accuracy assessments) or classes to ensure adequate samples of all classes are represented in support of class-specific accuracy assessments (Strahler, Boschetti et al. 2006). To create a statistically valid sample, all areas in the map must be available for sampling. If only areas of homogeneous land cover are available, then this requirement will not be satisfied (Stehman and Czaplewski 1998; Strahler, Boschetti et al. 2006). However, collecting heterogeneous pixels creates additional issues, in both the classification and validation of a map (particularly if a 'hard' classifier is used). To generate a statistically valid sample of homogeneous areas, the landscape can be stratified into homogenous/accessible and heterogeneous/inaccessible areas, and then random samples can be generated from each of these strata (Stehman and Czaplewski 1998).

While a stratified random sampling approach will provide a statistically robust data sample, inevitably some classes will be underrepresented in the final data set, creating issues for accuracy assessment. Targeted sampling can be used to supplement data from the probability sampling design which provides the core data set, provided samples from the targeted sampling are chosen in a probabilistic method (Stehman and Czaplewski 1998; Strahler, Boschetti et al. 2006). For example, Clark et al (2010) created a separate selection of sites for underrepresented classes and then took a random sample of these to ensure all classes had adequate samples. Other authors have also implemented a two-stage approach to ground data collection (Muchoney and Strahler 2002; Wulder, White et al. 2007).

As a guiding principle, a minimum of 50 samples per class is required for validation purposes (Hay 1979; Congalton 1991; Brogaard and Olafsdottir 1997). However, as the area to validate increases, so too does the number of samples required. Generally, between 50 and 100 samples per class are required (Hay 1979; Congalton 1991). If accuracy assessments involve a number of strata to be assessed independently, or there are a large number of classes, additional sites should be collected (Hay 1979; Congalton 1991). McRoberts (2011) found two accuracy assessment methods (model fitting and bootstrapping) performed better with larger sample numbers. Hay (1979) suggests a running total of sites per class is kept to ensure the sample is balanced and all classes are represented adequately. When a confusion matrix is to be used in the accuracy assessment, there needs to be sufficient samples to represent all classes and the potential confusion between classes.

Errors in the ground data including geospatial location errors and the assumption (for hard classifications) that pixels belong to a single class (Strahler, Boschetti et al. 2006) can result in a misrepresentative accuracy assessment. Using finer resolution imagery to provide validation data for coarser resolution imagery is a widely accepted practice, but the interpreted results from finer resolution imagery themselves are subject to errors (Xie, Sha et al. 2008; Clark, Aide et al. 2010). Accuracy of a map is essentially based on how it agrees with validation ground data, not strictly according to thematic accuracy (Congalton 1991; Strahler, Boschetti et al. 2006).

Validation of classified maps should involve a suite of measures including: overall map accuracy, accuracy on a class basis and a measure of the spatial variation in accuracy. The use of a confusion matrix is often suggested as this provides measures of overall accuracy and class-specific accuracy (Congalton 1991; Strahler, Boschetti et al. 2006). Errors of omission and commission are also derived from the confusion matrix. Some classification approaches enable a pixel 'rule' confidence map to be generated; this map shows the accuracy associated with each pixel. While not a statistically rigorous assessment, these confidence maps provide a spatial distribution of likely errors and a metric of classification quality per pixel, as well as providing complementary information to that derived from a confusion matrix (Strahler, Boschetti et al. 2006). Bootstrapping is a low cost method for assessing accuracy because it uses the same set of data for training and validation, where a sub-set of training data is withheld to be used for validation (McRoberts 2011). Bootstrapping can overestimate the accuracy slightly but is still a good indicator of relative errors across map classes (Strahler, Boschetti et al. 2006).

The process of collecting calibration and validation data for a state wide land cover product for Victoria in 2012 is discussed in this paper. This paper discusses the important elements of calibration and validation data collection as described in the literature and outline a procedure that from the evidence in the literature can effectively calibrate and validate a state wide land cover map in a timely and cost effective way.

The Victorian Land Use Information System (VLUIS)

The VLUIS project began in 2007. The objective of the project was to create a consistent and repeatable land use product for the state of Victoria. After an initial three year development period a method that used integration of state wide datasets was completed and a baseline VLUIS 2009 product created. The VLUIS product classifies the land tenure (public or private), the land use (based on the Australian Valuation Property Classification Code AVPCC) and land tenure as described in this paper in Table 1. Throughout the initial development period (2007-2010), land cover calibration and validation data were collected across the state. Due to the size and nature of the desired product outcome, it has taken many years of learning to fully realise an achievable system of on ground measurement

The VLUIS 2009 product is now widely used across government, industry and education sectors. Therefore, it is important to ensure that all future VLUIS products are of equivalent accuracy and comparable classification for these user groups. This paper explains the current development based on ground data collection experience since 2009 and describes the proposed method in light of current and future budget restrictions. Results and discussion of land cover product accuracy assessments from 2009, 2010 and 2011 are not presented in this paper.

Method

The Victorian Land Use Information System (VLUIS) produces an annual land information product for the state of Victoria (Morse-McNabb 2011). The following section discusses the methodological approach that underpins the classification and collection of calibration and validation data for the land cover component of the VLUIS product.

The land cover component of the product is created using a remote sensing approach. Using a time-series stack of MODIS EVI 16-day image composites (NASA LP DAAC 2011) collected over 12 months, phenology metrics such as start, end and length of the growing season are derived using the TIMESAT program (Eklundh 2010). These metrics are used as input into the C5.0 program (Rulequest 2009) to derive a rule set for each land cover class to produce an annual state wide land cover classification product.

Dominant Land Cover Classification

A two-tiered land cover classification was developed for the VLUIS land cover map product. The primary classes and their descriptions are shown in Table 1. These classes have been determined from three years (2009, 2010 and 2011) of field data collection experience. They represent land cover types that can be separated using imagery with coarse spatial resolution such as MODIS, based on temporal phenological or spectral differences. The term 'dominant' is used because each MODIS pixel covers approximately 6.25 ha and will often contain a mix of covers, making any single class erroneous (Strahler, Boschetti et al. 2006). A simple example of 'dominance' occurs in plantation forestry. Here the trees (either hardwood or softwood) when planted are very small (< 20 cm high and 10 cm diameter) and therefore the pasture (or bare ground) surrounding them remains the 'dominant' land cover type for at least 2-3 years whilst the seedlings establish.

Table 1: Dominant Land Cover Classification for Victoria

DOMINANT LAND COVER CLASS	DESCRIPTION
Bare ground and non-photosynthetic vegetation	Very little or no green vegetation has been measured in these areas for most of the year
Brassica	Annual broad acre crops grown for oil from the Brassicaceae family. E.g. Canola
Cereals	Broad acre crops in the Poaceae family. E.g. Wheat, Barley, Oats
Deciduous woody horticulture	Deciduous woody plants that are grown for many years for fruit each year.
Evergreen woody horticulture	Evergreen woody plants that are grown for many years for fruit each year
Hardwood plantation	Hardwood trees at least 2 years old that are planted to large continuous areas in rows.
Legumes	Broad acre crops in the Fabaceae family. E.g. Lentils, Chickpeas and Faba beans. A varied class. This class does not include lucerne or clover
Native woody cover	Naturally established trees, multiple ages, understory may or may not be present not in rows
Non-woody horticulture	Herbaceous, mostly annual, paddock grown fruit and vegetables
Pasture/grassland	Any herbaceous ground cover that is present for most of the year. Management can vary from grazing to cutting many times in one year. The plants may be annual or perennial and may be grasses or legumes (e.g. lucerne) or medics (e.g. clover)
Softwood plantation	Softwood trees at least 2 years old that are planted to large continuous areas in rows.
Water	Large water bodies such as large dams, reservoirs, coastlines etc.

Sample Site Selection:

There are five stages of site selection. First, the state is stratified into Primary Production Landscapes (PPL). These are regions grouped on soil, landscape, land use and climate and are used to obtain a geographic spread of ground samples (MacEwan, Robinson et al. 2008). Second, 1:25,000 map sheets are randomly selected from each PPL strata. The number of grids sampled in each PPL is based on the size of the PPL. However, due to resource limitations in 2012 fewer could be selected. Third, cadastral parcels larger than 25 hectares that are within the grid cells (Figure 1) and that are within 20 m of roads are selected. Fourth, all parcels are buffered with a 150 m internal buffer. Steps three and four are included because we are using MODIS imagery. The coarse spatial resolution of MODIS cannot map linear features and as linear features are associated with parcel boundaries (i.e. roads, fence lines, waterways, railway lines) then these areas are removed from the sample. This process is not used to select homogeneous areas or reduce thematic error, but to account for potential errors in mis-registration of the data (Strahler, Boschetti et al. 2006). Fifth, MODIS pixels that fall entirely inside the buffered area within each parcel boundary are available and may be selected for mapping (Figure 2).

Based on the literature search conducted, in 2012 we aim to collect 50-100 sites for each class as validation data (Hay 1979). This sums to between 600 and 1200 validation sites for the state for the 12 dominant land cover classes. Based on an 80/20 calibration/validation data split, we also aim to collect 200-400 sites per class for calibration purposes, giving a total of 250-500 sites per class in total. This sample size will result in 3000-6000 sites in total from all data sources (600-1200 validation, 2400-4800 calibration).
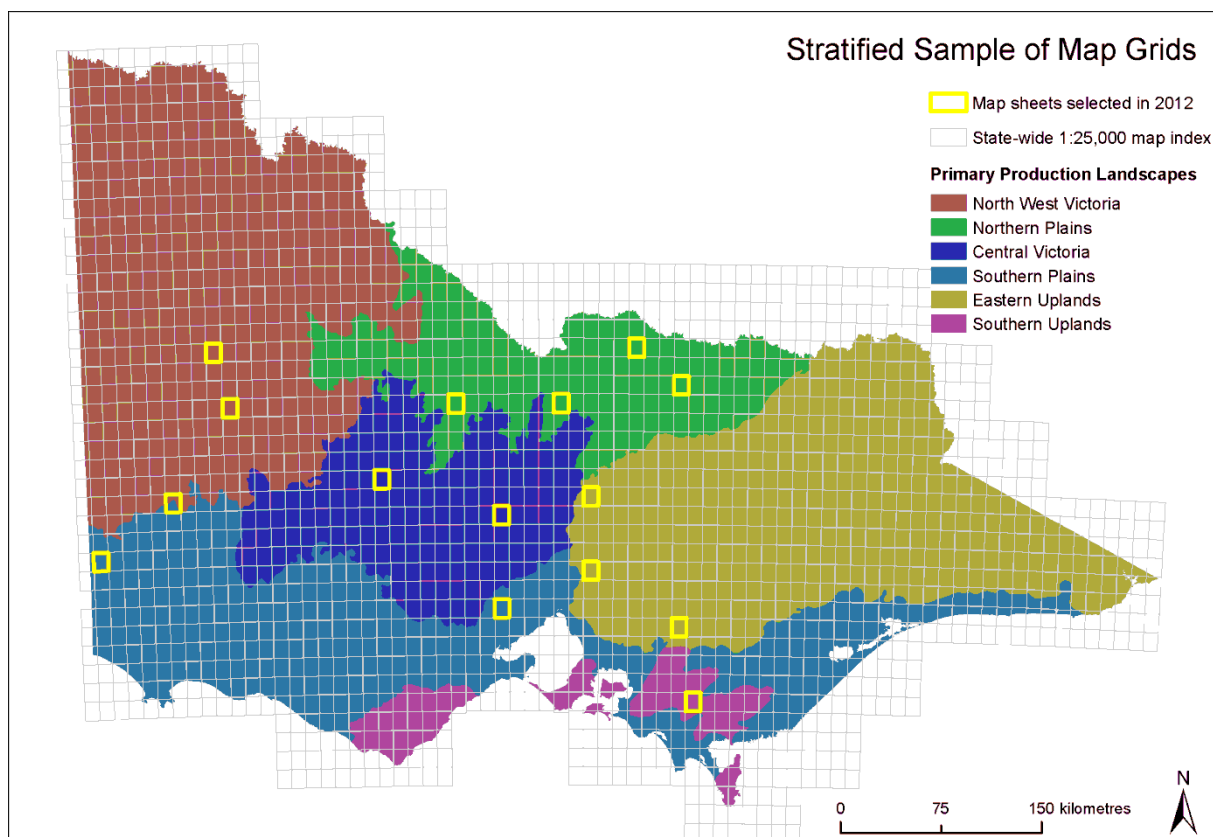

Figure 1: Fifteen map grids have been selected for sampling in 2012

Only 15 map sheets have been selected for 2012 due to resource and budgets constraints. MODIS pixels within parcels have been identified and these will be used as a minimum sampling unit. From the data collected within map sheets, single pixels within 50-100 parcels per class will be selected to use as validation data, with the remaining samples, up to a predetermined maximum sample size, to be used for calibration. Additional calibration sites will be withheld to ensure the data set does not introduce bias into the map product.
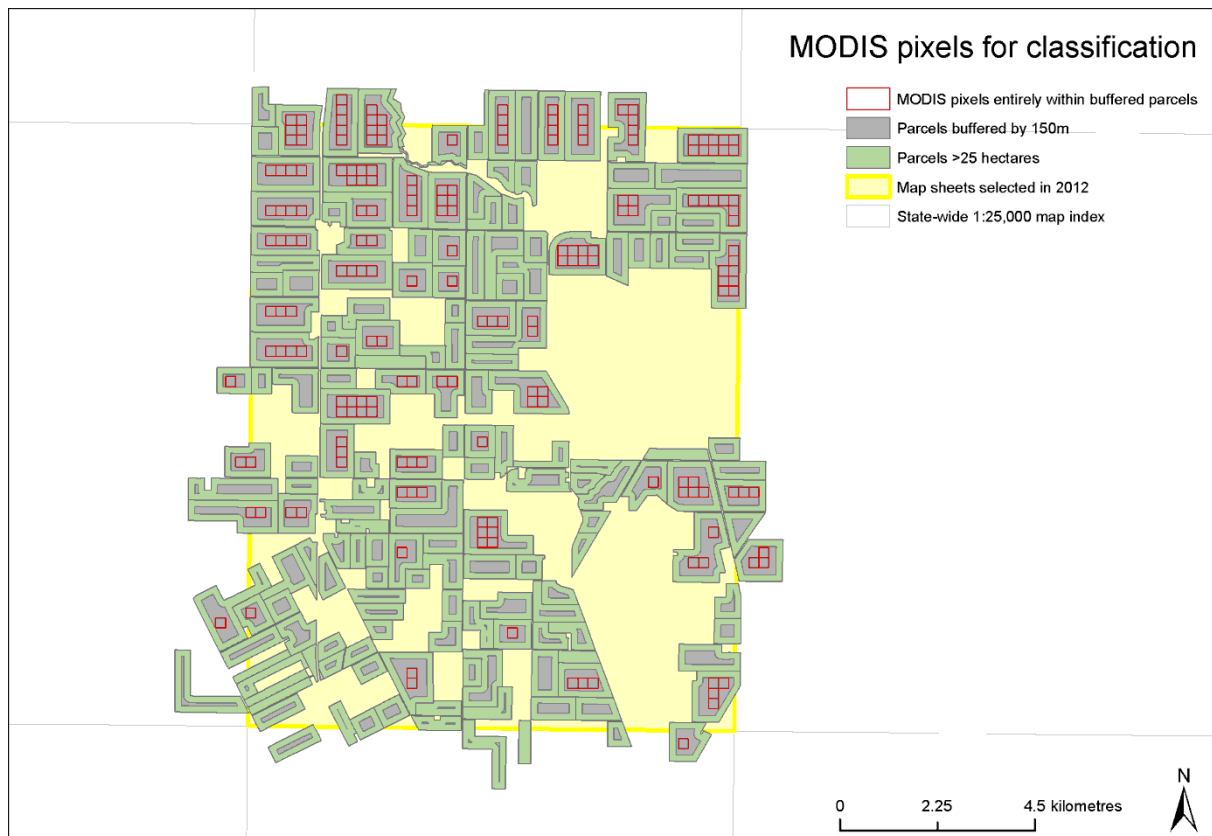
Figure 2: Parcels greater than 25 hectares that intersect with map grids are selected and then an internal buffer of 150 m is made before MODIS pixels that fall entirely within are selected for classification.

This basic sampling framework has been used to collect ground data since 2009. To minimise the cost of data collection, some map grids from this framework are sampled each year, while others are randomly added or taken out of the target sample. Three of the 15 grids selected for 2012 were sampled in 2011. Time and resource restrictions in 2012 have limited the validation sample to 15 grids; however, greater effort has been made to ensure an equal number of validation and calibration samples are taken for each class. The land use information from the VLUIS product, which originates from the Victorian Valuer-General (VVG), has been used to identify sample locations for classes that are underrepresented. Parcels within a 20 km radius of selected map sheets have been selected to optimise project resources (namely time and money). MODIS pixels within these parcels have been identified and will be sampled as an opportunistic targeted survey (Figure 3). From these surveys, sites will be randomly selected to fulfil the validation quota for each class, with the remainder to be used for calibration.
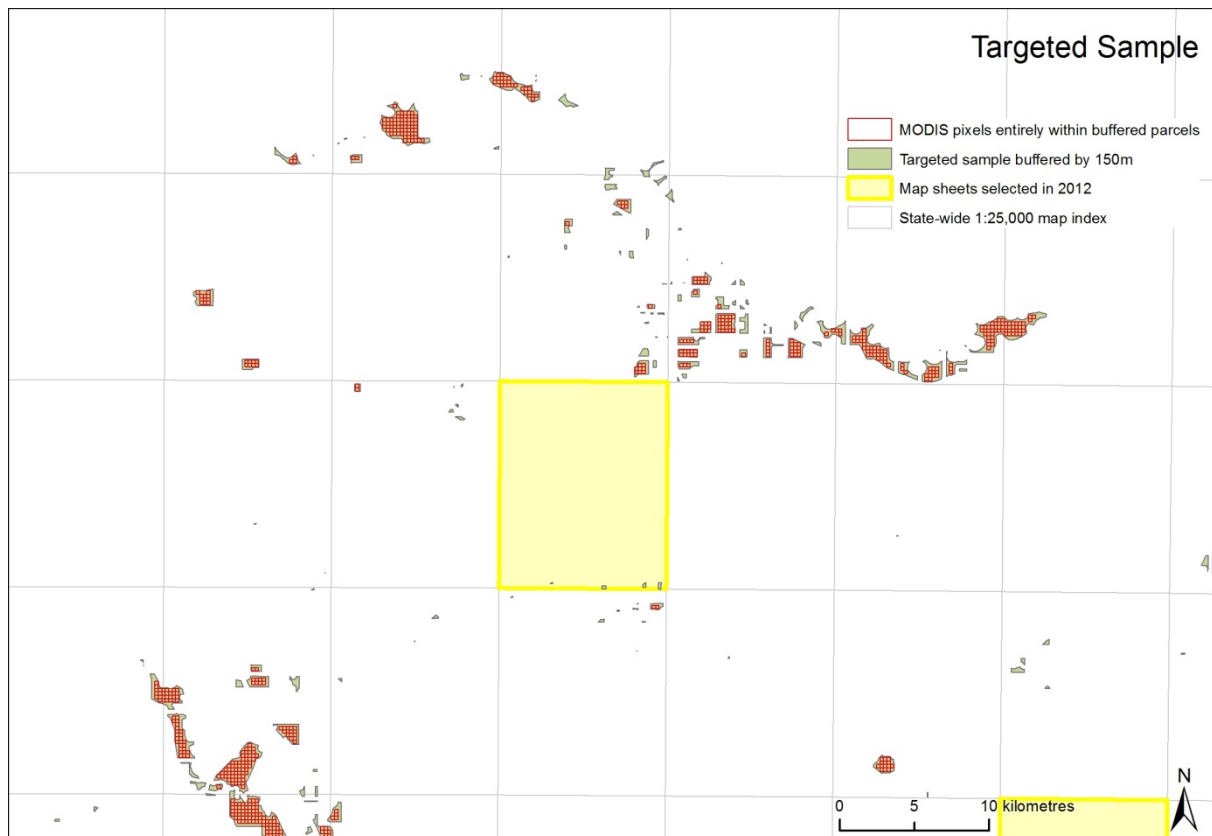
Figure 3: Parcels that contain rare land cover classes are identified within 20 km of the map grids and MODIS pixels that fall entirely within the parcel are selected for classification.

Field Data Entry

October to December is the best time to visually discriminate production environments (particularly crops). As the aim is to classify the image dataset, the minimum mapping unit is the MODIS pixel at a spatial resolution of approximately 250m x 250m. In 2012 a simple set of characteristics will be noted for each minimum mapping unit. These are: primary cover (from Table 1), secondary cover (that has a hierarchical relationship to the primary code), reliability (that describes spatial homogeneity) and some free text area for notes. All other ancillary spatial information i.e. parcel location, roads, waterways, aerial photography will be loaded into ESRI ArcGIS ArcView™ (ESRI 2012) on a field laptop prior to field work (Figure 4).
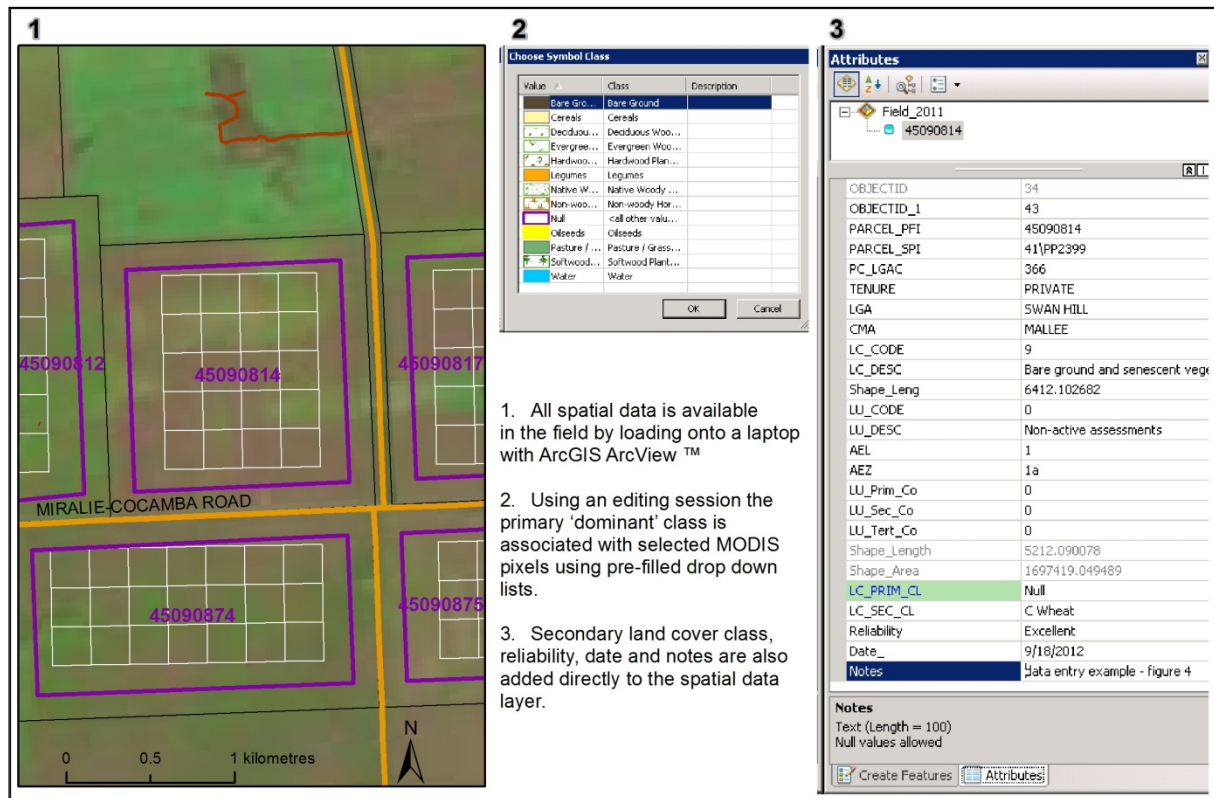
Figure 4: An ESRI ArcMap project is used in the field to categorise the minimum mapping units (MODIS pixels).

Results

Table 2 shows the number of parcels and individual pixels in land covers grouped by land use type (broadscale agriculture, horticulture, commercial forestry, woody native vegetation and water). Parcels and pixels were selected using the processes of random sample site selection and targeted sampling outlined in the Method section. The number of pixels and parcels is shown to clarify the true independent sample number. Not all pixels will be used, to avoid issues with spatial auto-correlation. By using the stratified random sample approach, with a limit of only 15 grids (due to cost restraints), it is estimated that the horticultural land cover categories are greatly under-represented.

Table 2: The number of parcels, and pixels contained within the parcels, which have been selected through a random and targeted approach.

Land Covers	Random Selection Parcels	Random Selection Pixels	Targeted Selection Parcels	Targeted Selection Pixels
Bare Ground, Brassica, Cereals, Legumes, Pasture (i.e. agricultural cropping activities)	1,920	3,379	0	0
Evergreen, Deciduous and Non – Woody Horticulture	20	16	465	901
Softwood & Hardwood Plantations	26	514	590	5,817
Native Woody Cover	24	635	970	57,243
Water	0	0	0	0

The VLUIS land use data has been used to estimate potential land cover types in the random sample site selection. As some land uses can have a range of land cover types; Table 2 denotes the assumed groupings. The first group, relates to agricultural cropping and grazing. The number of randomly selected parcels, in the first group, is adequate for calibration (200) and validation (50) alone and therefore a targeted selection is not required. The second group (horticultural land cover types) is not adequately sampled using the random approach and the targeted random selection approach also fails to have an adequate number of parcels for both calibration and validation. The ratio of parcels to pixels shows that horticultural parcels are small; horticultural properties are also found in small discrete areas across the state. These factors make it difficult to gather an adequate sample from either a targeted or random approach. Therefore, in addition to the random and targeted data in Table 2, survey data from SPC Ardmona Census and Sunrise21 will be used to supplement the calibration data. The SPC Ardmona Census maps all varieties of fruit, by block, that supply the SPC/Ardmona factories and Sunrise21 is a Mildura based company that maps each fruit variety in the Mildura district by block each year.

Softwood and Hardwood plantations and Native Woody Cover classes have too few parcels in the random selection but are adequately sampled in the targeted selection. Both of these groups have large parcel areas. Only the required number of pixels will be sampled to avoid spatial auto correlation, bias and redundancies.

Water is not targeted in the site sampling selection as all water bodies are well mapped in Victoria and therefore both validation and calibration data can be obtained from a purely desk-top study (Department of Sustainability and Environment 2012).

Discussion

The collection of ground data to develop calibration rules and to validate broad scale land cover products is a difficult task that requires significant planning, resources and time. Understanding the impacts of what is to be measured, in terms of classification, and the follow on effects in accuracy assessment is important. The land cover classification, created for this land cover mapping project has 12 primary classes. It is a simplistic classification based on land cover types that can be distinguished by time series analysis of coarse resolution MODIS vegetation imagery. However, 12 classes still require at least 3000 samples (2400 calibration and 600 validation) across the state, a large investment in time and resources. Nevertheless, we can demonstrate that most of the classes (9 from 12) have both an adequate calibration and validation sample and the remaining three horticultural classes can be sampled using targeted methods and supplementary datasets without reducing statistical validity.

Without a statistically valid random sample of validation sites it is not possible to assess the accuracy of a land cover classification. The process outlined in this paper has used methods and frameworks described by Hay (1979), Congalton (1991), Stehman and Czaplewski (1998) and reiterated in the review by Strahler (2006). Every attempt has been made to meet three main principles. First, each category must have an equal number of calibration and validation points. For this principle we used both random and targeted selection; though, even with targeted selection we found it difficult to collect enough rare (horticulture) classes. Second, a sufficient number of samples must be collected to represent each class and any confusion between each class. Due to time and resource constraints we are only able to collect the minimum in a recommended range from 50 -100 per class. There are no hard rules about the number of samples required for validation in the literature but the general consensus is at least 50-100 and generally more samples are better (Hay 1979; Congalton 1991). However, a smaller probability based sample is better than a large non-probability sample as this cannot be used for generalised accuracy assessment (Stehman and Czaplewski 1998). Third, when the minimum mapping unit is the pixel it is very tempting to collect many pixels from one area; yet, this introduces spatial auto-correlation that can introduce bias into the results. Undertaking a representative sample, whilst understanding the problems of autocorrelation; are the two main issues to arrive at a statistically independent accuracy assessment (Muchoney and Strahler 2002). Nevertheless; Curran and Williamson (1986) suggest that random sampling can still introduce the effect of spatial auto-correlation as it is inevitable that some sample sites will be close and therefore each measurement will include information about the neighbour. As we have information on land use at the cadastral parcel level we have used this information to identify groups of pixels with similar land management influences and apart from rare classes will only take one pixel from each parcel group. Rare classes, such as non-woody horticulture, are difficult to capture in a random selection because they exist in discrete areas. They are also difficult to collect in a targeted manner as many of the areas are very small and if the parcel clumping method is used the parcels are too small to provide even single pixels for selection. Muchoney and Strahler (2002) suggest two options; either remove the rare class from the classification or take extra targeted samples. Before removing the class from the classification we will use supplementary external data sets from reliable sources and collect neighbouring pixels to increase the sample size and assess the results.

We have used stratification in this site selection method to ensure an adequate geographical spread across Victoria. Due to resource restrictions in 2012, we estimate that the on ground sampling of only 15 map grids will be possible. As there are six main PPL areas, of variable size, across the state it is unlikely that we have achieved an adequate stratification. The lack of stratification does not mean the sample is not statistically valid as all areas had equal chance of selection (Strahler, Boschetti et al. 2006). Nevertheless, the sample size for 2012 is reduced and has not been effectively stratified.

Conclusion

The aim of the sampling procedure outlined in this paper is to provide a statistically valid dataset that can be used to generate and assess the accuracy of the resultant land cover map. In order to achieve this aim; a simple but mutually exclusive and exhaustive 12 class land cover classification has been created. With only 12 land cover classes it is possible to create a sampling framework where each of the classes is adequately sampled for both calibration data and validation data, either by using a statistically valid random sampling approach or through a secondary targeted sampling of underrepresented classes. The bias that can be introduced through spatial auto-correlation of image pixels has been addressed using single pixels from separate cadastral parcels. Most importantly the sampling procedure has been designed within current project constraints. Therefore, based on the method presented in this paper, it will be possible to make an assessment of the overall accuracy, errors of omission and commission and regional errors in the land cover map in 2012.

Acknowledgements

References

Brogaard, S. and R. Olafsdottir (1997). Ground-truths or Ground-lies? Department of Physical Geography, Lund University. **1**.

Brown de Colstoun, E. C., M. H. Story, et al. (2003). "National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier." Remote Sensing of Environment **85**(3): 316-327.

Clark, M. L., T. M. Aide, et al. (2010). "A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America." Remote Sensing of Environment **114**(11): 2816-2832.

CNES (2012). "Vegetation." Earth observation mission. Retrieved 18th Sept,, 2012.

Congalton, R. G. (1991). "A review of assessing the accuracy of classifications of remotely sensed data." Remote Sensing of Environment **37**(1): 35-46.

Curran, P. J. and H. D. Williamson (1986). "Sample size for ground and remotely sensed data." Remote Sensing of Environment **20**(1): 31-41.

Department of Sustainability and Environment (2012). Vicmap Hydro 1:25,000. Melbourne, Victorian Spatial Data Directory.

Eklundh, L. J., P (2010). TIMESAT 3.0 software Manual.

ESRI (2012). ArcGIS Desktop 10 Service Pack 5. Redlands CA, Environmental Systems Research Institute: ArcInfo.

Fassnacht, K. S., W. B. Cohen, et al. (2006). "Key issues in making and using satellite-based maps in ecology: A primer." Forest Ecology and Management **222**(1â€'3): 167-181.

Foody, G. M., A. Mathur, et al. (2006). "Training set size requirements for the classification of a specific class." Remote Sensing of Environment **104**(1): 1-14.

Hay, A. H. (1979). "Sampling Designs to Test Land-Use Map Accuracy." Photogrammetric Engineering & Remote Sensing **45**(4): 529-533.

Justice, C. O., E. Vermote, et al. (1998). "The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research." Geoscience and Remote Sensing, IEEE Transactions on **36**(4): 1228-1249.

MacEwan, R., N. Robinson, et al. (2008). Primary Production Landscapes of Victoria. 14th Australian Society of Agronomy Conference. Adelaide.

McRoberts, R. E. (2011). "Satellite image-based maps: Scientific inference or pretty pictures?" Remote Sensing of Environment **115**(2): 715-724.

Morse-McNabb, E. M. (2011). The Victorian Land Use Information System(VLUIS): A new method for creating land use data for Victoria, Australia. Spatial Sciences & Surveying Biennial Conference. Wellington, New Zealand.

Muchoney, D. M. and A. H. Strahler (2002). "Pixel- and site-based calibration and validation methods for evaluating supervised classification of remotely sensed data." Remote Sensing of Environment **81**(2â€'3): 290-299.

NASA LP DAAC (2011). "MODIS Land, Vegetation Indices MOD13Q1 Collection 5,." Retrieved 14/04/2011, from http://modis-land.gsfc.nasa.gov/vi.htm.

NOAA Satellite and Information Service (2012). "AVHRR." Retrieved 18th Sept,, 2012.

Nusser, S. M. and E. E. Klaas (2003). "Survey methods for assessing land cover map accuracy." Environmental and Ecological Statistics **10**: 309-331.

Rulequest, R. (2009). See5. St Ives NSW, Australia.

Stehman, S. V. and R. L. Czaplewski (1998). "Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles." Remote Sensing of Environment **64**: 331-344.

Strahler, A. H., L. Boschetti, et al. (2006). Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps. GOFC-GOLD Report No. 25. Italy, European Commission,.

Wulder, M. A., J. C. White, et al. (2007). "Validation of a large area land cover product using purpose-acquired airborne video." Remote Sensing of Environment **106**(4): 480-491.

Xie, Y., Z. Sha, et al. (2008). "Remote sensing imagery in vegetation mapping: a review." Jounal of Plant Ecology **1**(1): 9-23.